

# Rotation-Invariant Texture Classification Using a Complete Space-Frequency Model

George M. Haley and B. S. Manjunath, *Member, IEEE*

**Abstract**—A method of rotation-invariant texture classification based on a complete space-frequency model is introduced. A polar, analytic form of a two-dimensional (2-D) Gabor wavelet is developed, and a multiresolution family of these wavelets is used to compute information-conserving *microfeatures*. From these *microfeatures* a *micromodel*, which characterizes spatially localized amplitude, frequency, and directional behavior of the texture, is formed. The essential characteristics of a texture sample, its *macrofeatures*, are derived from the estimated selected parameters of the micromodel. Classification of texture samples is based on the *macromodel* derived from a rotation invariant subset of *macrofeatures*. In experiments, comparatively high correct classification rates were obtained using large sample sets.

**Index Terms**—Gabor filters, texture classification, wavelets.

## I. INTRODUCTION

THE SPECTRUM of texture analysis techniques ranges from those focusing on structural features to those emphasizing statistical modeling. In most statistically oriented techniques within the last 15 years [6], [11], [15], [16], [20], the image is modeled as a Markov random field (MRF) of pixels. In these approaches, the relationships between the intensities of neighboring pixels are statistically characterized. These methods have proven very effective for texture segmentation and classification. More recently, feature-based approaches have been introduced. Features are typically extracted using Gabor functions [3], [4], [7], [24], [26], [28] or wavelet basis functions [8], [12], [23]. Feature-based methods are often less computationally intensive and more effective than MRF-based approaches.

The general approach to developing rotation-invariant techniques has been to modify successful nonrotation-invariant techniques. Since general MRF models are inherently dependent on rotation, several methods have been introduced to obtain rotation invariance. Kashyap and Khotanzad [21] developed the “circular autoregressive” model with parameters that are invariant to image rotation. Choe and Kashyap [9] introduced an autoregressive fractional difference model that has rotation- (as well as tilt- and slant-) invariant parameters. Cohen *et al.* [10] extended a likelihood function to incorporate

rotation (and scale) parameters. To classify a sample, an estimate of its rotation (and scale) is required.

For feature-based approaches, rotation-invariance is achieved by using anisotropic features. Porat and Zeevi [27] use first- and second-order statistics based upon three spatially localized features, two of which (dominant spatial frequency and orientation of dominant spatial frequency) are derived from a Gabor-filtered image. Leung and Peterson [22] present two approaches, one that transforms a Gabor-filtered image into rotation-invariant features and the other of which rotates the image before filtering; however, neither utilizes the spatial resolving capabilities of the Gabor filter. You and Cohen [29] use filters that are tuned over a training set to provide high discrimination among its constituent textures. Greenspan *et al.* [17] and Haley and Manjunath [19] use rotation-invariant structural features obtained via multiresolution Gabor filtering. In both of those approaches, rotation invariance is achieved by using the magnitude of a discrete Fourier transform (DFT) in the rotation dimension. Haley and Manjunath also use statistical features that are rotation invariant.

The approach defined herein is novel in the following respects:

- the basis for classification is a complete feature space model;
- Rotation invariance is achieved by transforming Gabor features into rotation invariant features (using autocorrelation and DFT magnitudes) and by utilizing rotation invariant statistics of rotation dependent features;
- A polar form of a two-dimensional (2-D) Gabor function that is truly analytic (frequency causal) is introduced.

The motivation for the complete feature space model is to exploit the benefits of both model-based and feature-based approaches. The underlying assumption is that conserving information (spatial, spectral and directional characteristics) will result in maximizing classification performance. For classification purposes, “information” is defined as data useful in discriminating between different texture types. Unless the set of texture types is known and fixed, all feature data is potentially important for discrimination. To ensure that the method is as robust as possible, minimal data is discarded.

To conserve information in feature space, an invertible transform is required. While there are several viable options, including orthogonal wavelet transforms, Gabor wavelets were chosen for their desirable properties, as follows.

- Gabor functions achieve the theoretical minimum space-frequency bandwidth product [13], [14], [18], i.e., spatial resolution is maximized for a given bandwidth.

Manuscript received July 23, 1996; revised November 5, 1997. This work was supported in part by the National Science Foundation under Grant 94-1130 and by IRI under Grant 9704785. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. C.-C. Jay Kuo.

G. M. Haley is with Ameritech, Hoffman Estates, IL 60169 USA (e-mail: george.m.haley@ameritech.com).

B. S. Manjunath is with the Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106-9560 USA (e-mail: manj@ece.ucsb.edu).

Publisher Item Identifier S 1057-7149(99)00968-9.

- Gabor functions are used as (nonorthogonal) basis functions for exact signal representation.
- A narrowband Gabor function closely approximates an analytic function. Signals convolved with an analytic function are also analytic, allowing separate analysis of the magnitude (envelope) and phase characteristics in the spatial domain.
- The magnitude response of a Gabor function in the frequency domain is well-behaved, having no sidelobes.
- Gabor functions appear to share many properties with the human visual system [25].

While Gabor functions are a good choice, the standard forms can be further improved. Under certain conditions, very low frequency effects (e.g., due to illumination and shading variations) can cause a significant response in a Gabor filter, leading to misclassification. An analytic form is introduced to minimize these undesirable effects. In the flower petal configuration of the 2-D Gabor function, the polar form allows for superior frequency domain coverage, improves rotation-invariance and simplifies analysis, compared to the standard 2-D form.

The method for establishing a model is as follows.

- 1) Transform image samples into Gabor space.
- 2) Transform the Gabor space samples into a *microfeature* space that completely segregates the rotation dependent information from the rotation independent.
- 3) Estimate parameters of the microfeature-based *micro-model*, which characterizes spatially localized amplitude, frequency and directional behavior of the texture in microfeature space. A sample's micromodel parameters, its *macrofeatures*, provide a more global description of the sample.
- 4) Estimate parameters of the macrofeature-based *macro-model*.

Classification of texture samples is based on the rotation-invariant components of the macromodel. Section II provides a review of Gabor space analysis and presents the true analytic and 2-D polar forms of the Gabor function. Section III explains the transformation of the Gabor space sample into microfeatures. In Section IV, the micromodel and macromodel are developed. Experimental results are presented in Section V and conclusions, in Section VI.

## II. GABOR FUNCTIONS

### A. One-Dimensional Gabor Function

A Gabor function is the product of a Gaussian function and a complex sinusoid. Its general one-dimensional (1-D) form is

$$g_S(x, \omega_C, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \cdot \exp\left(\frac{-x^2}{2\sigma^2}\right) \cdot \exp(j\omega_C x) \quad (1)$$

$$G_S(\omega, \omega_C, \sigma) = \exp\left[\frac{-\sigma^2(\omega - \omega_C)^2}{2}\right]. \quad (2)$$

Thus, Gabor functions are bandpass filters. Gabor functions are used as complete, albeit nonorthogonal, basis sets. It has

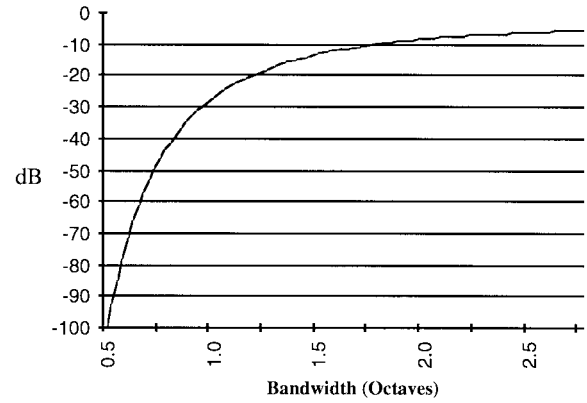


Fig. 1. Response of Gabor function at  $\omega = 0$  versus bandwidth.

been shown that a function  $i(x)$  is represented exactly [18] as

$$i(x) = \sum_{n=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \beta_{n,k} \cdot h_{n,k}(x) \quad (3)$$

where  $h_{n,k}(x) = g_S(x - nX, k\Omega, \sigma)$ , and  $\sigma$ ,  $X$ , and  $\Omega$  are all parameters and  $X\Omega = 2\pi$ .

### B. An Analytic Gabor Function

$G_S(\omega, \omega_C, \sigma)$  exhibits a potentially significant response at  $\omega = 0$  and at very low frequencies. This manifests itself as an undesirable response to interimage and intrainage variations in contrast and intensity due to factors unrelated to the texture itself, potentially causing misclassification. Cases include:

- sample images of a texture with differences in average intensity;
- images with texture regions having differences in contrast and/or intensity (Bovik [3] has demonstrated that region boundaries defined in segmentation using unmodified Gabor filters vary according to these differences between the regions);
- images with uneven illumination.

The response to a constant-valued input (i.e.,  $\omega = 0$ ) relative to the response to an input of equal magnitude at  $\omega = \omega_C$  can be computed as a function of octave bandwidth [3]:

$$|G_S(0)|/|G_S(\omega_C)| = 2^{-\gamma} \quad (4)$$

where  $\gamma = (2^B + 1)/(2^B - 1)$  and  $B = \log_2((\omega_C + \delta)/(\omega_C - \delta))$  and  $\delta$  is the half bandwidth. This plot is shown in Fig. 1. It is interesting to note that the response at  $\omega = 0$  depends upon the  $B$  but not  $\omega_C$ .

There are two approaches to avoiding these problems: preprocessing the image or modifying the Gabor function. Normalizing each image to have a standard average intensity and contrast corrects for interimage, but not intrainage, variations. Alternative methods of image preprocessing are required to compensate for intrainage variations, such as point logarithmic processing [3] or local normalization.

An equally effective and more straightforward approach is to modify the Gabor function to be analytic<sup>1</sup> by forcing the real

<sup>1</sup>Since  $G_S(\omega) \neq 0$  for  $\omega \leq 0$ , a Gabor function only approximates an analytic function.

and imaginary parts to become a Hilbert transform pair. This is accomplished by replacing the real part of  $g_S(x)$ ,  $g_{S\text{Re}}(x)$  with the inverse Hilbert transform of the imaginary part  $-\hat{g}_{S\text{Im}}(x)$ :

$$g_A(x) = -\hat{g}_{S\text{Im}}(x) + jg_{S\text{Im}}(x). \quad (5)$$

The Fourier transforms of the real and imaginary parts of  $g_S(x)$  are, respectively, conjugate symmetric and conjugate antisymmetric, resulting in cancellation for  $\omega \leq 0$ :

$$G_A(\omega) = \begin{cases} G_S(\omega) - G_S^*(-\omega), & \omega > 0 \\ 0, & \omega \leq 0. \end{cases} \quad (6)$$

Because it is analytic,  $G_A(\omega)$  possesses several advantages over for  $G_S(\omega)$  for many applications including texture analysis:

- improved low frequency response since  $|G_A(\omega)| < |G_S(\omega)|$  for small  $\omega$  and  $|G_A(0)| = 0$ ;
- simplified frequency domain analysis since  $G_A(\omega) = 0$  for  $\omega \leq 0$ ;
- reduced frequency domain computations since  $G_A(\omega) = 0$  for  $\omega \leq 0$ ;

These advantages are achieved without requiring additional processing. Thus, it is an attractive alternative for most texture analysis applications.

### C. Two-Dimensional Gabor Function: Cartesian Form

The Gabor function is extended into two dimensions as follows. In the spatial frequency domain, the Cartesian form is a 2-D Gaussian formed as the product of two 1-D Gaussians from (2):

$$\begin{aligned} G_C(\omega_x, \omega_y, \omega_{Cx'}, \omega_{Cy'}, \theta, \sigma_{x'}, \sigma_{y'}) \\ = G(\omega_{x'}, \omega_{Cx'}, \sigma_{x'}) \cdot G(\omega_{y'}, \omega_{Cy'}, \sigma_{y'}) \end{aligned} \quad (7)$$

where  $\theta$  is the orientation angle of  $G_C$ ,  $x' = x \cos \theta + y \sin \theta$  and  $y' = -x \sin \theta + y \cos \theta$ . In the spatial domain,  $G_C$  is separable into two orthogonal 1-D Gabor functions from (1) that are, respectively, aligned to the  $x'$  and  $y'$  axes:

$$\begin{aligned} g_C(x, y, \omega_{Cx'}, \omega_{Cy'}, \theta, \sigma_{x'}, \sigma_{y'}) \\ = g(x', \omega_{Cx'}, \sigma_{x'}) \cdot g(y', \omega_{Cy'}, \sigma_{y'}). \end{aligned} \quad (8)$$

As in (3), an image is represented exactly [1], [2]<sup>2</sup> as

$$\begin{aligned} i(x, y) = \sum_{n_x=-\infty}^{\infty} \sum_{n_y=-\infty}^{\infty} \sum_{k_x=-\infty}^{\infty} \sum_{k_y=-\infty}^{\infty} \beta_{k_x, k_y, n_x, n_y} \\ \cdot h_{k_x, k_y, n_x, n_y}(x, y) \end{aligned} \quad (9)$$

where  $h_{k_x, k_y, n_x, n_y}(x, y) = g(x - n_x X, k_x \Omega_x, \sigma_x) \cdot g(y - n_y Y, k_y \Omega_y, \sigma_y)$ ;  $\sigma_x, \sigma_y, X, Y, \Omega_x$ , and  $\Omega_y$  are constants; and  $X\Omega_x = Y\Omega_y = 2\pi$ . Approximations to  $\beta_{k_x, k_y, n_x, n_y}$  are obtained by using [25]

$$\begin{aligned} \hat{\beta}_{k_x, k_y, n_x, n_y} \\ = i(x, y) * h_{k_x, k_y, n_x, n_y}(x, y) \approx \beta_{k_x, k_y, n_x, n_y} \end{aligned} \quad (10)$$

provided that the parameters are chosen appropriately.

<sup>2</sup>The proofs in the references are based on the standard, not analytic, form of the Gabor function.

### D. Two-Dimensional Gabor Function: Polar Form

An alternative approach to extending the Gabor function into two dimensions is to form, in the frequency domain, the product of a 1-D analytic Gabor function  $G(\omega)$  (the subscript is omitted to indicate that the concepts are generally applicable to the standard form as well) of radial frequency  $\omega$  and a Gaussian function of orientation  $\theta$ :

$$\begin{aligned} G_P(\omega, \theta, \omega_C, \theta_C, \sigma_\rho, \sigma_\theta) \\ = G(\omega, \omega_C, \sigma_\rho) \cdot \exp \left[ \frac{-\sigma_\theta^2 (\theta - \theta_C)^2}{2} \right] \end{aligned} \quad (11)$$

$$\begin{aligned} g_P(x, y, \omega_C, \theta_C, \sigma_\rho, \sigma_\theta) \\ = \iint G_P \left( \sqrt{\omega_x^2 + \omega_y^2}, \tan^{-1}(\omega_y/\omega_x), \omega_C, \theta_C, \sigma_\rho, \sigma_\theta \right) \\ \cdot \exp[j2\pi(\omega_x x + \omega_y y)] d\omega_x d\omega_y \end{aligned} \quad (12)$$

where  $\omega = \sqrt{\omega_x^2 + \omega_y^2}$  and  $\tan(\theta) = \omega_y/\omega_x$ . Thus, (11) is a 2-D Gaussian in the polar, rather than Cartesian, spatial frequency domain. The frequency domain regions of both polar and Cartesian forms of Gabor functions are compared in Fig. 2.

In the Cartesian spatial frequency domain, the  $-3$  dB contour of the Cartesian form is an ellipse, while the polar form has a narrower response at low  $\omega$  and a wider response at high  $\omega$ . When arranged as “flower petals” (equally distributed along a circle centered at the origin), the polar form allows for more uniform coverage of the frequency domain, with less overlap at low frequencies and smaller gaps at high frequencies. The polar form is more suited for rotation invariant analysis since the response always varies as a Gaussian with rotation. The Cartesian form varies with rotation in a more complex manner, introducing an obstacle to rotation invariance and complicating analysis.

### E. Multiresolution Representation with Gabor Wavelets

The Gabor function is used as the basis for generating a wavelet family for multiresolution analysis. Wavelets have two salient properties: the octave bandwidth  $B$  and the octave spacing  $\Delta = \log_2(\omega_{s+1}/\omega_s)$  are both constant, where  $\omega_s$  is the center frequency. The filter spacing is achieved by defining

$$\omega_s = \omega_0 \cdot 2^{-s\Delta}, \quad s \in \{0, 1, 2, \dots\} \quad (13)$$

where  $\omega_0$  is the highest frequency in the wavelet family. Constant bandwidth requires that  $\sigma_\rho$  be inversely proportional to  $\omega_s$ :

$$\sigma_{\rho_s} = \frac{1}{\kappa \omega_s} \quad (14)$$

where

$$\kappa = \frac{2B - 1}{\sqrt{2} \ln 2 \cdot (2B + 1)}$$

is a constant. The orientations of the wavelets are defined as

$$\theta_r = \theta_0 + \frac{2\pi r}{R} \quad (15)$$

where  $\theta_0$  is the starting angle, the second term is the angular increment, and  $r$  and  $R$  are both integers such that  $0 \leq r < R$ .

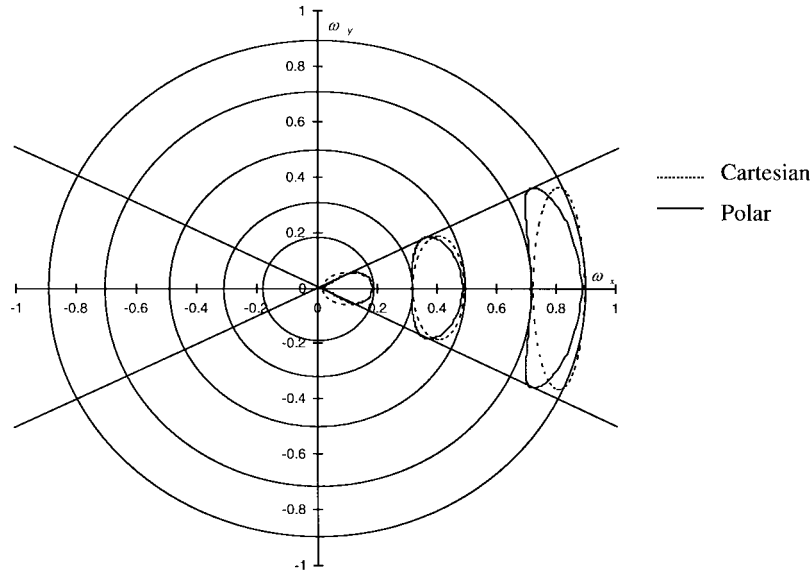


Fig. 2. The  $-3$  dB contours of Cartesian and polar Gabor functions of varying bandwidths. The angular  $-3$  dB width of the polar Gabor functions is  $45^\circ$ .

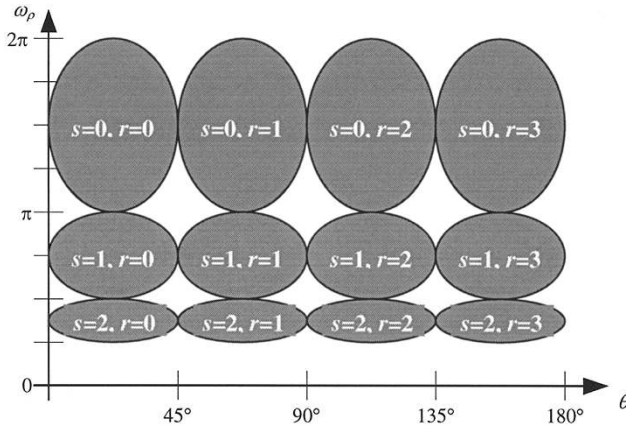


Fig. 3. A family of 2-D polar form Gabor wavelets in the polar spatial frequency domain, with  $S = 3$ ,  $R = 8$ ,  $\omega_0 = 3\pi/2$ , and  $\theta_0 = 22.5^\circ$ . Regions of  $\geq -3$  dB response is indicated by shading. The scale ( $s$ ) and orientation ( $r$ ) indexes are indicated inside each filter's response region.

Using (13)–(15) in (11), the 2-D Gabor wavelet family is defined as

$$\begin{aligned} H_{s,r}(\omega_x, \omega_y) &= G_P\left(\sqrt{\omega_x^2 + \omega_y^2}, \tan^{-1}(\omega_y/\omega_x), \omega_s, \theta_r, \frac{1}{\kappa\omega_s}, \sigma_\theta\right) \\ &= G\left(\sqrt{\omega_x^2 + \omega_y^2}, \omega_s, \frac{1}{\kappa\omega_s}\right) \\ &\quad \cdot \exp\left[\frac{-\sigma_\theta^2(\tan^{-1}(\omega_y/\omega_x) - \theta_r)^2}{2}\right] \end{aligned} \quad (16)$$

$$\begin{aligned} h_{s,r,n_x,n_y}(x, y) &= g_P\left(x - n_x X_s, y - n_y Y_s, \omega_s, \theta_r, \frac{1}{\kappa\omega_s}, \sigma_\theta\right) \end{aligned} \quad (17)$$

where  $X_s$  and  $Y_s$ , the sampling intervals, are inversely proportional to the bandwidths corresponding to  $s$ . Fig. 3 depicts an example wavelet family of this form.

As in (9), an image is represented using the polar wavelet form of the Gabor function from (17):

$$i(x, y) = \sum_{n_x=-\infty}^{\infty} \sum_{n_y=-\infty}^{\infty} \sum_{s=0}^{\infty} \sum_{r=0}^{R-1} \beta_{s,r,n_x,n_y} \cdot h_{s,r,n_x,n_y}(x, y). \quad (18)$$

Approximations to  $\beta_{s,r,n_x,n_y}$  are obtained as in (10):

$$\hat{\beta}_{s,r,n_x,n_y} = i(x, y) * h_{s,r,n_x,n_y}(x, y) \approx \beta_{s,r,n_x,n_y} \quad (19)$$

and parameters  $X_s$ ,  $Y_s$ ,  $\omega_0$ ,  $\kappa$ , and  $\sigma_\theta$  are chosen appropriately. Instead of a rectangular lattice, a polar Gabor wavelet representation has the shape of a cone.

### III. MICROFEATURE REPRESENTATION

#### A. Transformation into Gabor Space

As described in Section II, a set of 2-D Gabor wavelets can represent an image. Assuming that the image is spatially limited to  $0 \leq x < N_x X_s$ ,  $0 \leq y < N_y Y_s$ , where  $N_x$  and  $N_y$  represent the number of samples in their respective dimensions, and is bandlimited to  $0 < \omega \leq \omega_H$ ,<sup>3</sup> the number of Gabor wavelets needed to represent the image is finite. Substituting  $\hat{\beta}_{s,r,n_x,n_y}$  from (19) for  $\beta_{s,r,n_x,n_y}$  in (18), a texture image is approximately represented using the polar wavelet form of the Gabor function as

$$i(x, y) \approx \sum_{n_x=0}^{N_{sx}-1} \sum_{n_y=0}^{N_{sy}-1} \sum_{s=0}^{S-1} \sum_{r=0}^{R-1} \hat{\beta}_{s,r,n_x,n_y} \cdot h_{s,r,n_x,n_y}(x, y) \quad (20)$$

<sup>3</sup>For sampled texture images, the upper frequency bound is enforced, although aliasing may be present since natural textures are generally not bandlimited. It is both reasonable and convenient to assume that, for textures of interest, a lower frequency bound  $\omega_L > 0$  exists below which there is no useful discriminatory information.

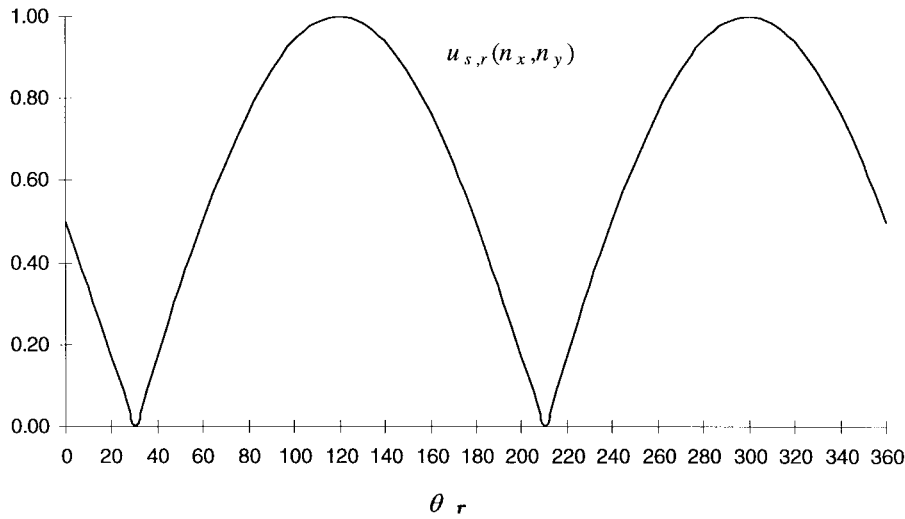


Fig. 4.  $u_{s,r}(n_x, n_y)$  in response to a sinusoidal input texture oriented at  $120^\circ$ .

where parameters  $S, R, X_s, Y_s, \omega_0, \kappa$ , and  $\sigma_\theta$  are chosen appropriately. Note that the  $s$  subscript is added to  $N_x$  and  $N_y$  to indicate their dependencies on  $X_s$  and  $Y_s$ . Thus, a texture image is represented with relatively little information loss by the coefficients  $\hat{\beta}_{s,r,n_x,n_y}$ .

Following Bovik *et al.* [4],  $\hat{\beta}_{s,r,n_x,n_y}$  is interpreted as a channel or band  $b_{s,r}(n_x, n_y)$  of the image  $i(x, y)$  tuned to the carrier frequency  $\omega_s = \omega_0 \cdot 2^{-s\Delta}$  (13), oriented at angle  $\theta_r = \theta_0 + 2\pi r/R$  (15) and sampled in the spatial domain at intervals of  $X_s$  and  $Y_s$ . Since  $b_{s,r}(n_x, n_y)$  is formed by convolution with a narrowband, analytic function (19),  $b_{s,r}(n_x, n_y)$  is also narrowband and analytic, and is therefore decomposable into amplitude and phase components that can be independently analyzed as follows:

$$b_{s,r}(n_x, n_y) = a_{s,r}(n_x, n_y) \cdot \exp(j\psi_{s,r}(n_x, n_y)) \quad (21)$$

where  $a_{s,r}(n_x, n_y) = |b_{s,r}(n_x, n_y)|$  and  $\psi_{s,r}(n_x, n_y) = \arg(b_{s,r}(n_x, n_y))$ .  $a_{s,r}(n_x, n_y)$  contains information about the amplitude and amplitude modulation (AM) characteristics of the texture's periodic features within the band, and  $\psi_{s,r}(n_x, n_y)$  contains information about the phase, frequency and frequency modulation (FM) characteristics. For textures with low AM in band  $(s, r)$ ,  $a_{s,r}(n_x, n_y)$  is approximately constant over  $(n_x, n_y)$ . For textures with low FM in band  $s, r$ , the slope of  $\psi_{s,r}(n_x, n_y)$  with respect to  $(n_x, n_y)$  is nearly constant.

Both  $a_{s,r}(n_x, n_y)$  and  $\psi_{s,r}(n_x, n_y)$  are rotation-dependent and periodic in  $r$  such that

$$a_{s,((r+R/2))_R}(n_x, n_y) = a_{s,r}(n_x, n_y) \quad (22)$$

$$\psi_{s,((r+R))_R}(n_x, n_y) = \psi_{s,r}(n_x, n_y) \quad (23)$$

$$\psi_{s,((r+R/2))_R}(n_x, n_y) = -\psi_{s,r}(n_x, n_y). \quad (24)$$

Rotating  $i(x, y)$  by  $\theta^\circ$  produces a circular shift in  $r$  of  $-R\theta/180^\circ$  for  $a_{s,r}(n_x, n_y)$  and  $-R\theta/360^\circ$  for  $\psi_{s,r}(n_x, n_y)$ .

The characteristics of  $b_{s,r}(n_x, n_y)$  are clearly illustrated by using a sinusoidal input,  $i(x, y) = \cos[\omega \cdot (x \cos \theta + y \sin \theta)]$ . Both AM and FM are nonexistent. The Gabor space representation that follows from

(21) is  $b_{s,r}(n_x, n_y) = a_{s,r} \exp(j\psi_{s,r}(n_x, n_y))$ , where  $\psi_{s,r}(n_x, n_y) = \omega(n_x \cos \theta + n_y \sin \theta)$  and  $a_{s,r}$  is a constant for a given  $s$  and  $r$ . From (16)

$$\begin{aligned} a_{s,r} &= |H_{s,r}(\omega_x, \omega_y)| \\ &= G\left(\omega, \omega_s, \frac{1}{\kappa\omega_s}\right) \cdot \exp\left[\frac{-\sigma_\theta^2(\theta - \theta_r)^2}{2}\right]. \end{aligned}$$

For each  $s$  and  $r$ ,  $\psi_{s,r}(n_x, n_y)$  is a plane in the spatial domain with a slope of  $(\omega_t \cos \theta_t, \omega_t \sin \theta_t)$ .

### B. Local Frequency Estimation

While  $\psi_{s,r}(n_x, n_y)$  contains essential information about a texture, it is not directly usable for classification. However, local frequency information can be extracted from  $\psi_{s,r}(n_x, n_y)$  as follows:

$$\phi_{s,r}(n_x, n_y) = \begin{cases} \theta_\nabla & |\theta_r - \theta_\nabla| \leq 90^\circ \\ \theta_\nabla + 180^\circ & |\theta_r - \theta_\nabla| > 90^\circ \end{cases} \quad (25)$$

and

$$\begin{aligned} u_{s,r}(n_x, n_y) &= \sqrt{\nabla_x^2(\psi_{s,r}(n_x, n_y)) + \nabla_y^2(\psi_{s,r}(n_x, n_y))} \\ &\quad \cdot \cos(\theta_r - \phi_{s,r}(n_x, n_y)) \\ &= \sqrt{\nabla_x^2(\psi_{s,r}(n_x, n_y)) + \nabla_y^2(\psi_{s,r}(n_x, n_y))} \\ &\quad \cdot |\cos(\theta_r - \theta_\nabla)| \end{aligned} \quad (26)$$

where  $\nabla_x()$  and  $\nabla_y()$  are gradient estimation functions,  $\theta_r$  is the orientation of the Gabor function, and  $\theta_\nabla = \tan^{-1}(\nabla_y(\psi_{s,r}(n_x, n_y))/\nabla_x(\psi_{s,r}(n_x, n_y)))$  is the direction of the gradient vector.  $u_{s,r}(n_x, n_y)$  is a spatially localized estimate of the frequency along the direction  $\theta_r$ , and  $\phi_{s,r}(n_x, n_y)$  is the direction of maximal phase change rate, i.e., highest local frequency. For the preceding sinusoidal input example,  $\phi_{s,r}(n_x, n_y) = 120^\circ$  for  $30^\circ \leq \theta_r < 210^\circ$ , and  $\phi_{s,r}(n_x, n_y) = -120^\circ$  when  $\theta_r$  is outside that range. The behavior of  $u_{s,r}(n_x, n_y)$  for the example is shown in Fig. 4.

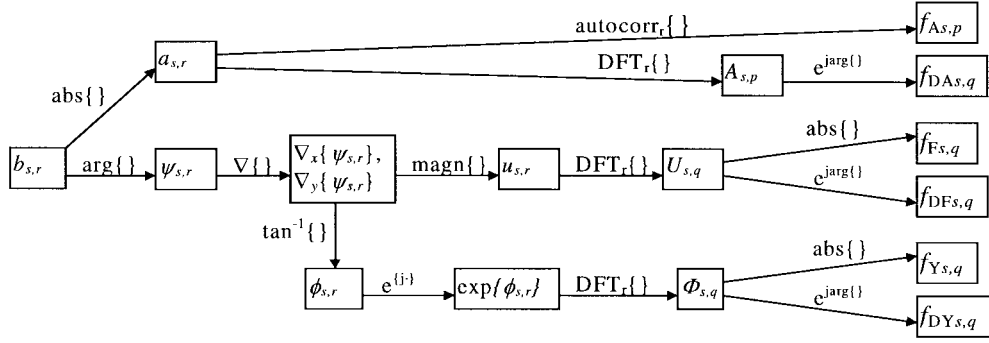


Fig. 5. Transformation from Gabor coefficients into microfeatures.

### C. Transformation into Microfeatures

To facilitate discrimination between textures,  $b_{s,r}(n_x, n_y)$  is further decomposed into microfeatures that contain local amplitude, frequency, phase, direction, and directionality characteristics. This decomposition is depicted in Fig. 5. In the following, for simplicity,  $R$  is assumed to be even. The microfeatures are defined to be

$$f_{As,p}(n_x, n_y) = \sum_{r=0}^{R/2-1} a_{s,r}(n_x, n_y) \cdot a_{s,((r+p)R)}(n_x, n_y), \quad 0 \leq p \leq R/4 \quad (27)$$

$$f_{Fs,q}(n_x, n_y) = \left| \sum_{r=0}^{R/2-1} u_{s,r}(n_x, n_y) \cdot \exp\left(-\frac{2\pi j r q}{R/2}\right) \right|, \quad 0 \leq q \leq R/4 \quad (28)$$

$$f_{Ys,q}(n_x, n_y) = \left| \sum_{r=0}^{R-1} \exp(j\phi_{s,r}(n_x, n_y)) \cdot \exp\left(-\frac{2\pi j r q}{R}\right) \right|, \quad q = 1, 3, \dots, R-1 \quad (29)$$

$$f_{DA s,q}(n_x, n_y) = \arg \left[ \sum_{r=0}^{R/2-1} a_{s,r}(n_x, n_y) \cdot \exp\left(-\frac{2\pi j r q}{R/2}\right) \right], \quad 1 \leq q \leq R/4 \quad (30)$$

$$f_{DF s,q}(n_x, n_y) = \arg \left[ \sum_{r=0}^{R/2-1} u_{s,r}(n_x, n_y) \cdot \exp\left(-\frac{2\pi j r q}{R/2}\right) \right], \quad 1 \leq q \leq R/4 \quad (31)$$

$$f_{DY s,q}(n_x, n_y) = \arg \left[ \sum_{r=0}^{R-1} \exp(j\phi_{s,r}(n_x, n_y)) \cdot \exp\left(-\frac{2\pi j r q}{R}\right) \right], \quad q = 1, 3, \dots, R-1. \quad (32)$$

$f_{As,p}(n_x, n_y)$  contains the amplitude envelope information from  $b_{s,r}(n_x, n_y)$ . Because of the  $R/2$  periodicity of  $a_{s,r}$  (22),

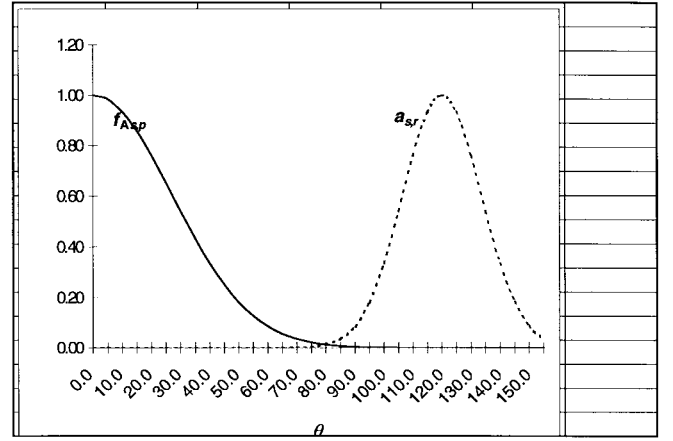


Fig. 6.  $a_{s,r}(n_x, n_y)$  and  $f_{As,p}(n_x, n_y)$  in response to a sinusoidal input texture oriented at  $120^\circ$ .  $f_{As,p}(n_x, n_y)$  does not depend upon the rotation of the input.

only  $R/2$  components are needed in the sum in (27). Eliminating the redundant components from the circular autocorrelation allows complete representation by the  $0 \leq p \leq R/4$  components of  $f_{As,p}(n_x, n_y)$ . It is rotation-invariant because the autocorrelation operation eliminates the dependence on  $r$ , and thus, on  $\theta$ . For a sinusoid input,  $a_{s,r}(n_x, n_y)$  is a Gaussian in  $r$ , and  $f_{As,p}(n_x, n_y)$  is the autocorrelation of a Gaussian in  $r$ , as shown in Fig. 6.

$f_{Fs,q}(n_x, n_y)$  contains the frequency envelope information from  $b_{s,r}(n_x, n_y)$ . Similar to  $a_{s,r}(n_x, n_y)$ ,  $u_{s,r}(n_x, n_y)$  has  $R/2$  periodicity. Since  $u_{s,r}(n_x, n_y)$  is real  $f_{Fs,q}(n_x, n_y)$  is conjugate symmetric in  $q$ , and consequently, its  $0 \leq q \leq R/4$  components are sufficient for complete representation. It is rotation-invariant because the DFT operation maps rotationally induced shifts into the complex numbers' phase components, which are removed when the magnitude operation is performed.

$f_{Ys,q}(n_x, n_y)$  contains the directionality information from  $b_{s,r}(n_x, n_y)$ . Since  $\phi_{s,((r+R/2)R)}(n_x, n_y) = \phi_{s,r}(n_x, n_y) + 180^\circ$ , only the components with odd  $q$  are nonzero. For the same reason as  $f_{Fs,q}(n_x, n_y)$ ,  $f_{Ys,q}(n_x, n_y)$  is rotation-invariant.

$f_{DA s,q}(n_x, n_y)$ ,  $f_{DF s,q}(n_x, n_y)$ , and  $f_{DY s,q}(n_x, n_y)$  contain the direction information from  $b_{s,r}(n_x, n_y)$ . Because  $f_{DA s,q}(n_x, n_y)$  and  $f_{DF s,q}(n_x, n_y)$  are conjugate symmetric



Fig. 7. Textures with similar microfeatures.

in  $q$ , they are represented completely by their  $0 \leq q \leq R/4$  components. However, the  $q = 0$  component is always zero since the DFT's are on real sequences in both cases.  $f_{DYs,q}(n_x, n_y)$  has the same nonzero indexes as  $f_{Ys,q}(n_x, n_y)$ .  $f_{DAs,q}(n_x, n_y)$ ,  $f_{DFs,q}(n_x, n_y)$ , and  $f_{DYs,q}(n_x, n_y)$  are inherently rotation-variant since the phases of the DFT contain all of the direction information.

Since all transformations in this decomposition are invertible (assuming boundary conditions are available), it is possible to exactly reconstruct  $b_{s,r}(n_x, n_y)$  from their microfeatures. Thus,  $f_{As,p}(n_x, n_y)$ ,  $f_{Fs,q}(n_x, n_y)$ ,  $f_{Ys,q}(n_x, n_y)$ ,  $f_{DAs,q}(n_x, n_y)$ ,  $f_{DFs,q}(n_x, n_y)$ , and  $f_{DYs,q}(n_x, n_y)$  provide a nearly exact representation of  $i(x, y)$ .

#### IV. THE TEXTURE MODEL

##### A. The Texture Micromodel

A texture may be modeled as a vector-valued random field  $\mathbf{f} = [\mathbf{f}_A \ \mathbf{f}_F \ \mathbf{f}_Y \ \mathbf{f}_{DA} \ \mathbf{f}_{DF} \ \mathbf{f}_{DY}]^T$ , where  $\mathbf{f}_A, \mathbf{f}_F, \mathbf{f}_Y, \mathbf{f}_{DA}, \mathbf{f}_{DF}$ , and  $\mathbf{f}_{DY}$  are vectors containing the microfeature components for all  $s$  and  $p$  or  $q$  indexes. It is assumed that  $\mathbf{f}$  is stationary and has a multivariate Gaussian distribution. For simplicity, a non-Markovian model is chosen.

Given these assumptions, the *micromodel* for texture  $t$  is stated as

$$p(\mathbf{f}|t) = \frac{1}{\sqrt{(2\pi)^{N_f} |\mathbf{C}_{ft}|}} \cdot \exp\left(\frac{-(\mathbf{f} - \mu_{ft}) \mathbf{C}_{ft}^{-1} (\mathbf{f} - \mu_{ft})^T}{2}\right) \quad (33)$$

where  $\mu_{ft} = E\{\mathbf{f}|t\}$  and  $\mathbf{C}_{ft} = E\{\mathbf{f} \cdot \mathbf{f}^T | t\} - E\{\mathbf{f}|t\} \cdot E\{\mathbf{f}^T | t\}$ , are the mean and covariance of  $\mathbf{f}$ , respectively, and  $N_f$  is the number of microfeatures.

##### B. Macrofeatures

While microfeatures can be used to represent a texture sample, microfeatures are spatially localized and do not characterize global attributes of textures. For instance, consider the textures in Fig. 7. Most of the spatial samples in the upper-right and lower-left quadrants of texture A would be classified as texture B based on microfeatures alone. Furthermore,  $\mathbf{f}_{DA}, \mathbf{f}_{DF}$ , and  $\mathbf{f}_{DY}$  are rotation-dependent, making them unsuitable for rotation-invariant classification.

For classification, a better texture model is derived from the micromodel parameters,  $\mu_{ft}$  and  $\mathbf{C}_{ft}$ . For instance, for the two textures shown in Fig. 7, the standard deviations of

$\mathbf{f}_{DA}, \mathbf{f}_{DF}$ , and  $\mathbf{f}_{DY}$  provide excellent discrimination information not available in the microfeatures themselves. A texture  $t$ 's *macrofeatures* are defined to be

$$\mathbf{F} = \begin{bmatrix} F_{CA} \\ F_{CF} \\ F_{CY} \\ F_{AM} \\ F_{FM} \\ F_{YM} \\ F_{DMA} \\ F_{DMF} \\ F_{DMY} \end{bmatrix} \begin{bmatrix} E\{\mathbf{f}_A|t\} \\ E\{\mathbf{f}_F|t\} \\ E\{\mathbf{f}_Y|t\} \\ \sqrt{E\{\mathbf{f}_A^2\} - E\{\mathbf{f}_A\}^2} \\ \sqrt{E\{\mathbf{f}_F^2\} - E\{\mathbf{f}_F\}^2} \\ \sqrt{E\{\mathbf{f}_Y^2\} - E\{\mathbf{f}_Y\}^2} \\ \sqrt{E\{\mathbf{f}_{DA}^2\} - E\{\mathbf{f}_{DA}\} \times E\{\mathbf{f}_{DA}\}^*} \\ \sqrt{E\{\mathbf{f}_{DF}^2\} - E\{\mathbf{f}_{DF}\} \times E\{\mathbf{f}_{DF}\}^*} \\ \sqrt{E\{\mathbf{f}_{DY}^2\} - E\{\mathbf{f}_{DY}\} \times E\{\mathbf{f}_{DY}\}^*} \end{bmatrix} \quad (34)$$

where  $\mathbf{f}^2 = \mathbf{f} \times \mathbf{f} = [f_1 \cdot f_1 \ f_2 \cdot f_2 \dots]^T$ . For texture  $t$ ,  $F_{CA}, F_{CF}$ , and  $F_{CY}$  describe amplitude, frequency and directionality characteristics, respectively, of the “carrier.”  $F_{AM}, F_{FM}$ , and  $F_{YM}$  describe a texture’s amplitude modulation, frequency modulation, and directionality modulation characteristics, respectively.  $F_{CA}, F_{CF}, F_{CY}, F_{AM}, F_{FM}$ , and  $F_{YM}$  are all rotation-invariant because the microfeatures upon which they are based are rotation-invariant.  $F_{DMA}, F_{DMF}$ , and  $F_{DMY}$  capture the directional modulation characteristics. While  $\mathbf{f}_{DA}, \mathbf{f}_{DF}$ , and  $\mathbf{f}_{DY}$  are rotation-dependent, their variances are not. Means of  $\mathbf{f}_{DA}, \mathbf{f}_{DF}$ , and  $\mathbf{f}_{DY}$  are directional in nature and are not used as classification features. For simplicity, off-diagonal covariance terms are not used, although they may contain useful information.

##### C. The Texture Macromodel

For purposes of classification, a texture  $t$  is modeled as a vector-valued Gaussian random vector  $\mathbf{F}$  with the conditional probability density function

$$p(\hat{\mathbf{F}}|t) = \frac{1}{\sqrt{(2\pi)^{N_F} |\mathbf{C}_{ft}|}} \cdot \exp\left(\frac{-(\hat{\mathbf{F}} - \mu_{ft}) \mathbf{C}_{ft}^{-1} (\hat{\mathbf{F}} - \mu_{ft})^T}{2}\right) \quad (35)$$

where  $\mu_{ft} = E\{\mathbf{F}|t\}$  and  $\mathbf{C}_{ft} = E\{\mathbf{F} \cdot \mathbf{F}^T | t\} - E\{\mathbf{F}|t\} \cdot E\{\mathbf{F}^T | t\}$  are the mean and covariance of  $\mathbf{F}$ , respectively,  $N_F$  is the number of macrofeatures, and  $\hat{\mathbf{F}}$  is an estimate of  $\mathbf{F}$  based on a sample of texture  $t$ . This is the texture *macromodel*.

The parameters  $\mu_{ft}$  and  $\mathbf{C}_{ft}$  are estimated from statistics over  $M$  samples for each texture  $t$ :

$$\hat{\mu}_{ft} = \frac{1}{M} \sum_{m=1}^M \hat{\mathbf{F}}_m$$

and

$$\hat{\mathbf{C}}_{ft} = \frac{1}{M-1} \sum_{m=1}^M (\hat{\mathbf{F}}_m - \hat{\mu}_{ft}) \cdot (\hat{\mathbf{F}}_m - \hat{\mu}_{ft})^T \quad (36)$$

where  $\hat{\mathbf{F}}_m$  is the estimate of  $\mathbf{F}$  based on sample  $m$  of texture  $t$ .

#### V. EXPERIMENTAL RESULTS

Experiments were performed on two groups of textures. The first group comprises 13 texture images [30] digitized from the Brodatz album [5] and other sources. Each texture was

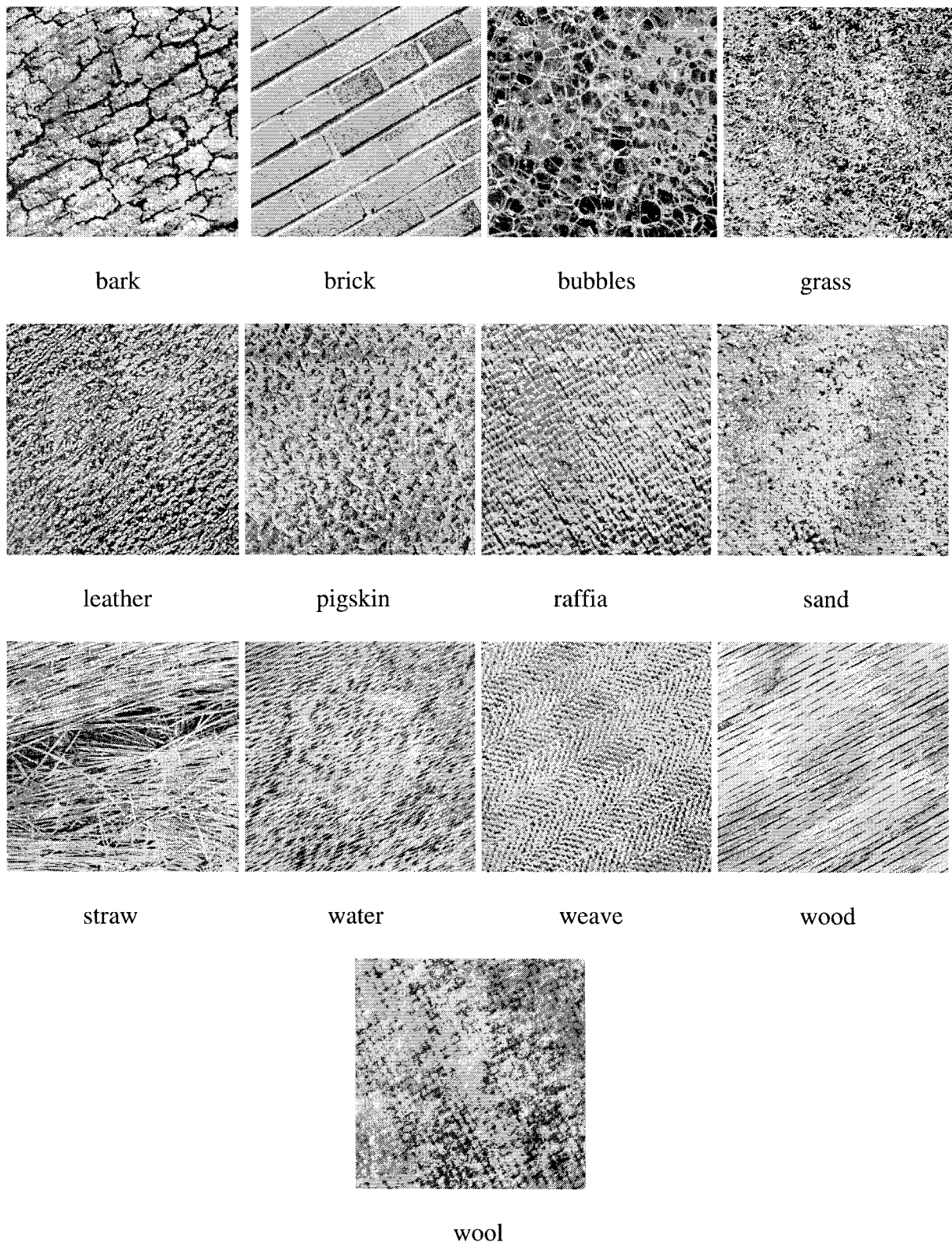


Fig. 8. Textures from the first group. Each texture was digitized at rotations of  $0^\circ$ ,  $20^\circ$ ,  $60^\circ$ ,  $90^\circ$ ,  $120^\circ$ , and  $150^\circ$ . Table I summarizes the results for rotation invariant classification for these textures.

digitized at rotations of  $0^\circ$ ,  $30^\circ$ ,  $60^\circ$ ,  $90^\circ$ ,  $120^\circ$ , and  $150^\circ$  as  $512 \times 512$  pixels, each of which was then subdivided into  $16 \times 16 \times 16 \times 16 \times 16 \times 16$  subimages. Fig. 8 present the  $120^\circ$  rotations of these images. The second group comprises 109 texture images from the Brodatz album digitized at  $0^\circ$  with  $512 \times 512$  pixels at 300 dpi resolution, each of which was then

subdivided into  $16 \times 16 \times 16 \times 16 \times 16 \times 16$  subimages. A polar, analytic Gabor transform was used with parameter values  $\omega_0 = 0.8\pi$ ,  $\theta_0 = 0^\circ$ ,  $S = 4$ ,  $R = 16$ ,  $\kappa = 0.283$  ( $B = 1$  octave) and  $\sigma_\theta = 0.0523/^\circ$  ( $-3$  dB width of  $90^\circ$ ).

Two types of feature analysis were performed for the group of 13 textures. Feature parametric analysis characterizes



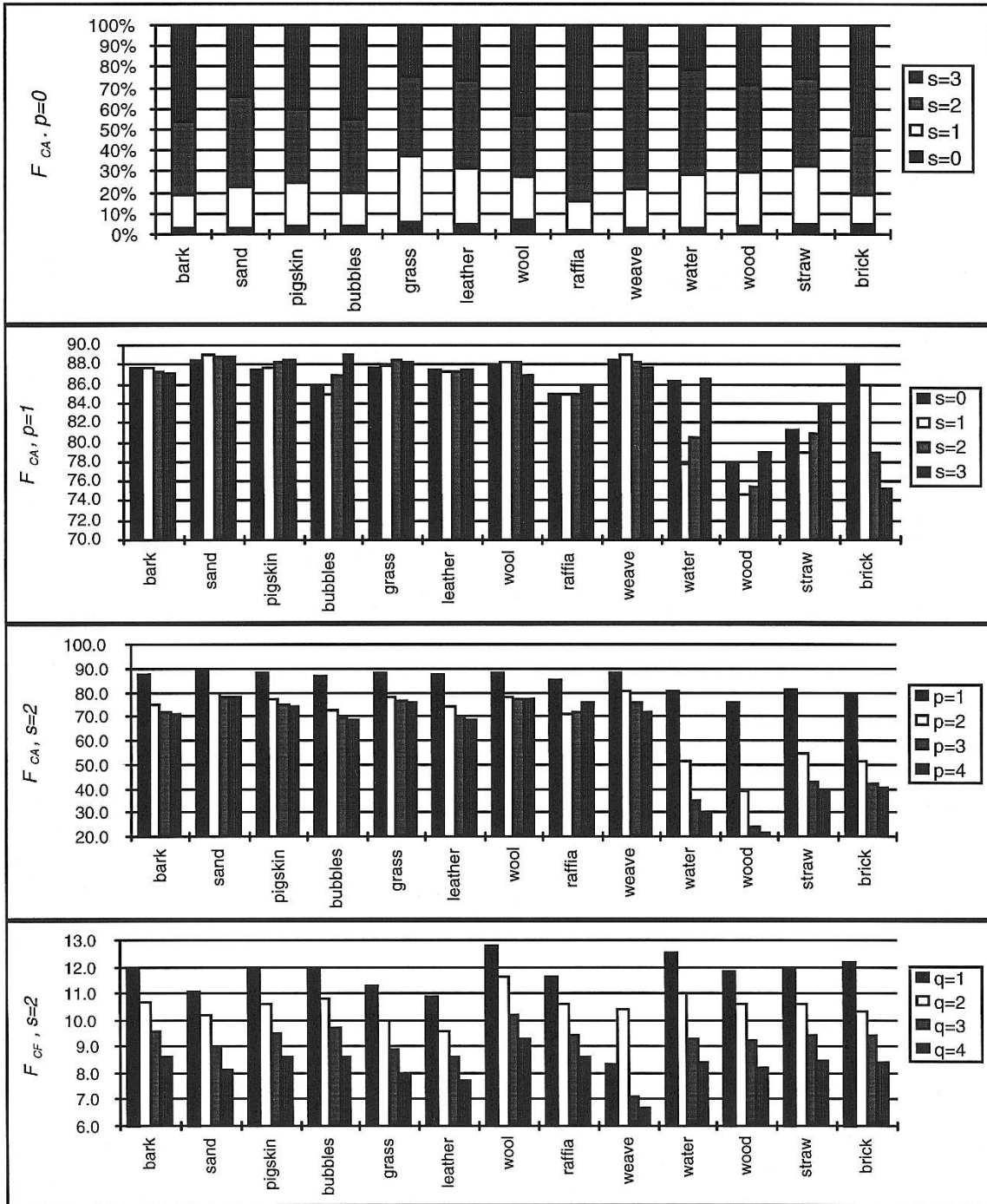


Fig. 9. Feature parametric analysis for  $F_{CA}$  ( $p = 0$ ,  $p = 1$ , and  $s = 2$ ); and for  $F_{CF}$  ( $s = 2$ ). The presentation format for  $F_{CA}$  ( $p = 0$ ) is unique since the features are normalized to have a sum of 1.0 for each texture.

feature types by the scale ( $s$ ), rotation frequency ( $q$ ) and/or autocorrelation index ( $r$ ). Quality analysis focuses on the discrimination capabilities of a single feature across all textures. Many of the feature values have been normalized for improved presentation.

Classification performance was demonstrated with both groups of textures. Half of the subimages (separated in a checkerboard pattern) were used to estimate the model parameters (mean and covariance) for each type of texture, while the other half were used as test samples. Features were extracted from all of the subimages in an identical manner.

To reduce filter sampling effects at high frequencies due to rotation, the estimation of model parameters was based on the features from subimages at all rotations in the first group of images.

#### A. Feature Parametric Analysis

Recall that  $F_{CA}$  represents the amplitude envelope information. Referring to Fig. 9, at  $p = 0$ ,  $F_{CA}$  is particularly high for weave at  $s = 2$  ( $> 60\%$ ) and brick at  $s = 3$  ( $> 50\%$ ). Visually distinct textures such as straw and leather have similar  $p = 0$  signatures. For  $p > 0$ , the most directional textures (water,

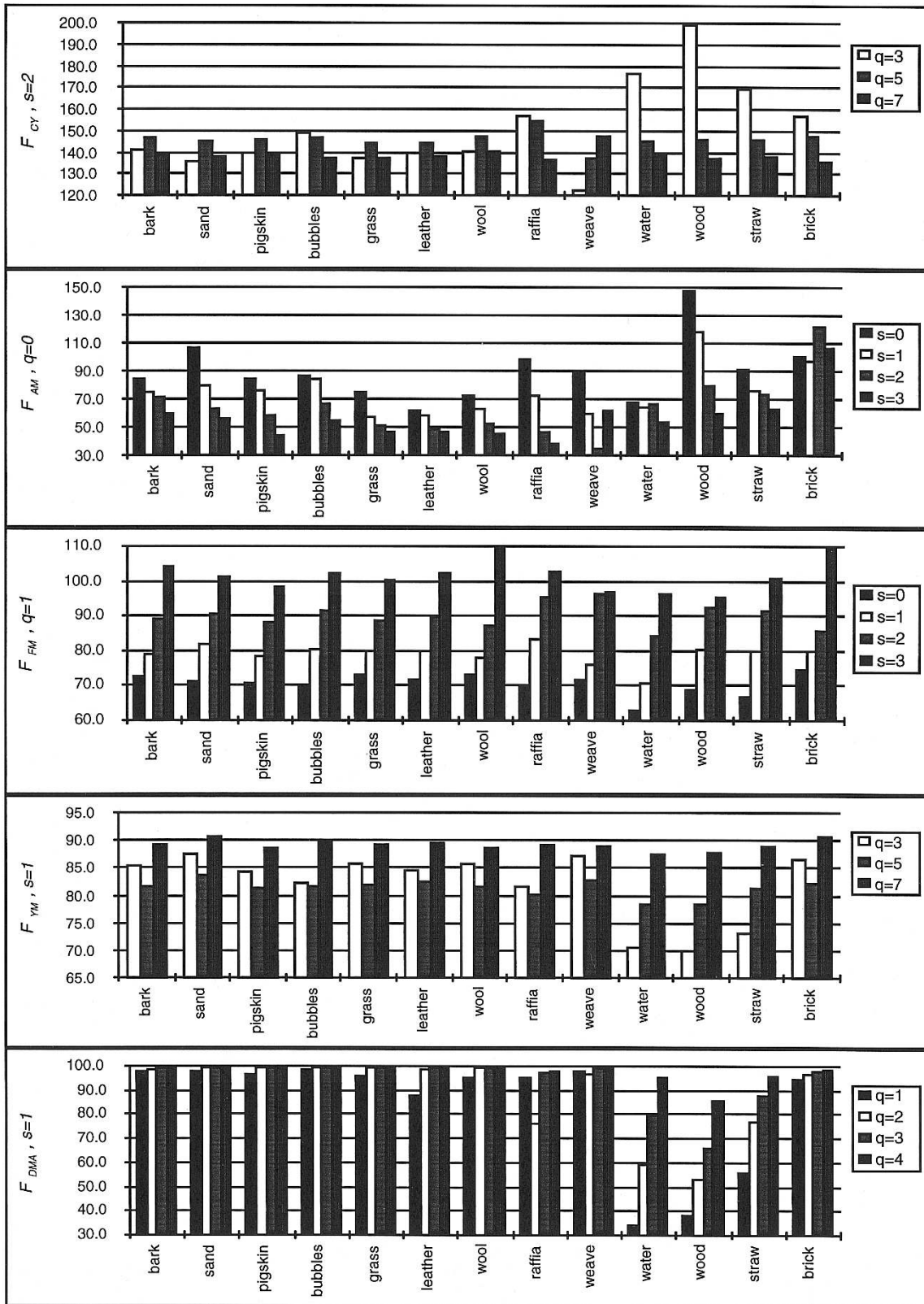


Fig. 10. Feature parametric analysis for  $F_{CY}$  ( $s = 2$ );  $F_{AM}$  ( $q = 0$ );  $F_{FM}$  ( $q = 1$ );  $F_{YM}$  ( $s = 1$ ); and  $F_{DMA}$  ( $s = 1$ ).

wood, straw, and brick) had the lowest values, especially at higher values of  $p$ . For all textures,  $F_{CA}$  generally decreased with increasing  $p$ , except for raffia at  $s = 2$ , which had two maxima ( $p = 0$  and  $p = 4$ ). The behavior of raffia is due to its bidirectional structure. The Gabor response of raffia has

two peaks separated by  $90^\circ$ , manifested as two maxima in the autocorrelation ( $p = 4$  corresponds to  $90^\circ$ ).

For  $F_{CF}$ , which represents the frequency envelope information from  $b_{s,r}$ , most textures had values that decreased with increasing  $q$ . Notable exceptions are wool and weave.

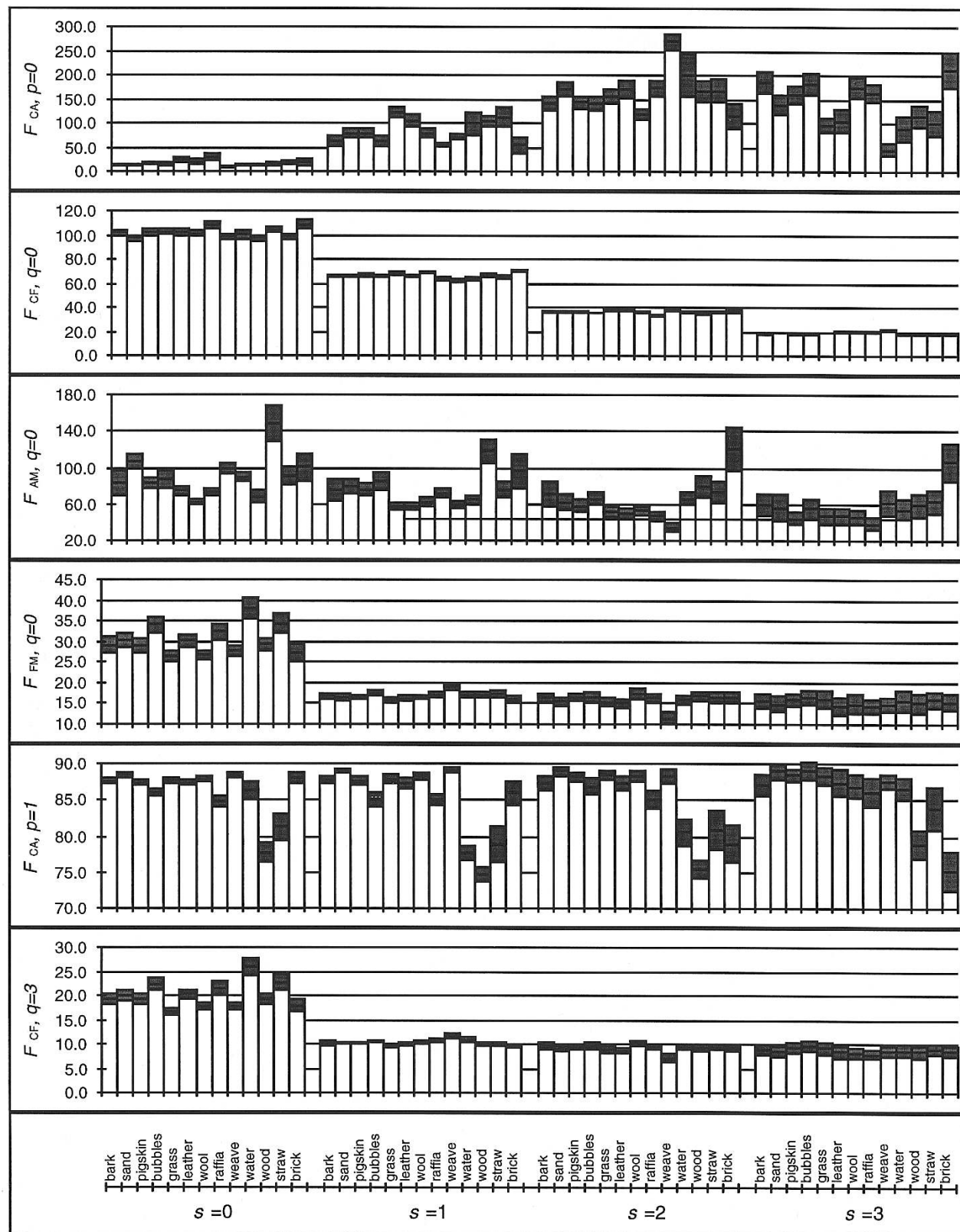


Fig. 11. Feature quality for carrier amplitude and carrier frequency characteristics,  $F_{CA}$  ( $p = 0$ ) and  $F_{CF}$  ( $q = 0$ ). Amplitude modulation and frequency modulation characteristics,  $F_{AM}$  ( $q = 0$ ) and  $F_{FM}$  ( $q = 0$ ). Feature quality for carrier directionality characteristics,  $F_{CA}$  ( $p = 1$ ) and  $F_{CF}$  ( $q = 3$ ).

The features for wool increased from having relatively low values at  $s = 0$  to having the highest values at  $s = 2$ . At  $s = 3$ , wool had the lowest value for  $q = 1$  but a relatively high value for  $q = 2$ . Weave had relatively low

values at  $s = 0$  but the highest values at  $s = 1$ . At  $s = 2$ , weave exhibited a very distinctive signature, with the lowest values for all  $q$  except  $q = 2$ , which had an average value.

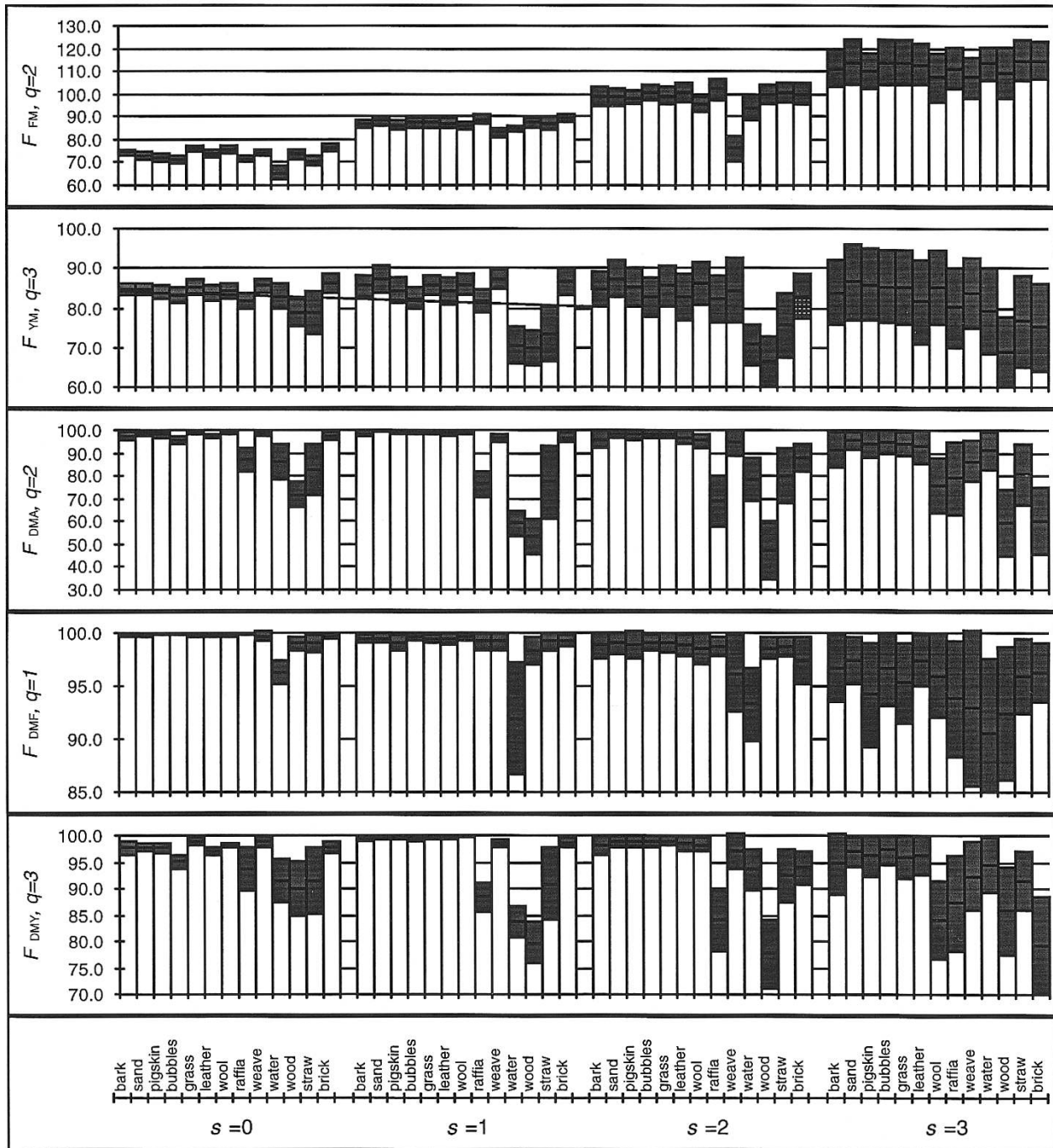


Fig. 12. Feature quality for carrier directionality  $F_{CY}$  ( $q = 3$ ), and directionality modulation characteristics,  $F_{FM}$  ( $q = 2$ ),  $F_{YM}$  ( $q = 3$ ),  $F_{DMA}$  ( $q = 2$ ),  $F_{DMF}$  ( $q = 1$ ), and  $F_{DMY}$  ( $q = 3$ ).

Referring to Fig. 10, for  $F_{CY}$ , higher the values corresponded to stronger directionality. Strongly unidirectional textures had values that peaked at  $q = 3$  and decreased rapidly with increasing  $q$ . Textures with low directionality tended to peak at  $q = 5$ . Raffia, bubbles and weave had distinctive signatures, with raffia and bubbles having relatively high values at both  $q = 3$  and  $q = 5$ , and weave increasing with  $q$  (at  $s = 2$ ) instead of decreasing or peaking at  $q = 5$ .

For  $F_{AM}$ , the most amplitude-uniform textures, raffia and weave, had the lowest values overall. The differences in intensity levels between bricks is reflected in its relatively high

values, particularly at low frequencies (higher  $s$ ). The large variations in the strength of the wood grain patterns resulted in high values for higher frequencies. Weave and brick had unique signatures for  $q = 0$ .

For  $F_{FM}$ , the multidirectional textures weave and raffia (for higher frequencies) and wool and brick (for lower frequencies) had the highest values for  $q = 1$ . Weave had a particularly complex variation over  $q$  and  $s$ .

For  $F_{YM}$ , the textures with a single predominant orientation monotonically increased with  $q$ ; all others had a minimum at  $q = 5$  and a maximum at  $q = 7$ .

TABLE I  
CLASSIFICATION PERFORMANCE FOR FIRST GROUP OF TEXTURES

Sample Type	% Classified Correctly
Bark	87.5%
Sand	97.9%
Pigskin	95.8%
Bubbles	100%
Grass	95.8%
Leather	93.8%
Wool	91.7%
Raffia	100%
Weave	100%
Water	97.9%
Wood	97.9%
Straw	100%
Brick	100%

For  $F_{DMA}$ , the textures with the most consistent direction, water and wood, had the lowest values. The value for leather for  $q = 1$  was significantly lower than that for grass, one of the few features distinguishing the two. Textures generally increased with  $q$  except for multidirectional textures (wool, raffia, brick), for which the values at  $q = 1$  were the lowest at their principal frequencies. Again, weave had a peculiar signature over all  $s$ .

The values for  $F_{DMF}$  (not shown) were generally similar to those for  $F_{DMA}$ . However, there were several notable exceptions. The signatures for water and wood were quite different. For instance, for  $s = 1$ ,  $q = 1$  their values were similar in  $F_{DMA}$ , but in  $F_{DMF}$  the values for wood were much higher in wood than in water. Wool had a particularly dramatic depression at  $s = 3$ ,  $q = 2$ .

For all textures  $F_{DMY}$  at  $q = 1$  decreased with increasing  $s$ , except for weave which had its minimum at  $s = 2$ . At  $q = 3$ , wood and water had the lowest values for  $s = 1$ , raffia and wood for  $s = 2$  and wool and brick for  $s = 3$ . At  $q = 7$ , weave had a very low value for  $s = 3$ .

### B. Feature Quality Analysis

Figs. 11 and 12 graphically present examples of the relative quality of the features used. The shaded area at the top of each bar represents  $\pm 1\sigma$ . For a given feature, the greater the vertical separation between shaded areas for a set of textures, the greater the discrimination capabilities provided by that feature for those textures. The following identifies textures for which the example features provided notably good discrimination (defined as roughly nonoverlapping  $\pm 1\sigma$  regions between a texture and most similar textures): Most features

$F_{CA}(p = 0)$	grass ( $s = 0, 1$ ); wool ( $s = 2$ ); weave ( $s = 2, 3$ ); brick ( $s = 3$ );
$F_{CF}(q = 0)$	sand, wood ( $s = 0$ ); weave ( $s = 1, 3$ ); raffia ( $s = 2$ );
$F_{AM}(q = 0)$	sand, leather, water ( $s = 0$ ); wood ( $s = 0, 1$ ); weave ( $s = 2$ ); brick ( $s = 2, 3$ );
$F_{FM}(q = 0)$	bubbles, grass, wool, water, wood, straw, brick ( $s = 0$ ); weave ( $s = 1, 2$ );
$F_{CA}(p = 1)$	straw ( $s = 0$ ); bubbles ( $s = 0, 1$ ); raffia, wood ( $s = 0, 1, 2$ ); brick ( $s = 3$ );
$F_{CF}(q = 3)$	very similar to $F_{FM}(q = 0)$ ;

$F_{CY}(q = 3)$	sand, leather, water ( $s = 0$ ); wood ( $s = 0, 1$ ); weave ( $s = 2$ ); brick ( $s = 2, 3$ );
$F_{FM}(q = 2)$	water ( $s = 0$ ); weave ( $s = 2$ );
$F_{YM}(q = 3)$	none;
$F_{DMA}(q = 2)$	raffia ( $s = 0, 1, 2$ ); straw ( $s = 1$ ); wood ( $s = 2$ );
$F_{DMF}(q = 1)$	water ( $s = 0, 1, 2$ );
$F_{DMY}(q = 3)$	bubbles ( $s = 0$ ); raffia ( $s = 1, 2$ ); wood ( $s = 2$ );

provided very good discrimination for particular textures, at least for some values of  $q$  and  $s$ . While, for some features, discrimination quality generally deteriorated as  $q$  and/or  $s$  increased, exceptions were evident. For wool and pigskin, the most difficult texture pair to differentiate in the experiments,  $F_{DMA}$  ( $q = 2$ ,  $s = 3$ ) and  $F_{DMY}$  ( $q = 3$ ,  $s = 3$ ), were particularly useful discriminators.

### C. Classification

A model of each type of texture was established using half of its samples to estimate mean and covariance, the parameters required by (35). For the other half of the samples, each was classified as the texture  $t$  that maximized  $p(\mathbf{F}|t)$ . Because of rank deficiency problems in the covariance matrix due to high interfeature correlation, off-diagonal terms in the covariance matrix were set to zero. This is believed to have had a significant adverse effect on classification performance.

Classification performance for the first group of textures (at different rotations) is presented in Table I. Out of a total of 624 sample images, 604 were correctly classified (96.8%). The misclassification rate per competing texture type is  $(100 - 96.8\%)/12 = 0.27\%$ . Bark was misclassified as brick, bubbles, pigskin, sand, and straw; sand as bark; pigskin as bark and wool; grass as leather; leather as grass and straw; wool as bark and pigskin; water as straw; and wood as straw.

Classification performance for the second group of textures (the complete Brodatz album) is presented in Table II. Out of a total of 872 sample images, 701 were classified correctly (80.4%). The misclassification rate per competing texture type is  $(100 - 80.4\%)/108 = 0.18\%$ . Perhaps some comments are in order regarding the classification rate. Many of the textures in the Brodatz album are not homogeneous. Some examples of such textures with less than 50% classification are shown in Fig. 13. Although one can use a selected subset of textures, it will make comparisons between different algorithms more difficult. Notice (in Fig. 13) that the D58 texture has similar segments to those of the other three textures shown, and not surprisingly, 50% of misclassified images from these textures were classified as D58. Examples of some homogeneous textures that resulted in 100% classification are also shown in Fig. 13. Finally, for comparison purposes, when using the same subset of the Brodatz album used by Chang and Kuo [8], 100% of the samples were correctly classified.

## VI. CONCLUSIONS

The approach described herein has proven to be an effective method of rotation-invariant texture classification. However, there is potential for substantial improvement in the classification stage. Performance was limited by using a model

TABLE II  
CLASSIFICATION PERFORMANCE FOR THE SECOND GROUP OF TEXTURES

Sample Texture	% Classification	Sample Texture	% Classification	Sample Texture	% Classification	Sample Texture	% Classification
D1	100%	D29	25%	D57	87.5%	D85	100%
D2	12.5%	D30	62.5%	D58	50%	D86	87.5%
D3	87.5%			D59	25%	D87	100%
D4	100%			D60	25%	D88	50%
D5	62.5%	D33	62.5%	D61	62.5%	D89	50%
D6	87.5%	D34	100%	D62	75%	D90	87.5%
D7	12.5%	D35	100%	D63	75%	D91	25%
D8	75%	D36	100%	D64	100%	D92	100%
D9	100%	D37	100%	D65	100%	D93	87.5%
D10	75%	D38	100%	D66	62.5%	D94	100%
D11	87.5%	D39	75%	D67	100%	D95	100%
D12	87.5%	D40	100%	D68	75%	D96	87.5%
D13	37.5%	D41	87.5%	D69	75%	D97	75%
D14	100%	D42	62.5%	D70	62.5%	D98	50%
D15	75%	D43	87.5%	D71	87.5%		
D16	100%	D44	75%	D72	75%	D100	100%
D17	100%	D45	87.5%	D73	12.5%	D101	87.5%
D18	100%	D46	87.5%	D74	87.5%	D102	100%
D19	87.5%	D47	87.5%	D75	100%	D103	50%
D20	100%	D48	100%	D76	100%	D104	75%
D21	100%	D49	100%	D77	100%	D105	62.5%
D22	100%	D50	100%	D78	100%	D106	75%
D23	62.5%	D51	75%	D79	100%	D107	50%
D24	100%	D52	100%	D80	100%	D108	87.5%
D25	87.5%	D53	100%	D81	75%	D109	75%
D26	87.5%	D54	87.5%	D82	100%	D110	50%
D27	25%	D55	100%	D83	100%	D111	100%
D28	75%	D56	100%	D84	100%	D112	100%

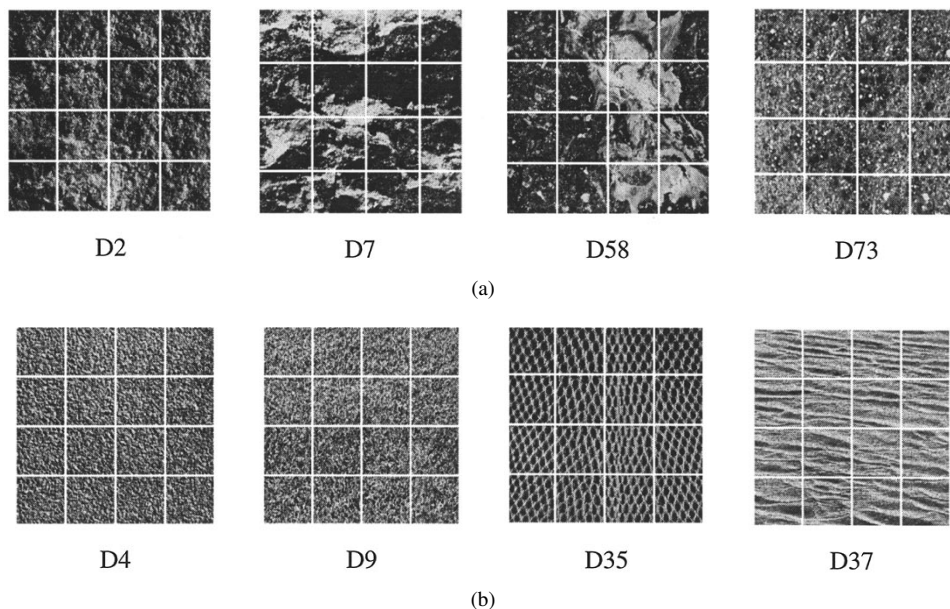


Fig. 13. (a) Some examples of nonhomogeneous textures. About 50% of the sample patterns from D2, D7, and D73 were classified as D58. (b) Few examples of homogeneous texture patterns with 100% classification rate.

based on a variance vector rather than a covariance matrix. This implies that the features are independent, when in fact there is a high degree of interdependence as manifested in the covariance matrices. This is comparable to using Euclidean distance rather than a Bayesian distance, which has been shown to result in over a tenfold reduction in performance

in a similar application [8]. An alternate approach would be to orthogonalize the features using a Karhunen–Loeve transform. This approach would have the ancillary benefit of fitting the Gaussian assumption better, since many of the histograms observed were very non-Gaussian. However, a much larger sample set would be required.

The concepts described herein may be used in other applications. The two level modeling approach facilitates an integrated texture segmentation/classification scheme. Segmentation could be performed using only the micromodel, followed by classification using the macromodel. Microfeatures are potentially useful features for classification of images other than textures.

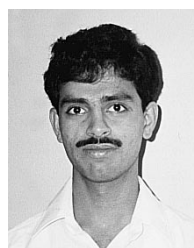
Finally, an approach worthy of future consideration is the development of a rotation invariant Markovian model based on microfeatures. While the complexities would be considerable, it may provide significant improvements in classification (as well as segmentation) performance by combining the strengths of Markovian model-based and feature-based approaches.

## REFERENCES

- [1] M. J. Bastiaans, "A sampling theorem for the complex spectrogram, and Gabor's expansion of a signal in Gaussian elementary signals," *Opt. Eng.*, vol. 20, pp. 594–598, 1981.
- [2] ———, "Gabor's signal expansion and degrees of freedom of a signal," *Opt. Acta*, vol. 29, pp. 1223–1229, 1982.
- [3] A. C. Bovik, "Analysis of multichannel narrow-band filters for image texture segmentation," *IEEE Trans. Signal Processing*, vol. 39, pp. 2025–2042, Sept. 1991.
- [4] A. C. Bovik, M. Clark, and W. S. Geisler, "Multichannel texture analysis using localized spatial filters," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, pp. 55–73, Jan. 1990.
- [5] P. Brodatz, *Textures: A Photographic Album for Artists and Designers*. New York: Dover, 1966.
- [6] R. Chellappa and S. Chatterjee, "Classification of textures using Gaussian Markov random field models," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 33, pp. 959–963, Aug. 1985.
- [7] R. Chellappa, R. L. Kashyap, and B. S. Manjunath, "Model-based texture segmentation and classification," in *Handbook of Pattern Recognition and Computer Vision*, C. H. Chen, L. F. Pau, and P. F. P. Wang, Eds. Singapore: World Scientific, 1992.
- [8] T. Chang and C. C. J. Kuo, "Texture analysis and classification with tree-structured wavelet transform," *IEEE Trans. Image Processing*, vol. 2, pp. 429–441, Oct. 1993.
- [9] Y. Choe and R. L. Kashyap, "3-D shape from a shaded and textural surface image," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, pp. 907–918, Sept. 1991.
- [10] F. S. Cohen, Z. Fan, and M. A. Patel, "Classification of rotated and scaled textured image using Gaussian Markov random field models," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, pp. 192–202, Feb. 1991.
- [11] G. R. Cross and A. K. Jain, "Markov random field texture models," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 5, pp. 25–39, Jan. 1983.
- [12] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis," *IEEE Trans. Inform. Theory*, vol. 36, pp. 961–1005, Sept. 1990.
- [13] J. G. Daugman, "Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 1169–1179, July 1988.
- [14] ———, "Uncertainty relation for resolution in space, spatial frequency and orientation optimized by two-dimensional visual cortical filters," *J. Opt. Soc. Amer.*, vol. 2, pp. 1160–1169, 1985.
- [15] H. Derin and H. Elliott, "Modeling and segmentation of noisy and textured images using Gibbs random fields," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-9, pp. 39–55, Jan. 1987.
- [16] H. Derin, H. Elliott, R. Cristi, and D. Geman, "Bayes smoothing algorithms for segmentation of binary images modeled by Markov random fields," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-6, pp. 707–720, Nov. 1984.
- [17] H. Greenspan, S. Belongie, R. Goodman, and P. Perona, "Rotation invariant texture recognition using a steerable pyramid," in *Proc. IEEE Int. Conf. Image Processing*, Jerusalem, Israel, Oct. 1994.
- [18] D. Gabor, "Theory of communication," *J. Inst. Elect. Eng.*, vol. 93, pp. 429–457, 1946.
- [19] G. M. Haley and B. S. Manjunath, "Rotation invariant texture classification using modified Gabor filters," in *Proc. IEEE Int. Conf. Image Processing*, Washington, DC, Oct. 1995.
- [20] R. L. Kashyap, R. Chellappa, and A. Khotanzad, "Texture classification using features derived from random field models," *Pattern Recognit. Lett.*, pp. 43–50, Oct. 1982.
- [21] R. L. Kashyap and A. Khotanzad, "A model-based method for rotation invariant texture classification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 8, pp. 472–481, July 1986.
- [22] M. M. Leung and A. M. Peterson, "Scale and rotation invariant texture classification," in *Proc. 26th Asilomar Conf. Signals, Systems and Computers*, Pacific Grove, CA, Oct. 1992.
- [23] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, pp. 672–693, Jan. 1990.
- [24] B. S. Manjunath and R. Chellappa, "A computational approach to boundary detection," in *Proc. IEEE Computer Soc. Conf. Computer Vision and Pattern Recognition*, Lahaina, Hawaii, June 3–6, 1991.
- [25] S. Marcelja, "Mathematical description of the responses of simple cortical cells," *J. Opt. Soc. Amer.*, vol. 70, pp. 1297–1300, 1980.
- [26] M. Porat and Y. Y. Zeevi, "The generalized scheme of image representation in biological and machine vision," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 10, pp. 452–468, July 1988.
- [27] ———, "Localized texture processing in vision: Analysis and synthesis in the Gaborian space," *IEEE Trans. Biomed. Eng.*, vol. 36, pp. 115–129, Jan. 1989.
- [28] M. R. Turner, "Texture discrimination by Gabor functions," *Biol. Cybern.*, vol. 55, pp. 71–82, 1986.
- [29] J. You and H. A. Cohen, "Classification and segmentation of rotated and scaled textured images using 'tuned' masks," *Pattern Recognit.*, vol. 26, pp. 245–258, 1993.
- [30] Signal and Image Processing Inst., Univ. Southern Calif., Los Angeles (URL sipi.usc.edu).

**George M. Haley** received the B.S. degree from the University of Notre Dame, South Bend, IN, in 1984, and the M.S. degree from the University of California, Santa Barbara, in 1998.

He is currently a Director at Ameritech, Hoffman Estates, IL. Since 1984, he has been involved in the development of several acoustic, IR, and RF sensor systems in the areas of algorithm development, system definition and analysis, software development and management. His current technology interests are in architecture; sensor-based systems; modular, reusable software implementations; and strategic technology management.



**B. S. Manjunath** (S'88–M'91) received the B.E. degree in electronics (with distinction) from Bangalore University, Bangalore, India, in 1985, the M.E. degree (with distinction) in systems science and automation from the Indian Institute of Science in 1987, and the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, in 1991.

He joined the Electrical and Computer Engineering Department at the University of California, Santa Barbara, in 1991, where he is now an Associate Professor. During the summer of 1990, he worked at the IBM T. J. Watson Research Center, Yorktown Heights, NY. His current research interest include computer vision, learning algorithms, image/video databases and digital libraries.

Dr. Manjunath is currently an Associate Editor of the IEEE TRANSACTIONS ON IMAGE PROCESSING and is a Guest Editor of a 1999 special issue of the TRANSACTIONS on image and video processing in digital libraries. He has served on the program committees of many international conferences and workshops, and was on the organizing committee of the 1998 International Conference on Image Processing. He was a recipient of the National Merit Scholarship (1978–1985) and was awarded the University Gold Medal for the best graduating student in electronics engineering in 1985 from Bangalore University.