

Multichannel Texture Analysis Using Localized Spatial Filters

ALAN CONRAD BOVIK, MEMBER, IEEE, MARIANNA CLARK, MEMBER, IEEE, AND WILSON S. GEISLER

Abstract—A computational approach for analyzing visible textures is described. Textures are modeled as irradiance patterns containing a limited range of spatial frequencies, where mutually distinct textures differ significantly in their dominant characterizing frequencies. By encoding images into multiple narrow spatial frequency and orientation channels, the slowly-varying channel envelopes (amplitude and phase) are used to segregate textural regions of different spatial frequency, orientation, or phase characteristics. Thus, an interpretation of image texture as a *region code*, or *carrier* of region information, is emphasized.

The channel filters used, known as the 2-D Gabor functions, are useful for these purposes in several senses: they have tunable orientation and radial frequency bandwidths, tunable center frequencies, and optimally achieve joint resolution in space and in spatial frequency. By comparing the channel amplitude responses, boundaries between textures can be detected. By locating large variations in the channel phase responses, discontinuities in texture phase can be detected. Examples are given of both types of texture processing using a variety of real and synthetic textures.

Index Terms—Amplitude demodulation, complex modulation, computer vision, Gabor function, multiple channel, phase demodulation, texture analysis, texture segmentation.

I. INTRODUCTION

TEXTURE can be used in the analysis of images in several ways: in the segmentation of scenes into distinct objects and regions, in the classification or recognition of surface materials, and in the computation of surface shape. However, an exact definition of texture as either a surface property or as an image property has never been adequately formulated. While the concept of a *surface texture* as a pattern of variations in macroscopic surface topology is easy to accept, real-world surface textures are extremely difficult to model mathematically. Modeling texture in an image as a function of surface texture is yet more complex, since an accurate model must incorporate descriptions of both the optical properties of the surface materials and of the geometries of the lighting sources and imaging system. The modeling problem is made effectively unapproachable by the infinite variety of

illumination, shadowing, and surface reflectance patterns, surface topologies and self-occlusions, and perspective distortion effects that must be considered.

However, much progress can be made towards developing texture analysis techniques by examining the problem in the post-image formation setting. In doing so, a *perceptual surface texture* may be informally defined to be a spatial pattern of local surface radiances giving rise to a perception of surface homogeneity. Within this context, an *image texture* may be defined as a local arrangement of image irradiances projected from a surface patch of perceptually homogeneous radiances.

Attributes giving rise to a sense of perceptual homogeneity have been construed to include such spatio/temporal surface properties as color, relative depth, motion, and flicker rate [1]. While integrating these properties into a scene analysis system can contribute powerfully to the computation of spatial form, the scope of this paper is confined to the analysis of textures occurring in static, monocular, monochromatic images. However, it is possible that the type of analysis used here may be extended to the analysis of depth or color patterns, or patterns that vary with time.

The analysis of texture requires the identification of those texture attributes which can be used for segmentation, discrimination, recognition, or shape computation, and the development of computational approaches for accomplishing these tasks. Here, low-level segmentation of images based on texture is considered, viz. without scrupulousness, higher-level interpretation, or *a priori* knowledge of the texture types that may be encountered [1]–[3]. Thus, discrimination is modeled as well, since the discrimination problem is embedded in the segmentation problem. However, recognition of real-world surface textures can only be achieved in a rudimentary sense using low-level processing. As with shape recognition, texture recognition requires the integration of computed low-level information with high-level knowledge. Thus, recognition of real-world textures is not considered; only textures occurring in images are considered, and *texture* will be regarded as being synonymous with image texture unless otherwise indicated.

In this paper, textures are modeled as irradiance patterns which are distinguished by a high concentration of localized spatial frequencies. Distinct textures are characterized as differing significantly in their dominant spatial frequencies. This model of texture allows the devel-

Manuscript received April 18, 1988; revised June 15, 1989. Recommended for acceptance by W. E. L. Grimson. This work was supported in part by the Texas Advanced Research Program under Grant 3546.

A. C. Bovik is with the Department of Electrical and Computer Engineering, University of Texas, Austin, TX 78712.

M. Clark is with Lockheed Missiles and Space Corporation, Austin Division, Austin, TX.

W. S. Geisler is with the Department of Psychology, University of Texas, TX 78712.

IEEE Log Number 8931857.

opment of a computational framework for texture segmentation. In the approach taken here, textured images are encoded into multiple narrow spatial frequency and orientation channels. The output of each channel is a complex-modulated subimage whose instantaneous amplitude and phase envelopes describe the spatial support of the frequencies and/or orientations to which the channel is tuned. Thus, at a global scale of perception, texture is regarded simply as a "carrier" of region information. Within this framework, the local structure of a texture is described by the orientations and frequencies of the carriers (viz. by the center frequencies and orientations of the active channels), whereas information describing the spatial extent of the texture is contained in the envelopes of the channel outputs. Thus, the analysis of texture is reduced to the analysis of the envelopes of the active channel outputs.

There are many plausible choices for the channel filters used in encoding textured images; from a practical point of view, it may be that certain filters may be useful for some segmentation tasks but not for others. However, the 2-D Gabor filters have been shown to be particularly useful for analyzing textured images containing highly specific frequency or orientation characteristics [4]–[6]. Gabor filters have been applied by Clark *et al.* [4], [6] to both natural and artificial textures, and by Turner [5] to artificial textures similar to those often used in psychophysical experiments. The segmentations achieved bear a striking resemblance to perception.

The 2-D Gabor filters are appropriate for textural segmentation in several senses: they have tunable orientation and radial frequency bandwidths, tunable center frequencies, and optimally achieve joint resolution in space and spatial frequency. The demodulated Gabor channel envelopes generally contain only low spatial frequencies which are optimally localized in both domains. By comparing the channel amplitude envelopes, the boundaries between textured regions differing markedly in their spatial frequency content can be accurately located. Discontinuities in texture phase, possibly arising from surface discontinuities, can be located by detecting large variations in the phase envelopes.

The framework for textural analysis developed here has its origin in recent physiological evidence of the architecture of the neural receptive fields in the striate visual cortex. However, the approach is intended primarily as a computer vision method for textural segmentation, and it is argued that the model is appropriate within such a computational framework. While it is possible that the Gabor-shaped receptive fields are fundamental to biological processing of texture, any extension of the approach described here for modeling analogous biological processing must inevitably be verified by extensive psychophysical and electrophysiological experimentation.

A. Organization of the Paper

In this paper, texture segmentation is modeled as a process whereby textured image regions are segregated by

localizing spatial changes in the frequency, orientation, or phase of the textures. In Section II, a class of channel filters are described (the 2-D Gabor filters) that are used to encode images into multiple complex-modulated subimages. If the density of the channel filters is sufficiently high, the image can be exactly recovered from the subimages. The energy in each subimage is concentrated in both spatial frequency (within the passband of the channel filter) and in space (within the support of the textures to which the channel is tuned), allowing accurate localization of boundaries occurring between textures having different dominant frequency characteristics. The salient properties of the Gabor filters which make them advantageous for this purpose are reviewed, and the necessary sampling densities required to accurately represent the Gabor filters in discrete form are derived. Methods for recovering (demodulating) the Gabor channel amplitude responses are described in Section III. The amplitude demodulation model is then applied to the texture segmentation problem, where the boundaries between textures are localized by comparing the demodulated amplitude responses. It is also shown that the amplitude responses can be resampled to provide an efficient and complete representation of the segmented textural regions. The selection of an appropriate set of Gabor filters to obtain an accurate segmentation is also considered. Examples are given of the segmentation technique as applied to a variety of images.

Methods for extracting useful information from the Gabor phase responses are considered in Section IV. In particular, it is found that abrupt transitions in the computed channel phases may be used to detect phase discontinuities within a texture. Texture phase discontinuities occur when one part of a texture is shifted with respect to another, possibly arising from surface discontinuities or from joins of textures which are identical except for a spatial shift. Thus, phase discontinuities are only located within textures, viz. after the textures have been segregated using amplitude demodulation. Section V surveys research in computational vision, visual physiology, and visual psychophysics that directly pertains to or motivates this work. The paper is concluded in Section VI with a discussion of possible extensions of the multiple channel model to other aspects of spatial vision.

II. THE MULTIPLE CHANNEL FILTER MODEL

In this section, the basic properties of the Gabor channel filters are described. In Section II-A, the 2-D Gabor functions are defined: they are complex sinusoidal (carrier) gratings modulated by 2-D Gaussian functions in the space domain, and shifted Gaussians in the spatial frequency domain. Thus, the Gabor functions are complex-valued functions on \mathbf{R}^2 , except when the carrier sinusoid has zero center frequency. In this case, the Gabor functions reduce to the 2-D Gaussian functions. In Section II-B, those properties of Gabor filters which make them useful for localizing image regions containing highly specific spatial frequency information are described. The Ga-

bor filters can be configured to have various shapes, bandwidths, center frequencies, and orientations by the adjustment of suitable parameters. By varying these parameters, a filter can be made to pass any elliptical region of spatial frequencies. Regardless of the region of frequencies passed, the 2-D Gabor functions uniquely minimize the 2-D space-frequency uncertainty principle for complex-valued functions on \mathbf{R}^2 . Moreover, for reasonably narrow bandwidths the real and imaginary components of a Gabor function (or an image filtered by a Gabor function) closely approximate a Hilbert transform pair. Hence, the Gabor functions can be interpreted as the product of a modulating amplitude envelope with a complex carrier function whose argument is a modulating phase envelope. The amplitude and phase envelopes can be computed and analyzed separately, forming the basis for the segmentation strategies detailed in Sections III and IV. Finally, in Section II-D, discretization of the Gabor filters is considered. In particular, the minimum sampling density required to adequately represent a Gabor filter in discrete form is derived.

A. Gabor Channel Filters

The convolution version of the complex 2-D Gabor functions have the general form ($j = \sqrt{-1}$)

$$h(x, y) = g(x', y') \cdot \exp [2\pi j(Ux + Vy)], \quad (1)$$

where $(x', y') = (x \cos \phi + y \sin \phi, -x \sin \phi + y \cos \phi)$ are rotated coordinates, and where

$$g(x, y) = \left(\frac{1}{2\pi\lambda\sigma^2} \right) \cdot \exp \left[-\frac{(x/\lambda)^2 + y^2}{2\sigma^2} \right]. \quad (2)$$

Thus, $h(x, y)$ is a complex sinusoidal grating modulated by a 2-D Gaussian function with aspect ratio λ , scale parameter σ , and major axis oriented at an angle ϕ from the x -axis. If $\lambda = 1$, then ϕ need not be specified since then $g(x, y)$ is circularly symmetric. The spatial frequency response of the Gabor function (1) is

$$H(u, v) = \exp \left\{ -2\pi^2\sigma^2[(u' - U')^2\lambda^2 + (v' - V')^2] \right\} \quad (3)$$

where $(u', v') = (u \cos \phi + v \sin \phi, -u \sin \phi + v \cos \phi)$ and (U', V') is a similar rotation of the center frequency (U, V) . Thus, $H(u, v)$ is a bandpass Gaussian with minor axis oriented at an angle ϕ from the u -axis, aspect ratio $1/\lambda$, radial center frequency $F = \sqrt{U^2 + V^2}$ (measured in cycles/image) and orientation $\theta = \tan^{-1}(V/U)$ (degrees or radians measured from the u -axis). Fig. 1(a) shows intensity plots of the real and imaginary parts of several Gabor filters for $\phi = 0$ and several values of θ , and their spatial frequency responses, while Fig. 2 is a perspective plot of the real and imaginary components of a Gabor function.

Although it is not necessarily desirable to eliminate from consideration Gabor filters having arbitrarily oriented modulating Gaussians, it is usually most convenient to consider filters whose modulating Gaussians have the

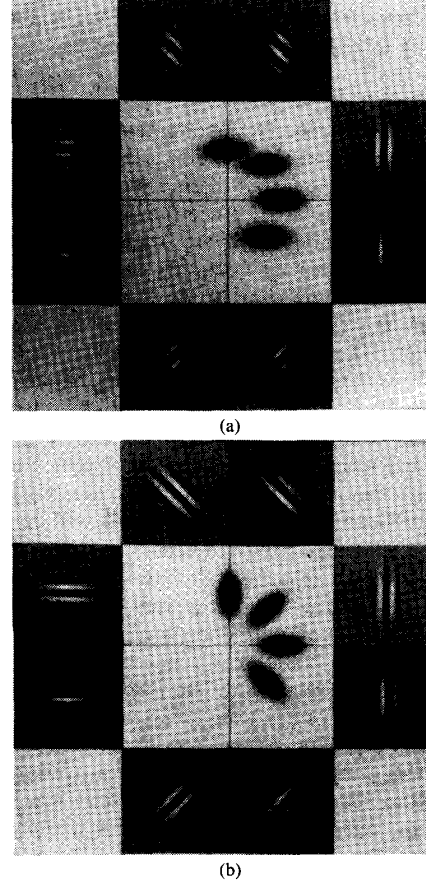


Fig. 1. Intensity plots of 2-D Gabor filters having center frequency $F = 32$ cycles/image, aspect ratio $\lambda = 1/2$, and orientations $\theta = 0^\circ, \pm 45^\circ$, and 90° . The complex components of the impulse response are shown in pairs along the perimeter; the spatial frequency responses are shown in the center. (a) Gabor filters (1) with $\phi = 0$; (b) "daisy petal" Gabor filters (4) with $\phi = \theta$.

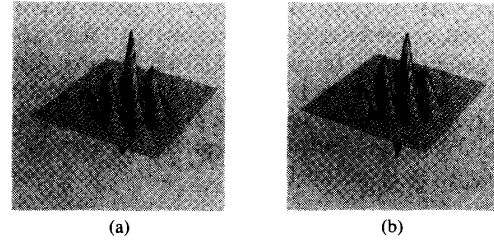


Fig. 2. Perspective views of the (a) real and (b) imaginary components of a Gabor filter with center frequency $F = 50$ cycles/image and unity aspect ratio.

same orientation as the complex sine grating ($\phi = \theta$). In this case, (1) and (3) reduce to

$$h(x, y) = g(x', y') \cdot \exp (2\pi j F x') \quad (4)$$

and

$$H(u, v) = \exp \left\{ -2\pi^2\sigma^2[(u' - F)^2\lambda^2 + (v')^2] \right\}. \quad (5)$$

Fig. 1(b) shows intensity plots of several 2-D Gabor functions having this "daisy petal" configuration.

B. Properties of Gabor Filters

It is intuitive to specify Gabor filters in terms of their spatial frequency and orientation bandwidths. However, the bandwidths of the Gabor filters having the form (1) do not have a clear meaning, since the range of radial frequencies passed varies with the orientation of the frequencies. The frequency (octave) and orientation (radian) half-peak bandwidths of the "daisy petal" Gabor filter (4) are $B = \log_2 [(\pi F \lambda \sigma + \alpha)/(\pi F \lambda \sigma - \alpha)]$ and $\Omega = 2 \tan^{-1} [\alpha/(\pi F \sigma)]$, respectively, where $\alpha = \sqrt{(\ln 2)/2}$, as depicted in Fig. 3. By varying the free parameters B , Ω , F , and θ , a Gabor filter of arbitrary orientation and bandwidth characteristics can be generated spanning any elliptical portion of the spatial frequency domain with major axis passing through the origin. Generally, only Gabor filters of the form (4) will be considered except where otherwise indicated.

It is well-known that the 1-D Gaussian function $g(x) \propto \exp(-x^2/2\sigma^2)$ with Fourier transform $G(u) \propto \exp(-2\pi^2\sigma^2 u^2)$ is the only function mapping $\mathbf{R} \rightarrow \mathbf{R}$ which achieves the lower bound of the uncertainty relationship $\Delta x \cdot \Delta u \geq 1/4\pi$ relating the variances (normalized by energy) Δx and Δu of $g(x)$ and $G(u)$, respectively. However, the more general complex-valued 1-D Gabor functions $h(x) = g(x) \cdot \exp(2\pi j U x)$ were proved by Gabor in 1946 [8] to also minimize the uncertainty relationship, and to be the only class of 1-D functions mapping $\mathbf{R} \rightarrow \mathbf{C}$ achieving the lower bound. Thus, shifting the frequency response of a Gaussian function does not change its localization in either domain. Daugman [9] extended the uncertainty relationship to 2-D by demonstrating the 2-D Gabor functions $h(x, y)$ to be the only functions mapping $\mathbf{R}^2 \rightarrow \mathbf{C}$ simultaneously achieving the lower bounds of the uncertainty inequalities $\Delta x \cdot \Delta u \geq 1/4\pi$ and $\Delta y \cdot \Delta v \geq 1/4\pi$. Thus, the 2-D Gabor filters optimally and uniquely achieve simultaneous localization in space and in spatial frequency. This implies that they can be defined to have very narrow frequency and orientation responses while maintaining spatial localization.

The Gabor function (4) has real and imaginary components given respectively by

$$h_c(x, y) = g(x', y') \cos(2\pi F x')$$

and

$$h_s(x, y) = g(x', y') \sin(2\pi F x'). \quad (6)$$

The functions $h_c(x, y)$ and $h_s(x, y)$ are, respectively, even and odd symmetric along the preferred orientation direction θ . The results of convolving $h_c(x, y)$ and $h_s(x, y)$ with any 2-D function are approximately identical other than a difference of $\pi/2$ in the phase spectra along the direction θ , viz. $h_c(x, y)$ and $h_s(x, y)$ are approximately in phase quadrature.

The quadrature relationship of the complex components of the Gabor filter allows for a useful and unique approach

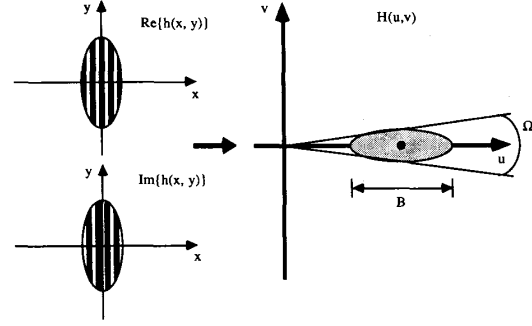


Fig. 3. Half-peak radial frequency (B) and orientation (Ω) bandwidths of the daisy petal Gabor filter given in (4).

to texture processing. The sinusoidal carrier acts to translate the frequency response of the Gabor filter along the modulation axis by a distance F from the origin. For small spatial frequency bandwidths, $H(u, v)$ is negligible for $u' < 0$; e.g., for radial frequency bandwidths $B = 0.7$, 1.0, and 1.3 octaves (typical of those used here), $H^2(u, v)$ falls below 212 dB, 108 dB, and 68 dB from peak value at $(u', v') = (0, 0)$. Thus $h_c(x, y)$ and $-h_s(x, y)$ closely approximate a Hilbert transform pair for fixed values of y' :

$$\Gamma_\theta \{h_c(x, y)\} = \frac{-1}{\pi x'} * h_c(x, y) = -h_s(x, y), \quad (7)$$

where $\Gamma_\theta \{\cdot\}$ denotes the 1-D Hilbert transform taken along direction θ . The 2-D Gabor filter $h(x, y)$ is subsequently the *analytic function* associated with $h_c(x, y)$, and $-h_s(x, y)$ is the *quadrature filter* of $h_c(x, y)$ to within a good approximation [10]. Since any analytic function can be interpreted as the product of a modulating envelope or space-varying phasor and a complex carrier function, the instantaneous envelope of $h(x, y)$ (the modulating Gaussian) can be recovered by computing its complex magnitude. Moreover, since the convolution of an analytic function with another function is also analytic, filtering an image with a 2-D Gabor filter results in a complex-valued analytic image containing only a limited range of frequencies and orientations. The complex magnitude of the filtered image will be maximized over those regions of the original image characterized by the particular spatial frequency attributes to which the filter is tuned.

C. Gabor Expansion of an Image

Any finite-dimensional function can be expressed as a weighted sum of appropriately shifted Gabor functions, known as the *Gabor expansion* [8], [11], [12]. If $t(x, y)$ is an arbitrary function on \mathbf{R}^2 , then it can be expanded into the sum

$$t(x, y) = \sum_{r=-\infty}^{\infty} \sum_{s=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \cdot \beta_{rsmn} g(x - x_r, y - y_s) \cdot \exp \{2\pi j [U_m(x - x_r) + V_n(y - y_s)]\}$$

where the sequences of shifts $\{x_r\}$ and $\{y_s\}$ and modulation frequencies $\{U_m\}$ and $\{V_n\}$ have constant spacings X , Y , U , and V satisfying $XU = YV = 1$. The resulting grid of shifts and frequencies (four-dimensional in this case) is known as the *Gabor lattice*. The expansion coefficients $\{\beta_{rsmn}\}$ form a *complete representation* of $t(x, y)$, since $t(x, y)$ can be exactly reconstructed from the coefficients. Methods for computing the exact expansion coefficients are discussed in [11], [12]; generally it is a complex process. However, to a good approximation [13]

$$\beta_{rsmn} \approx t(x_r, y_s) * g(x_r, y_s) \cdot \exp \{2\pi j [U_m x_r + V_n y_s]\},$$

provided that the Gabor filters do not overlap too much in their frequency responses. Assuming that the modulating Gaussians are isotropic with space constant σ , then more than 99 percent of the energy of each filter is confined to one of the nonoverlapping $U \times V$ frequency rectangles $[U_m - U/2, U_m + U/2] \times [V_m - V/2, V_m + V/2]$ provided that $U, V > 1/2\sigma$. Thus, to close approximation the expansion coefficients can be computed by convolving the image with a set of Gabor functions having sufficiently spaced center frequencies. Moreover, a finite subset of the coefficients (finite number of Gabor filters) will adequately represent the important frequencies of the image if the image is approximately bandlimited, or if the image is sampled. Clearly, the highest radial frequency in any particular direction bounds the number of filters needed. Thus, the image can be closely reconstructed from the responses of a finite set of Gabor filters. More importantly, however, the expansion coefficients β_{rsmn} having relatively large magnitudes for a given (r, s) correspond with the dominant frequencies occurring *within the spatial vicinity* of (x_r, y_s) in the image. For our purposes, these may be interpreted as the dominant local texture frequency components. While the expansion may be generalized using “window” functions other than the Gaussian [11], [12], the measurement of spatially and spectrally localized instantaneous frequencies is achieved optimally using the Gaussian.

D. Sampling the Gabor Filters

Thus far, only the continuous form of the Gabor channel filters have been considered. Since the object here is the development of techniques for the computer analysis of textured images using Gabor filters, the effects of discretizing the filters must be considered. To achieve a fair representation of the filters in discrete form, adequate sampling and quantization densities must be used. Here, the minimum sampling densities required to ensure that the effects of aliasing do not become excessive are considered. The effects of quantizing the filter coefficients is beyond the scope of this work and is not considered.

Maintaining an adequate sampling density is particularly important when implementing Gabor filters for discrete convolution, since different choices of the filter parameters lead to very different sampling requirements. If care is not taken in the sampling, the attractive properties of the Gabor filters will be lost.

For simplicity, consider the Gabor filter (6) with $\theta = 0$. In a discrete spatial convolution, the real and imaginary components are computed separately. Since the magnitude of the frequency responses of the complex components are effectively equal, it suffices to consider a sampling of the real part. Moreover, assume that the critical sampling density will be determined along the direction θ (the x -direction), which is equivalent to assuming that the orientation sensitivity of the filter is reasonably specific. Thus, consider sampling of

$$h_c(x) \propto \exp \left(-\frac{x^2}{2\lambda^2\sigma^2} \right) \cdot \cos(2\pi Fx) \quad (8)$$

with spatial frequency response

$$H_c(u) = \left\{ \exp \left[-2[\pi\lambda\sigma(u - F)]^2 \right] + \exp \left[-2[\pi\lambda\sigma(u + F)]^2 \right] \right\} / 2.$$

An ideal sampling with period $X = 1/f_s$ is obtained by multiplying (8) by the 1-D impulse train $\sum_{n=-\infty}^{\infty} \delta(x - nX)$, yielding the discrete impulse response

$$h_c(n) = \exp \left[-\frac{(nX)^2}{2\lambda^2\sigma^2} \right] \cdot \cos(2\pi FnX)$$

with frequency response

$$\begin{aligned} H_c(u) * X^{-1} \sum_{m=-\infty}^{\infty} \delta(u - mX^{-1}) \\ = X^{-1} \sum_{m=-\infty}^{\infty} H_c(u - mX^{-1}). \end{aligned}$$

Since Gabor filters are not bandlimited, some aliasing will occur regardless of the sampling density. The most obvious constraint that must be imposed is that $f_s > 2F$. Since the main lobe of $H_c(u)$ (a shifted Gaussian) decays eternally, only the overlap between the upper main lobe of $X^{-1}H_c(u)$ and the lower main lobe of the first harmonic $X^{-1}H_c(u - X^{-1})$ need be considered. As pointed out in [14] and in [15] in the context of edge detection filters, the effect of this aliasing can be effectively measured by computing the percent of the energy of the upper main lobe of $X^{-1}H_c(u)$ that falls above the half sampling frequency $(2X)^{-1}$, i.e.,

$$\begin{aligned} \xi &= \int_{(2X)^{-1}}^{\infty} \exp \left\{ -4[\pi\lambda\sigma(u - F)]^2 \right\} du / \\ &\quad \int_0^{\infty} \exp \left\{ -4[\pi\lambda\sigma(u - F)]^2 \right\} du \\ &= \Phi \left\{ \sqrt{2}\pi\sigma[F - (2X)^{-1}] \right\} / \Phi \left\{ \sqrt{2}\pi\sigma F \right\}, \end{aligned}$$

where

$$\Phi(a) = \left(\frac{1}{\sqrt{2\pi}} \right) \int_{-\infty}^a \exp(-b^2/2) db.$$

Since $\sqrt{2}\pi\sigma F = \ln 2 \cdot (2^B + 1)/(2^B - 1) \equiv \gamma_B$, and defining $R_s = f_s/2F > 1$, then

$$\xi = \frac{\Phi[(1 - R_s)\gamma_B]}{\Phi(\gamma_B)}.$$

This expresses the percentage of aliased energy in $h_c(n)$ as a function of the desired filter bandwidth B and the ratio of the sampling frequency to the filter center frequency.

In order to avoid severe aliasing, some maximum value ξ_{\max} of the energy overlap must be enforced if a filter with bandwidth B is to be implemented. The minimum sampling frequency is then specified according to

$$R_s > 1 - \left(\frac{1}{\gamma_B} \right) \Phi^{-1} \left[\frac{\xi_{\max}}{\Phi(\gamma_B)} \right]. \quad (9)$$

An extreme upper bound on the percentage of aliased energy is $\xi_{\max} = 10\%$ [14]. Fig. 4 plots the minimum value of R_s required to maintain this upper bound for reasonable radial frequency bandwidths ranging over $\frac{1}{2} \leq B \leq 3$. The values of R_s range from a low value of about 1.3 to a high of 2.3 over these bandwidths. Here the necessary sampling rates are quite reasonable (relative to the filter center frequencies) due to the narrow Gabor filter bandwidths. For a more reasonable percentage of aliased energy $\xi_{\max} = 1\%$, the minimum values of R_s range from about 1.6 to a high of about 3.5.

The implications of the preceding derivation vary with the flexibility of the imaging system. If images of a preordained sampling density are available, (9) implies that implementation of Gabor filters with low center frequencies may be made more efficient by applying them to subsampled versions of the textured image. However, implementation of filters with high center frequencies (e.g., above 0.2 cycles/pixel) may require that the filters be sampled more densely than the available grid, unless the filter bandwidths B are specified sufficiently narrow. Unfortunately, this may necessitate the use of a larger set of filters if it is desired to cover the frequency plane; moreover, the use of extremely narrowband filters can result in excessive loss of spatial resolving capability between the channel responses.

If there is flexibility available in defining the subimage sampling densities, (9) implies that a hierarchy of images of variable resolutions can be separately filtered by multiple sets of Gabor filters of different sampling densities and reasonable bandwidths, where each set of filters has a common radial frequency band but varying orientation sensitivities. Thus, processes involving interactions between the channel responses can proceed first across the available orientations within a given sampling density (resolution), and then across the available resolutions. Since the channel outputs are generally demodulated as part of the processing, the different sampling densities could be reduced to a common value prior to making interchannel comparisons (see Section III-C).

III. TEXTURE SEGMENTATION VIA CHANNEL AMPLITUDE DEMODULATION

The computational approach to texture segmentation described in this section has its basis in the recovery and analysis of the amplitude envelopes of Gabor filtered images. Means by which the demodulation can be effected

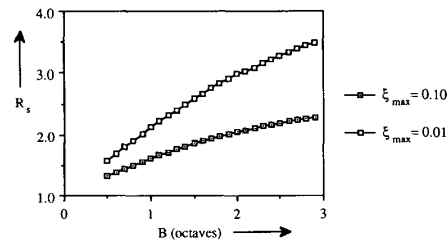


Fig. 4. Ratio of the minimum sampling frequency to twice the Gabor channel center frequency ($R_s = f_s / 2F$) against radial frequency bandwidth B .

are discussed in Section III-A. In Section III-B, the demodulation model is applied to the texture segmentation problem, where it is assumed that each textured region is characterized by a narrow range of dominant spatial frequencies differing significantly from the dominant frequencies of the other textures in the image. In Section III-C it is shown that the demodulated channel amplitude responses can be resampled to allow for more efficient representation and transmission of textural region information. Methods for selecting a set of filters to accomplish segmentation are discussed in Section III-D. This topic is important since it may not be computationally convenient or feasible to apply a large number of filters to an image. Finally Section III-E gives examples of the process as applied to a number of real and synthetic textures.

A. Demodulation of Channel Amplitudes

A continuous setting is assumed to simplicity. Suppose that $t(x, y)$ is a real-valued and continuous monochromatic image containing only a narrow range of spatial frequencies concentrated near (U, V) . Any such narrowband image $t(x, y)$ may be expressed in the form

$$\begin{aligned} t(x, y) &= 2 \left\{ c(x, y) \cos [2\pi(Ux + Vy)] \right. \\ &\quad \left. + s(x, y) \sin [2\pi(Ux + Vy)] \right\} \\ &= a(x, y) \cos [2\pi(Ux + Vy) + p(x, y)], \end{aligned} \quad (10)$$

where $a(x, y) = \sqrt{c^2(x, y) + s^2(x, y)}$ and $p(x, y) = -\tan^{-1} [s(x, y)/c(x, y)]$ are slowly-varying amplitude and phase terms. In the following, $t(x, y)$ will be regarded as a simple model for a texture, or a dominant component of a texture, that contains substantial energy over a single narrow range of spatial frequencies and orientations. The "texture" model (10) is admittedly somewhat idealized, since real-world image textures may generally contain multiple components of the form (10) distributed over the spatial frequency plane. Moreover, real-world textures may contain wideband components that are not adequately described by (10). However, since the global is to *segment* images based on texture, rather than to form a complete description of the textures, it is only necessary to identify narrow spectral regions over

which the frequency content of adjacent textures differ. Within this context, the model (10) is no more unreasonable than the usual ideal step edge model used in evaluating edge detection schemes.

Applying the complex-valued Gabor filter (4) to $t(x, y)$ yields a complex response $k(x, y) = k_c(x, y) + jk_s(x, y)$ with respective real and imaginary components

$$\begin{aligned} k_c(x, y) &= \text{Re} \{k(x, y)\} = h_c(x, y) * t(x, y) \\ &= [g(x', y') * c(x, y)] \cos(2\pi Fx') \\ &\quad + [g(x', y') * s(x, y)] \sin(2\pi Fx') \\ k_s(x, y) &= \text{Im} \{k(x, y)\} = h_s(x, y) * t(x, y) \\ &= [g(x', y') * c(x, y)] \sin(2\pi Fx') \\ &\quad - [g(x', y') * s(x, y)] \cos(2\pi Fx'). \end{aligned} \quad (11)$$

The response $k(x, y)$ can be computed either by calculating $k_c(x, y)$ and $k_s(x, y)$ separately by 2-D convolution or directly via Fourier transform inversion. By previous assumptions $k(x, y)$ is an analytic signal and $k_c(x, y)$ and $k_s(x, y)$ form a quadrature pair. The amplitude and phase envelopes of $k(x, y)$ may be recovered via

$$\begin{aligned} m(x, y) &= \sqrt{k_c^2(x, y) + k_s^2(x, y)} \\ &= \sqrt{[g(x', y') * c(x, y)]^2 + [g(x', y') * s(x, y)]^2} \\ \text{and} \\ \psi(x, y) &= \tan^{-1} \left[\frac{k_s(x, y)}{k_c(x, y)} \right] \\ &= 2\pi Fx' - \tan^{-1} \left[\frac{g(x', y') * s(x, y)}{g(x', y') * c(x, y)} \right] \end{aligned} \quad (12)$$

where $k(x, y) = m(x, y) \cdot \exp[j\psi(x, y)]$. The envelopes $m(x, y)$ and $\psi(x, y)$ may be thought of as smoothed versions of $a(x, y)$ and $p(x, y)$. Computation of $m(x, y)$ and $\psi(x, y)$ is a simple matter using a computer, although direct computation of $\psi(x, y)$ results in a “wrapped” phase which is approximately periodically discontinuous. However, in computing global texture or surface properties, the variation in $\psi(x, y)$ is of principal concern since the actual phase values are only meaningful with respect to some reference point. Methods for extracting useful information from $\psi(x, y)$ are considered in Section IV.

The Hilbert transform interpretation of the Gabor filtered image $k(x, y)$ allows for the simple and meaningful method (12) for recovering the amplitude response $m(x, y)$. However, there are other means by which the amplitude envelope of $k(x, y)$ may be estimated. For example, full-wave rectification results in the weighted envelope

$$\begin{aligned} &|k_c(x, y)| + |k_s(x, y)| \\ &= m(x, y) \left\{ |\cos(\psi(x, y))| + |\sin(\psi(x, y))| \right\}. \end{aligned}$$

Although this demodulation strategy has been successfully used for segmenting textured images [4], the sinusoidal terms introduce additional variation in the computed envelope. This variation can only be removed by a low-pass post-filter. Since the amplitude responses can generally be expected to contain additional fluctuations due to transitions in texture phase, it is generally advisable to compute the complex magnitude (12).

B. Texture Segmentation by Channel Amplitude Comparisons

The demodulation model described in the preceding did not require the assumption that the channel filters are Gaussian-modulated sinusoids. In fact, the results are equally valid if the Gaussian function $g(x, y)$ is replaced by another smoothing filter. The segmentation strategy detailed here also does not require that the channel filters be Gabor-shaped. However, as noted by Marr [16], texture boundaries in images should have geometrical origins, such as discontinuities in surface shape or reflectance; subsequently, boundaries between projected textures should be highly localized. Use of the Gaussian smoothing filter allows the localized detection of such boundaries [7].

Suppose now that the image $t(x, y)$ can be partitioned into N textured regions $\mathcal{R}_1, \dots, \mathcal{R}_N$, where the boundaries between adjacent regions are distinct and continuous. Further suppose that each texture contains its most significant energy over a unique narrow range of frequencies:

$$t(x, y) = \sum_{n=1}^N t_n(x, y) z_n(x, y) + t'(x, y)$$

where

$$\begin{aligned} t_n(x, y) &= 2[c_n(x, y) \cos(2\pi F_n x_n) + s_n(x, y) \\ &\quad \cdot \sin(2\pi F_n x_n)] \\ &= a_n(x, y) \cos[2\pi F_n x_n + p_n(x, y)], \end{aligned} \quad (13)$$

$(x_n, y_n) = (x \cos \theta_n + y \sin \theta_n, -x \sin \theta_n + y \cos \theta_n)$, and where $z_n(x, y)$ is the set indicator function associated with region \mathcal{R}_n , i.e., $z_n(x, y) = 1$ for $(x, y) \in \mathcal{R}_n$ and $z_n(x, y) = 0$ otherwise. Since it is assumed that the image is partitioned, $z_n(x, y) \cdot z_i(x, y) = 0$ for any (x, y) when $n \neq i$. Thus, $t_n(x, y)$ is a narrowband 2-D image uniquely associated with region \mathcal{R}_n . The term $t'(x, y)$ is a residual image containing frequencies of insignificant energy and any dc component, i.e., $t(x, y) - t_n(x, y) = t'(x, y)$ for $(x, y) \in \mathcal{R}_n$. Thus, it is not necessary that $t_n(x, y)$ completely characterize region \mathcal{R}_n , only that $t_n(x, y)$ is the dominant component of $t(x, y)$ over \mathcal{R}_n .

Equation (13) does not necessarily imply that the region-modulated image $t_n(x, y) \cdot z_n(x, y)$ is narrowband, since the discontinuous nature of $z_n(x, y)$ must result in spectral rippling. However, assume that $z_n(x, y)$ is reasonably well-behaved (smooth). In addition, it is assumed that the textures $t_n(x, y)$ have similar ranges or peak-to-

peak contrast, i.e., that the envelopes $a_n(x, y)$ have similar average values. Given a set of 2-D Gabor filters $h_n(x, y) = g_n(x, y) \cdot \exp(2\pi j F_n x_n)$ (where $g_n(x, y) = g(x_n, y_n)$ with parameters λ_n and σ_n), having sufficiently narrow bandwidths, then $h_n(x, y) * t_i(x, y) \ll h_n(x, y) * t_n(x, y)$ for every $(x, y) \in \mathbb{R}^2$ and $i \neq n$, since $t_i(x, y)$ is deficient in those frequencies making region \mathcal{R}_n spectrally distinct. Of course, the number of Gabor filters applied may generally exceed the number of differing texture components in the image. Since the responses to those filters not tuned to any occurring texture may be assumed to be very small (provided that the set of filters used do not overlap too heavily in their frequency responses), they can be safely ignored. Thus, for $1 \leq n \leq N$

$$\begin{aligned} k_n(x, y) &= h_n(x, y) * t(x, y) \\ &\approx h_n(x, y) * t_n(x, y) z_n(x, y) \\ &= m_n(x, y) \exp[j\psi_n(x, y)]. \end{aligned}$$

The region-modulated narrowband images $t_n(x, y)$ $z_n(x, y)$ are converted into Gaussian-smoothed, complex-modulated versions of the region-modulated envelopes $a_n(x, y) z_n(x, y)$. The amplitude envelopes $m_n(x, y) \approx g_n(x, y) * a_n(x, y) z_n(x, y)$ may be recovered using (12). If $a_n(x, y) \approx a_n = \text{constant over } \mathcal{R}_n$, then $m_n(x, y) \approx A_n [g_n(x, y) * z_n(x, y)]$, which is a smoothed and weighted version of the indicator $z_n(x, y)$. However, this will occur only when $t_n(x, y)$ is very narrowband. Although the use of narrower channel filters will generally yield smoother responses, if $t_n(x, y)$ is not very narrowband then $m_n(x, y)$ may still vary considerably. In the absence of these effects, a complete segmentation may be obtained using the region assignment

$$(x, y) \in \mathcal{R}_n \quad \text{if } n = \arg \left\{ \max_{1 \leq i \leq N} [m_i(x, y)] \right\}. \quad (14)$$

If a complete set of Gabor filters is used, then (14) approximately corresponds to finding the maximum coefficient in the Gabor expansion at each image coordinate.

However, the envelopes $m_n(x, y)$ are often insufficiently smooth to allow for a consistent segmentation using (14). If any of the underlying textures $t_n(x, y)$ do not have sufficiently narrow bandwidths, leakage will occur. The effects of leakage can be reduced by postfiltering the channel amplitudes with Gaussian filters having the same shape as the corresponding channel filters, but greater spatial extents. The modified region assignment

$$\begin{aligned} (x, y) \in \mathcal{R} \\ \text{if } n = \arg \left\{ \max_{1 \leq i \leq N} [g_i(\gamma x, \gamma y) * m_i(x, y)] \right\}, \end{aligned} \quad (15)$$

($0 < \gamma < 1$) will then yield a smoother segmentation. In our experiments, scaling the postfilters by $\gamma = 2/3$ has proven to be effective in every case. Fig. 5 shows a block diagram of the overall process for segmenting an image using (15); of course, any given textural region may be empty.

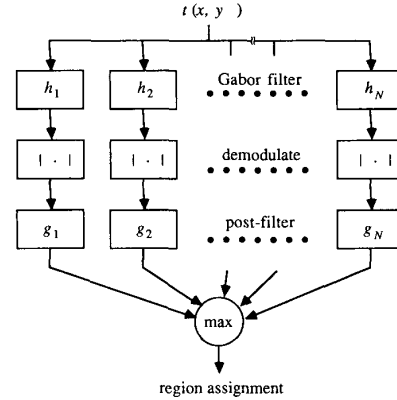


Fig. 5. Block diagram illustrating a system for segmenting an image into textured regions of different spatial frequency characteristics.

C. Recoding Region Information

There are three obvious motivations for encoding visual information: 1) to increase the signal-to-noise ratio (SNR) in the transmission and storage process; 2) to remove redundant or irrelevant information (compression); and 3) to represent information in a form that is useful for subsequent processing. Demodulation of the Gabor channel responses effectively removes high-frequency information describing the local structure of the image, although this information is at least partially embedded in the channel center frequencies. The information that remains describes the spatial extent of the textures to which the channels are tuned. By comparing this information from the various channels, textures may be segregated. However, it is possible that the envelope responses may be useful for other tasks, such as motion or shape analysis, or stereo. By recovering the slowly-varying channel envelopes, textural region information is made explicit.

As shown in Section II-D, Gabor filters must often be densely sampled to be responsive to the high-frequency textures which "carry" global region information. Once the channel responses have been demodulated, the low-frequency information which they contain can be more efficiently coded by resampling at a lower density prior to transmission or additional processing. The degree to which the sampling rate may be reduced can be determined from simple sampling theorem arguments.

Consider the Gabor filter (4) with parameters λ and σ . Since the linear radial bandwidth of the filter is of interest in using sampling theorem arguments, take $\theta = 0$ without loss of generality. The linear (half-peak) bandwidths of the filter along the horizontal and vertical directions are $B_x = 2\alpha/\pi\lambda\sigma$ and $B_y = 2\alpha/\pi\sigma$, respectively; the highest radial frequency within the passband is $F + \alpha/\pi\lambda\sigma$ (practically, F is always greater than $\alpha(1 - 1/\lambda)/\pi\sigma$). The half-peak bandwidths are used here for simplicity, since only the percent reduction in the sampling density is derived, rather than the actual sampling density required; the results are the same regardless of the bandwidth definition. Since it is easily shown that the achievable compression (in percent) is the same in any direction,

compression in the x -direction only will be considered. Specifically, the obtainable percent decrease in the sampling density of the amplitude envelope $m(x, y)$ is found.

Assuming that the filtered response $k(x, y) = h(x, y) * t(x, y)$ is approximately bandlimited to the passband of $h(x, y)$, the highest significant frequency in $m(x, y)$ is easily seen to be no more than $2\alpha/\pi\lambda\sigma$. The ratio of the necessary sampling density to accurately represent $m(x, y)$ to that to represent $k(x, y)$ is consequently $[2\alpha/\pi\lambda\sigma]/[F + \alpha/\pi\lambda\sigma] = 1 - 2^{-B}$. For a fairly large bandwidth ($B \approx 2.5$ octaves), the reduction in sampling density is about 18 percent; however, for a narrow filter ($B \approx 0.5$ octaves) the achievable reduction in the sampling density is larger than 70 percent. Thus, if a hierarchical filtering scheme is implemented as suggested in Section II-D, comparisons between the channel responses can be made more efficacious by first reducing sampling densities of the demodulated subimages to a common value.

D. Filter Selection

In order to effect an accurate segmentation of a textured image it is necessary to select a set of channel filters which will accomplish the task. The question is not critical if an extensive bank of filters is available in a parallel implementation, since the filters may then be selected so that the channel responses contain most of the information in the image. Methods for selecting a set of channel filters that adequately cover the spatial frequency domain have been considered in [17]. However, for conventional computer implementation, just a few values of the free parameters B , Ω , F , and θ can result in a very large set of filters. For example, only four values of each parameter used in all possible combinations results in 256 filters; accurate implementation of a "complete" Gabor expansion could entail a tremendous number of convolutions. If a relatively small number of filters are to be used to achieve segmentation, then some provision must be made for estimating the appropriate parameters to be used. This is a nontrivial problem having no simple solution.

In the examples, a simple peak-finding algorithm applied to the image power spectrum was used to guide the choice of the filter center frequencies. For a given textured image $t(x, y)$ having Fourier transform $T(u, v)$, the spectral peaks were defined as the maxima of $|T(u, v)|^2$ within square neighborhoods of a set size over the positive ($u > 0$) half-plane of frequencies. The peak locations were listed in polar form (F, θ) for $\theta \in (-\pi/2, \pi/2]$. Since the examples provided later are generally composed of images containing two discriminable textured regions, the number of Gabor filters used to effect each segmentation was limited to two. While a larger set of channel filters could be used in the segmentations, the use of only two responses demonstrates that the Gabor filtering/demodulation model is sufficiently powerful to achieve segmentation using only a subset of the information within each texture.

The filter parameters were chosen using a limited

amount of human intervention. For strongly oriented textures, the most significant spectral peak along the orientation direction was chosen. For periodic textures, the lower fundamental frequency was chosen. Finally, for nonoriented textures, the center frequencies were chosen from the two largest maxima. The radial frequency bandwidths B were chosen from among the values $\{0.7, 1.0, \text{ and } 1.3\}$ depending on the filter center frequency: smaller bandwidths were used at higher center frequencies. Except where specifically noted, the Gabor filters were defined with aspect ratio $\lambda = 1$ for simplicity. Specification of the parameters F , θ , and B constrains the remaining parameters Ω and σ , since for $\lambda = 1$, $\sigma = (\alpha/\pi F)(2^B + 1)/(2^B - 1)$ and $\Omega = 2 \tan^{-1} [(2^B - 1)/(2^B + 1)]$. The corresponding orientation bandwidths are $\{27^\circ, 37^\circ, \text{ and } 46^\circ\}$.

While the method for selecting filters outlined above is simple and convenient, a more elegant peak-finding approach is easily defined. Simply, the best match (among Gabor filters of a given shape and aspect ratio) to the image may found by convolving the image spectrum with a Gaussian. Thus, the local maxima of $|G_{\sigma, \lambda}(u, v) * T(u, v)|$ over half of the (u, v) plane may be found, where $G_{\sigma, \lambda}(u, v)$ is given by (3) with $U' = V' = 0$. A set of filters can then be selected by thresholding the maximal values to obtain a set of center frequencies (F, θ) . This approach has the added advantages of smoothing the image frequencies and providing a "best" match to the Gabor filter frequency responses.

E. Segmentation Examples

The examples given in this section illustrate the segmentation models (14) and (15) applied to both synthetically generated and naturally-occurring or man-made textures digitized to 256 gray levels with 256×256 pixels spatial resolution. The Gabor filtering was accomplished using a floating-point 2-D FFT implementation. Since a Gabor filter may have a very slight response to a level irradiance, a small threshold was applied to each filtered and demodulated image. At those coordinates where no envelope exceeded the threshold, no region assignment was given. Several examples were created by embedding one Brodatz [18] texture inside of another. In each case the textures were scaled to have the same range of gray levels in an effort to equalize their ranges. In all the examples the Gabor filter center frequency and frequency bandwidths are specified in units of cycles/image and octaves, respectively. Whenever the segmentation model (15) was applied, the spatial extent of the postfiltering Gaussian was increased by an amount $1/\gamma = 3/2$.

Fig. 6 depicts an image of two textures composed of dark lines of different orientations against a bright background. The periodic nature of the textures allowed for very accurate identification of the texture boundaries, since the "textures" contain very narrow ranges of frequencies. Segmentation was obtained using (14), since the computed channel amplitudes were very smooth.

The aperiodic placement of the oriented line segments

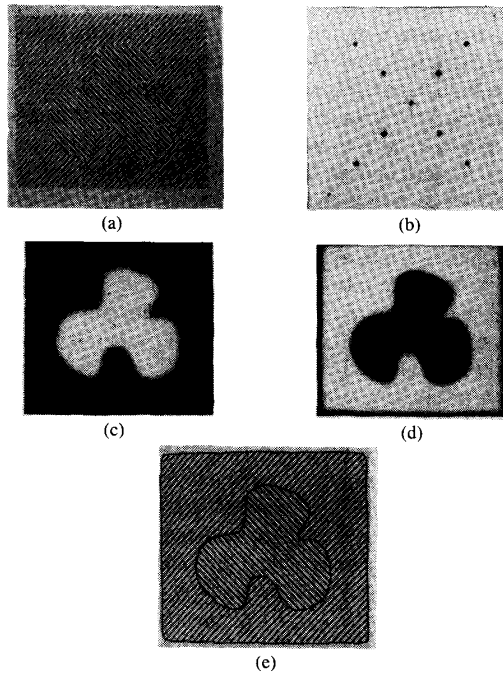


Fig. 6. Segmentation using two filters of different orientations: (a) original; (b) Fourier transform; (c), (d) channel amplitudes recovered (F, θ, B) = (50, 0.7, 42°) and (50, 0.7, -43°); (e) computed texture boundaries.

composing the textures in Fig. 7 (photographed from Fig. 5 in [19]) resulted in considerable fluctuation in the computed channel amplitudes. Application of (14) consequently resulted in the misclassification of a number of small regions. By postfiltering the envelope, a more consistent segmentation was obtained using (15).

Fig. 8 illustrates a similar texture, taken from Marr [16, Fig. 2-34], only in this case the inner texture is modulated by a triangular region with sharp corners. Since the boundary of the inner texture is not smooth, the spectrum of the associated region indicator must be spread. However, segmentation using (15) fully captured the shape of the region boundaries and allowed successful classification of all the line segments composing the textures.

It is interesting to note that in the last example, the boundary computed follows a tortuous path around the the oriented segments, while perception of the inner textured region is strongly that of a triangular shape. While this may be accomplished by higher-level associative processes, a closer approximation to a triangle could also be obtained by using either narrower channel filters, or a spatially broader postfilter.

Fig. 9 shows a composition of two Brodatz textures (handmade paper, D57, and netting, D34). This image has high spectral energy at several orientations, but the netting texture is finer, so there is more energy at higher spatial frequencies than in the paper texture. The segmentation was obtained using (15) with channel filters of the same orientation, but different center frequencies.

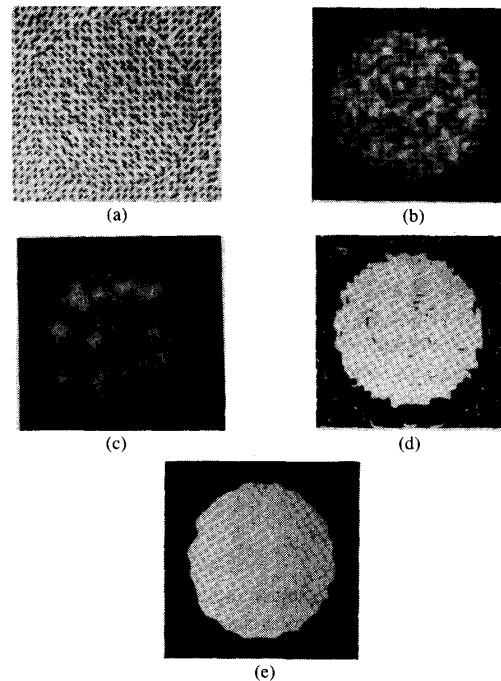


Fig. 7. Segmentation of oriented textures. (a) Original; (b) computed channel amplitude (F, B, θ) = (40, 1, 45°); (c) segmentation using (14), (F, B, θ) = (40, 1, 45°) and (40, 1, -45°); (d) Result of postfiltering (c); segmentation using (15), also shown overlaid on (a).

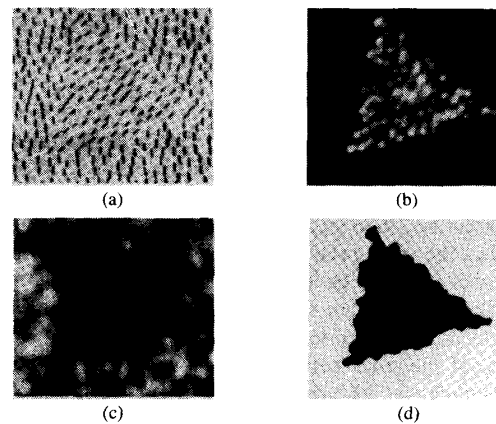


Fig. 8. Segmentation of oriented textures separated by a triangular boundary. (a) Original; (b) unfiltered amplitude response (F, θ, B) = (59, 1.0, -48°); (c) postfiltered amplitude response (F, θ, B) = (42, 1.0, 5°); (d) segmentation using (15). Note that the images in (b), (c) are "inverted" since they derive from different channels.

Fig. 10 is an image similar to that used by Marr in his description of the full primal sketch [16, Fig. 2-7]. The grouping of elements was obtained in a single computational stage using two Gabor filters. This example suggests that perceptual groupings may be effected using a computational framework similar to that we propose for textural segmentation. However, it is not clear that the assignment of boundaries around the groupings has any

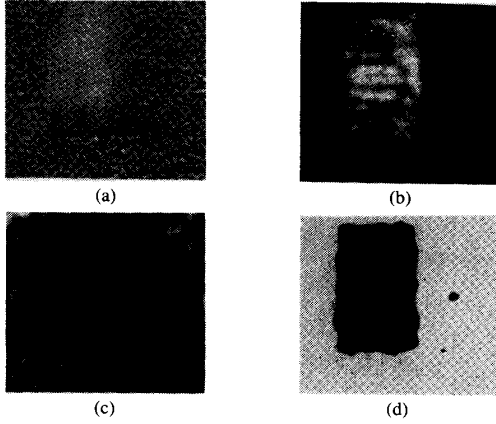


Fig. 9. Segmentation based on granularity. (a) Comparison of Brodatz textures handmade paper, D57, and netting, D34; (b) computed amplitude response $(F, \theta, B) = (80, 0.7, 89^\circ)$; (c) computed amplitude response $(F, \theta, B) = (40, 1.3, 89^\circ)$; (d) segmentation using (15).

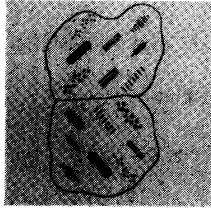


Fig. 10. Emulation of perceptual grouping using (15). $(F, \theta, B) = (16, 1.3, 45^\circ)$ and $(16, 1.3, -41^\circ)$.

real meaning, since it was observed that using different bandwidths produced slight but noticeable shifts in the computed boundaries.

The final example (Fig. 11) illustrates the segmentation of a herringbone material, the top half of which has been shifted to create a sharp textural boundary. It is interesting to note the shape of the computed boundaries. Instead of creating a set of rectangles, the boundaries often followed a path more in agreement with perception; the textures seem to be continuous across the corners.

IV. TEXTURE SEGMENTATION BY CHANNEL PHASE DEMODULATION

The amplitude envelope $m(x, y)$ of the complex-valued Gabor-filtered image $k(x, y) = h(x, y) t(x, y)$ only partially describes the nature of the image $t(x, y)$ over the frequencies to which the filter is tuned; the remaining information is contained in the phase envelope $\psi(x, y)$. In this section, methods for extracting useful information

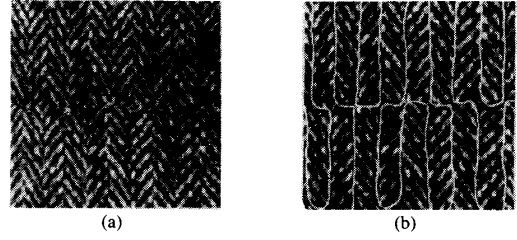


Fig. 11. Segmentation of herringbone material. (a) Original; (b) boundaries computed using (15) with $(F, \theta, B) = (25, 1.0, 34^\circ)$ and $(25, 1.0, -38^\circ)$.

from $\psi(x, y)$ are considered. In Section IV-A, the type of information which can be obtained from the channel phase responses is described. It is found that meaningful information can be obtained by computing the variations in the texture phase, where small variations correspond to regions of smooth texture phase, and large variations correspond to sudden transitions from one dominant phase to another. By localizing the large variations, the transitions between texture phases can be marked. This is accomplished by finding the zero crossings (ZC's) of a second-order differential operator (Laplacian) applied to the channel phase responses. In Section IV-B, a framework for detecting texture phase transitions is given. The detection of phase transitions is postponed until after the amplitude demodulation segmentation process, since differences in texture phase are generally meaningful only within a single texture type. In Section IV-C, examples are given of the process applied to a number of interesting textures.

A. Information in the Channel Phases

Consider again the narrowband image $t(x, y)$ given by (10) and the corresponding Gabor filtered image $k(x, y)$ given by (11). The smoothly-varying phase envelope of $k(x, y)$ is

$$\psi(x, y) = 2\pi Fx' - \tan^{-1} \left[\frac{g(x', y') * s(x, y)}{g(x', y') * c(x, y)} \right],$$

which will contain discontinuities which are approximately periodically spaced if the principal value of the inverse tangent is computed; to compute $\psi(x, y)$ without discontinuities, phase unwrapping is required. Fortunately, this difficulty can be avoided since the information in $\psi(x, y)$ at any given coordinate (x, y) is only meaningful relative to the phase at neighboring coordinates. A more meaningful source of information is the spatial variation in phase described by the directional instantaneous frequencies

$$\begin{aligned} \psi_x(x, y) &= D_x \psi(x, y) = \frac{[k_x(x, y) \cdot D_x k_s(x, y) - k_s(x, y) \cdot D_x k_c(x, y)]}{m^2(x, y)} \\ &= 2\pi F \cos \theta - \frac{[g(x', y') * c(x, y)] \cdot [D_x g(x', y') * s(x, y)] - [g(x', y') * s(x, y)] \cdot [D_x g(x', y') * c(x, y)]}{m^2(x, y)} \end{aligned}$$

and

$$\begin{aligned}\psi_y(x, y) &= D_y \psi(x, y) = \frac{[k_c(x, y) \cdot D_y k_s(x, y) - k_s(x, y) \cdot D_y k_c(x, y)]}{m^2(x, y)} \\ &= 2\pi F \sin \theta - \frac{[g(x', y') * c(x, y)] \cdot [D_y g(x', y') * s(x, y)] - [g(x', y') * s(x, y)] \cdot [D_y g(x', y') * c(x, y)]}{m^2(x, y)},\end{aligned}\quad (16)$$

where $D_x = \partial/\partial x$ and $D_y = \partial/\partial y$. The gradient magnitude $|\nabla \psi(x, y)| = \sqrt{\psi_x^2(x, y) + \psi_y^2(x, y)}$ will take large values near sharp changes in either the dominant texture phase or dominant texture frequencies. Clearly, phase unwrapping is unnecessary using (16), although the partial derivatives of $k_c(x, y)$ and $k_s(x, y)$ must be computed. This can be accomplished either by using a truncated Taylor's series discrete approximation to the differential operators [24], or by using the following more accurate approach. The modified "daisy petal" filters

$$\begin{aligned}h_x(x, y) &= D_x g(x', y') \cdot \exp(2\pi j F x') \\ &= h_{c,x}(x, y) + j h_{s,x}(x, y)\end{aligned}$$

and

$$\begin{aligned}h_y(x, y) &= D_y g(x', y') \cdot \exp(2\pi j F x') \\ &= h_{c,y}(x, y) + j h_{s,y}(x, y)\end{aligned}\quad (17)$$

are complex sinusoidal gratings modulated by x - and y -derivatives of Gaussians, respectively. Similar modifications of the more general Gabor filter (1) may also be defined. The filters $h_x(x, y)$ and $h_y(x, y)$ are not Gabor filters, but they are approximately analytic if the Gabor filter $h(x, y)$ is narrowband. Hence, the real and imaginary responses of the filters $h_x(x, y)$ and $h_y(x, y)$ are

$$\begin{aligned}k_{c,x}(x, y) &= h_{c,x}(x, y) * t(x, y) \\ &\approx [D_x g(x', y') * c(x, y)] \cos(2\pi F x') \\ &\quad + [D_x g(x', y') * s(x, y)] \sin(2\pi F x') \\ k_{s,x}(x, y) &= h_{s,x}(x, y) * t(x, y) \\ &\approx [D_x g(x', y') * c(x, y)] \sin(2\pi F x') \\ &\quad - [D_x g(x', y') * s(x, y)] \cos(2\pi F x')\end{aligned}$$

and

$$\begin{aligned}k_{c,y}(x, y) &= h_{c,y}(x, y) * t(x, y) \\ &\approx [D_y g(x', y') * c(x, y)] \cos(2\pi F x') \\ &\quad + [D_y g(x', y') * s(x, y)] \sin(2\pi F x') \\ k_{s,y}(x, y) &= h_{s,y}(x, y) * t(x, y) \\ &\approx [D_y g(x', y') * c(x, y)] \sin(2\pi F x') \\ &\quad - [D_y g(x', y') * s(x, y)] \cos(2\pi F x').\end{aligned}$$

It is easily verified that

$$\begin{aligned}\psi_x(x, y) &= 2\pi F \cos \theta \\ &\quad - \frac{[k_c(x, y) \cdot k_{s,x}(x, y) - k_s(x, y) \cdot k_{c,x}(x, y)]}{m^2(x, y)}\end{aligned}$$

and

$$\begin{aligned}\psi_y(x, y) &= 2\pi F \sin \theta \\ &\quad - \frac{[k_c(x, y) \cdot k_{s,y}(x, y) - k_s(x, y) \cdot k_{c,y}(x, y)]}{m^2(x, y)}.\end{aligned}\quad (18)$$

This approach allows the simultaneous calculation of all quantities required to compute the phase variations. The quantities $\psi_x(x, y)$ and $\psi_y(x, y)$ describe the smoothed variation in texture phase; by suppressing the nonmaxima of the gradient magnitude $|\nabla \psi(x, y)|$, the discontinuities in the texture phase may be localized, since local maxima in $|\nabla \psi(x, y)|$ will correspond to sharp transitions in the unwrapped phase $\psi(x, y)$. While these maxima can be found using local search techniques, as a simple alternative the predicate $\nabla^2 \psi(x, y) = 0$ may be applied to mark the abrupt phase transitions. Noting that

$$\begin{aligned}\nabla^2 \psi(x, y) &= \{[2\pi F \cos \theta - \psi_x(x, y)] \cdot D_x m^2(x, y) + [2\pi F \sin \theta - \psi_y(x, y)] \cdot D_y m^2(x, y) \\ &\quad + [g(x', y') * s(x, y)] \cdot [\nabla^2 g(x', y') * c(x, y)] \\ &\quad - [g(x', y') * c(x, y)] \cdot [\nabla^2 g(x', y') * s(x, y)]\} / m^2(x, y),\end{aligned}$$

and since

$$D_x m^2(x, y) = 2[k_c(x, y) k_{c,x}(x, y) + k_s(x, y) k_{s,x}(x, y)]$$

and

$$D_y m^2(x, y) = 2[k_c(x, y) k_{c,y}(x, y) + k_s(x, y) k_{s,y}(x, y)]$$

and using (18), then

$$\begin{aligned} \nabla^2 \psi(x, y) = & \{ [k_c^2(x, y) - k_s^2(x, y)] \cdot [k_{c,x}(x, y) k_{s,x}(x, y) + k_{c,y}(x, y) k_{s,y}(x, y)] \\ & + k_c(x, y) k_s(x, y) [k_{s,x}^2(x, y) - k_{c,x}^2(x, y) + k_{s,y}^2(x, y) - k_{c,y}^2(x, y)] \\ & + [g(x', y') * s(x, y)] \cdot [\nabla^2 g(x', y') * c(x, y)] \\ & - [g(x', y') * c(x, y)] \cdot [\nabla^2 g(x', y') * s(x, y)] \} / m^2(x, y). \end{aligned} \quad (19)$$

To facilitate computation of the terms containing Laplacians, define the modified filter

$$\begin{aligned} h_L(x, y) &= \nabla^2 g(x', y') \cdot \exp(2\pi j F x') \\ &= h_{c,L}(x, y) + j h_{s,L}(x, y) \end{aligned}$$

which is also approximately analytic for parameter values typical of those used here. The real and imaginary responses of this filter to image $t(x, y)$ are given by

$$\begin{aligned} k_{c,L}(x, y) &= h_{c,L}(x, y) * t(x, y) \\ &\approx [\nabla^2 g(x', y') * c(x, y)] \cos(2\pi F x') \\ &\quad + [\nabla^2 g(x', y') * s(x, y)] \sin(2\pi F x') \end{aligned}$$

and

$$\begin{aligned} k_{s,L}(x, y) &= h_{s,L}(x, y) * t(x, y) \\ &\approx [\nabla^2 g(x', y') * c(x, y)] \sin(2\pi F x') \\ &\quad - [\nabla^2 g(x', y') * s(x, y)] \cos(2\pi F x'). \end{aligned}$$

Substituting these quantities into (19) and using (11) yields the following expression for the Laplacian of the channel phase:

$$\begin{aligned} \nabla^2 \psi(x, y) = & \{ k_c(x, y) k_{s,L}(x, y) - k_s(x, y) k_{c,L}(x, y) \\ & + [k_c^2(x, y) - k_s^2(x, y)] \cdot [k_{c,x}(x, y) k_{s,x}(x, y) + k_{c,y}(x, y) k_{s,y}(x, y)] \\ & + k_c(x, y) k_s(x, y) [k_{s,x}^2(x, y) - k_{c,x}^2(x, y) + k_{s,y}^2(x, y) - k_{c,y}^2(x, y)] \} / m^2(x, y). \end{aligned} \quad (20)$$

Detecting phase discontinuities subsequently reduces to locating ZC's in the Laplacian of channel phase, without the need for phase unwrapping or local derivative approximations; the various calculations instead involve frequency-shifted derivatives of Gaussian functions, or derivatives of Gaussian-smoothed versions of the texture

envelopes. It is well-known that Gaussian smoothing regularizes the computation of discrete approximation of spatial derivatives [21]; thus, the derivative terms in (20) may be regarded as accurate and stable in a discrete implementation.

It is interesting to note that the filter $h_L(x, y)$ is actually a frequency-shifted version of the well-known Laplacian-of-a-Gaussian (LoG) edge detector [22]. The filter may subsequently be closely approximated as a difference of two Gabor filters having the same orientation, center frequency, and aspect ratio, but different space constants (the

ratio 1 : 1.6 yields a good approximation [16]):

$$\begin{aligned} h_L(x, y) &\approx g(x', y'; \sigma, \lambda) \cdot \exp(2\pi j F x') \\ &\quad - g(x', y'; 1.6\sigma, \lambda) \cdot \exp(2\pi j F x') \\ &= [g(x', y'; \sigma, \lambda) - g(x', y'; 1.6\sigma, \lambda)] \\ &\quad \cdot \exp(2\pi j F x') \\ &\approx \nabla^2 g(x', y'; \sigma, \lambda) \cdot \exp(2\pi j F x'), \end{aligned}$$

where $g(x', y'; \sigma, \lambda)$ is a 2-D Gaussian function with space constant σ and aspect ratio λ . As such, the operator is subject to accuracy limitations [23], but the detected discontinuities can be expected to be continuous and smooth [21].

It is appropriate at this point to consider the effect that an abrupt change in texture phase has on the computed channel envelopes. A phase discontinuity in an otherwise ideal narrowband texture is modeled as:

$$\begin{aligned} t(x, y) &= 2A \{ q(-x) \cdot \cos[2\pi F x' - P] \\ &\quad + q(x) \cdot \cos[2\pi F x' - P + \delta] \}, \end{aligned}$$

where $q(x)$ is the Heaviside unit step: $q(x) = 1$ for $x > 0$, $q(x) = 0$ for $x < 0$, $q(0) = 1/2$, and $|\delta| < \pi$ is a difference in texture phase across the vertical line $x = 0$. The amplitude response of a Gabor filter $h(x, y) = g(x, y) \cdot \exp(2\pi j F x')$ to $t(x, y)$ (letting $\lambda = 1$ for simplicity)

is then

$$m(x, y) = |A| \cdot [1 - 4\Phi(x/\sigma) \Phi(-x/\sigma) \cdot \sin^2(\delta/2)]^{1/2}.$$

Thus, $m(x, y)$ achieves a unique *minimum* along the line $x = 0$, whereas, for values of $|x|$ that are not small, $m(x, y) \approx A$. Moreover, since $m(0, y) = A \cdot |\cos(\delta/2)|$, the attenuation near the line $x = 0$ is greatest for $\delta = 180^\circ$, the maximum possible phase offset. It is interesting to note that the attenuation does not depend on the orientation θ of the dominant carrier frequency of the texture relative to the phase discontinuity. This implies that sudden changes in texture phase along any orientation can lead to local attenuations of the channel amplitude envelopes, which partially explains the "worminess" evident in the amplitude envelopes of many of the examples in Section III, and is one reason why postfiltering is needed to smooth the amplitude responses. Although the attenuation may obstruct amplitude-based segmentation using (14), by using (15) a successful segmentation can be obtained.

The directional variations in the phase of $h(x, y) * t(x, y)$ are given by

$$\psi_x(x, y) = 2\pi F \cos \theta - g(x) \sin \delta / [1 - 4\Phi(x/\sigma) \cdot \Phi(-x/\sigma) \sin^2(\delta/2)]$$

and

$$\psi_y(x, y) = 2\pi F \sin \theta,$$

where $g(x) = D_x \Phi(x/\sigma)$ is a 1-D symmetric Gaussian function with unit area and space constant σ . The gradient magnitude $|\psi(x, y)|$ is maximized at $x = 0$ since $g(x, y)$ and $\Phi(x/\sigma) \Phi(-x/\sigma)$ both achieve maxima there, and for the maximal phase shift $\delta = \pi$ (the maxima in this case is an isolated singularity along the line $x = 0$; for less idealized discontinuities, the maxima will be achieved smoothly). The Laplacian of phase is given by

$$\nabla^2 \psi(x, y) = D_{xx} \psi(x, y) = \left\{ (x/\sigma^2) g(x) \sin \delta [1 - 4\Phi(x/\sigma) \Phi(-x/\sigma) \sin^2(\delta/2)] + 4g(x, y) g(x) \sin \delta \sin^2(\delta/2) [2\Phi(x/\sigma) - 1] \right\} / [1 - 4\Phi(x/\sigma) \Phi(-x/\sigma) \sin^2(\delta/2)]^2,$$

which uniquely *crosses* zero at $x = 0$, since both terms in the denominator are always of the same sign, although $\nabla^2 \psi(x, y) = 0$ everywhere if $\delta = 0$ (if no phase discontinuity exists). Thus, texture phase discontinuity detection is an increasing function of δ for $\delta \in [0, \pi]$.

B. Segmentation Using Channel Phase

By locating discontinuities in texture phase, information not available in the amplitude envelopes of the textures is made explicit. In doing so it is less useful to compare the phases of the different channel outputs than it is to locate the phase discontinuities within each channel output. The assumption underlying this is that texture phase discontinuities should arise from offset surface textures which are indiscriminable when viewed separately

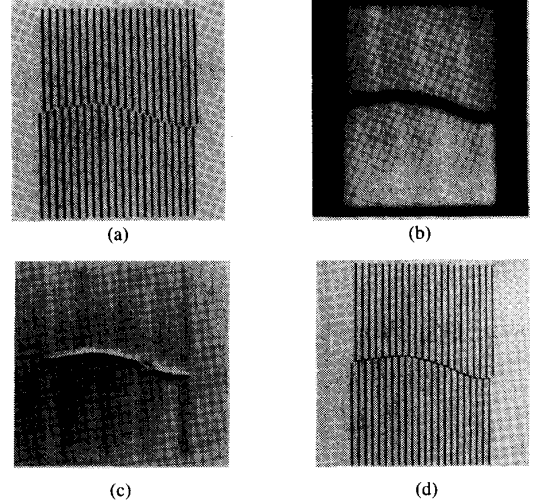


Fig. 12. Segmentation of illusory curve by phase discontinuity detection. (a) Original; (b) amplitude envelope of a Gabor filter tuned to the texture (F, θ, B) = (32, 0.7, 0°); (c) Laplacian of channel phase; (d) significant ZC's superimposed on the original.

(have the same dominant frequency components), but which should be interpreted as belonging to separate regions when viewed together. The texture in Fig. 12 illustrates this possibility. There is a distinct impression of two separate regions which are adjacent along a smooth curve. The illusory contour is clearly a byproduct of the difference in spatial phase of the texture in the two regions.

Consider again the model given in (10)–(15) for segmenting an image $t(x, y)$ by comparison of the channel amplitude responses. Along with each channel output $k_n(x, y) = h_n(x, y) * t(x, y)$ the Laplacian of phase $\nabla^2 \psi_n(x, y)$ is computed using (20). In order to detect phase discontinuities in the dominant texture frequencies,

define the combined phase response

$$\Psi(x, y) = \nabla^2 \psi_n(x, y) \quad \text{if } n = \arg \left\{ \max_{1 \leq i \leq N} [g_i(\gamma x, \gamma y) * m_i(x, y)] \right\} \quad (21)$$

[see (16)]. Detection of the significant phase discontinuities then reduces to locating the ZC's $\Psi(x, y) = 0$. Alternatively, the phase boundaries may be located via $\{(x, y): \nabla^2 \psi_n(x, y) = 0\} \cap \{(x, y): (x, y) \in \mathcal{R}_n\}$. While it is difficult to analyze the sources of error which may arise in this model, some of the weaknesses of the amplitude demodulation approach can be expected to propagate to (21), since the significant channel responses are still defined by the model (14), (15).

C. Segmentation Examples

In segmenting textures by finding discontinuities in channel phase via (21), the filter parameters were determined using the methods discussed in Section III-D. In configuring the filters $h_x(x, y)$, $h_y(x, y)$, and $h_L(x, y)$, the same relevant parameters were used as in the corresponding Gabor filter $h(x, y)$.

Fig. 12 depicts an image containing two apparently separate regions separated by a smooth illusory curve. The line patterns on each side of the texture are identical except for a difference in phase; if viewed separately under identical viewing conditions, the textures composing the two regions would be indistinguishable. While the model developed in the preceding provides a method for localizing the phase discontinuity using a single channel filter, the boundary between the textures may also be interpreted as composing a texture separate from those on either side. Thus, the amplitude response of an appropriately tuned Gabor filter is also shown; as expected, the response vanishes at the phase discontinuity, although Gaussian post-filtering smooths the response across the boundary. The magnitude of the Laplacian of phase obtained using the Gabor filter (and the modified filters h_x , h_y , and h_L of the same orientations and spatial extent) was very small except near the discontinuity, where a ZC contour occurred. The significant ZC's (scaled by the gradient magnitude of the Laplacian $|\nabla \cdot \nabla^2 \psi(x, y)|$ and thresholded) are shown superimposed on the original image. Although a different segmentation could possibly be obtained by applying a different set of Gabor filters oriented along the direction of the discontinuity (responsive to the offset line terminations), in our view phase segmentation using (21) supplies a more physically meaningful separation of the textures; moreover, the perception of two separate regions is stronger than of three.

Fig. 13 is composed of two textures composed of short line segments oriented along the horizontal and vertical directions of the same lengths and periodicities. However, in the outer texture the horizontal and vertical line segments form the shape “+”, while in the inner texture they form the shape “L”. Since both the inner and outer textures can be interpreted as a sum of identical textures (other than a phase difference), the textures have identical amplitude spectra. The Laplacian of phase response computed using (20) (tuned to the vertical line segments) yields a very strong response to the change in phase. The scaled ZC's obtained using (21) are shown superimposed on the original. It is interesting to note that, although the segmentation agrees closely with the perceived boundary, there is some deformation near the corners of the square. This effect is very similar to that observed by Berzins [23] in his analysis of Laplacian operators, and can probably be attributed to the use of the isotropic Laplacian operator rather than directional differential operators. An interesting, but much more expensive extension would be to compute the ZC's of the directional derivatives along the gradient orientation, similar to [24].

Finally, Fig. 14 shows the Brodatz texture D20, French canvas; the left and right halves have been offset from one

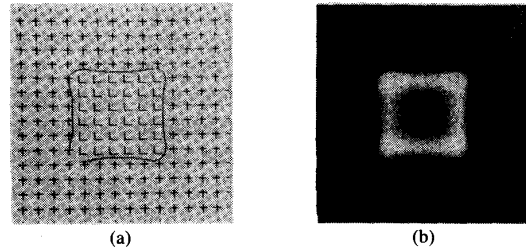


Fig. 13. Detection of phase discontinuity between textures having identical amplitude spectra. (a) Original; (b) Laplacian of phase (F, θ, B) = (16, 0.7, 0°). The computed boundary is overlaid on (a).

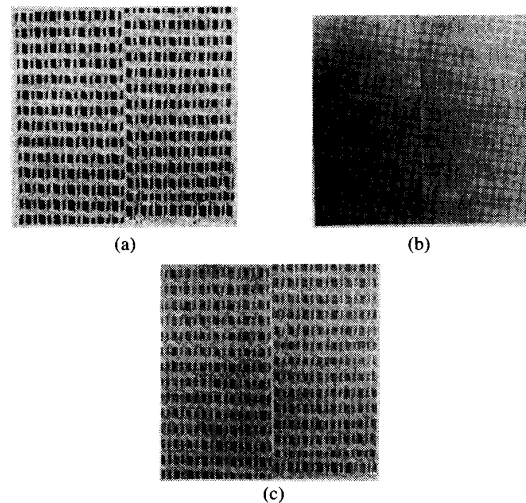


Fig. 14. Segmentation of Brodatz texture D20, French canvas. (a) Original; (b) Laplacian of phase (F, θ, B) = (14, 1.0, 86°); (c) significant ZC's overlaid on the original.

another to introduce a phase discontinuity. In this case the ZC's Laplacian of phase again provided a close approximation of the separating boundary, although the ZC contour diminished near the bottom of the image. An interesting effect is the change in *sign* of $\nabla^2 \psi(x, y)$ along the phase discontinuities [see also Figs. 12(c) and 13(b)].

V. SURVEY

This section surveys research in computational vision, visual physiology, and visual psychophysics that directly pertains to this work. Section V-A casts the Gabor multiple channel filtering model in the light of other spatial filtering and texture frequency analysis approaches. Section V-B considers relevant research in visual physiology and psychophysics and the potential role that the multiple Gabor channel model may play in unifying feature-based and frequency-based computational texture analysis paradigms.

A. Spatial Filtering Approaches for Texture Analysis

Numerous computational approaches for texture discrimination/segmentation have been advanced that can broadly be classified as involving spatial filtering. Methods utilizing frequency- and/or orientation-selective linear filters are of relevance here. Surveys and comparisons

of statistical, autocorrelation, structural, and co-occurrence techniques that in some sense involve localized image filtering are amply available, e.g., [25]–[31]. Fourier analysis for textural analysis in machine vision dates at least from Bajcsy [32] who uses windowed subimages in analyzing texture and texture shape [33]. Several authors use the outputs of multiple (space- and/or frequency-) localized filters as features/attributes for classifying textures via clustering [34]–[40]; Faugeras [34] describes an early pattern recognition based approach. Laws [35] extends the concept using multiple feature templates or filters in conjunction with a clustering algorithm. Some of the templates used by Laws are interesting in that they resemble space-limited (truncated) discrete Gabor filters. Pietikainen [36] reports a similar approach using the responses of multiple LoG filters for classification purposes. Segmentation/discrimination algorithms are reported for detecting periodic and oriented image regions by partitioning the image spectrum into bins [37] or by locating spectral peaks [38]. Energy features computed from the image spectra are used to classify textures in a pattern space. Knutson and Granlund [39] propose the use of image quadrature filters for texture analysis. While the Hilbert transform approach is similar in spirit to the multiple Gabor channel filter model, the filters do not share the same localization properties as the Gabor filters. Coggins and Jain [40] devise a multichannel segmentation algorithm using both orientation-sensitive and moderately narrow channel bandwidths. Again, segmentation is augmented by clustering the channel outputs. Wilson and Spann [41] use multiple low-uncertainty (prolate) filters in conjunction with hierarchical (quadtree) feature descriptions and a clustering algorithm. Clark and Bovik [4], [6] and Bovik *et al.* [7] use Gabor filters for segmentation purposes whereas Turner [5] demonstrates their utility for discrimination. In these approaches segmentation/discrimination is accomplished without clustering or classical pattern classification techniques. In view of the similarity of the filtering structures proposed here and in [41] clustering processes could be expected to be effectively applied to multiple Gabor channel outputs. However, the focus of this paper is to investigate fundamental aspects of segmentation rather than trying to optimize pixel assignments by complicated clustering techniques or other higher-level arbitration procedures. Nevertheless, the hierarchical approach taken in [41] suggests computationally convenient multiscale techniques for segmentation using a “Gabor pyramid” [42]. Daugman [43] recently describes a three-layer neural network for computing complete discrete Gabor expansions and uses the coefficients for a number of tasks, including segmentation. Rearick [44] and Voorhees and Poggio [45] attempt to unify feature detection and segmentation by computing densities of “textons” [1]. Limited results are obtained, primarily because of the lack of generally agreed upon (rigorous) definitions of textons and the complications involved in analyzing their densities. However, research along this direction is not inconsistent with multiple channel models which use spatially localized filters: the use of low-uncertainty filters can be

viewed as optimizing a tradeoff between feature detection and spatial frequency selection [5], [7].

B. Visual Physiology and Psychophysics

Computational models of low-level biological vision often derive from physiological and psychophysical investigations of the structure and function of neural receptive fields in the post-retinal ganglion cells, the lateral geniculate nucleus, and the striate cortex. A classic example of conflicting theories arising from such experiments are the (until recently) apparently incompatible *feature detection* [46], [47] and *Fourier decomposition* [48]–[50] models of visual processing. Although numerous experiments have provided further support for both the feature detector and Fourier decomposition models, neither interpretation has been shown to adequately describe cortical processing [51]. A possible unification of the models may develop from recent models of certain simple cell receptive fields and in the functional version of the uncertainty principle. A quantitative model of the highly oriented simple cell receptive fields was shown by Marcelja [13] to agree closely with cortical channel characteristics. The receptive field profiles of these simple cells closely approximate 1-D Gabor functions, the real (cosine) and imaginary (sine) components being associated with the bar and edge detecting cells, respectively. Daugman [9], [52] proposed the 2-D Gabor receptive field model, extending Gabor's result [8] by showing that the 2-D Gabor filters are the unique minimum-uncertainty 2-D filters. Subsequent researchers have independently confirmed the approximate validity of the Gabor receptive field model [53]–[55] while others dispute the Gabor model as too specific; other (similar) mathematical models may apply equally well, e.g., frequency-shifted prolate functions [56]. The point is, however, that the Gabor model provides an extremely useful and tractable representation of the simple cells, particularly in developing computational models of visual tasks. It seems unlikely that the weights in the spatial summations across the receptive fields uniformly adhere with great precision to any particular mathematical model across individuals or even across neurons. It probably suffices that the receptive fields possess spatio-frequency localization conveniently expressed as low uncertainty. Low-uncertainty channel filters represent a reasonable compromise between previous theories of visual neural function, since they are both highly responsive to localized features, and highly frequency- and orientation-selective.

Although the texture analysis model put forth here is motivated in large part by the Gabor receptive field model, it is advocated here only as a convenient computational formalism, not as a probable model of cortical processing; too many issues remain unresolved. Nevertheless, it is not unreasonable to suggest that the simple cells perform an important early stage of processing in textural perception. Thus, it is of interest to compare some of the parameters and filter definitions used here with those measured in biological systems. Three (half-peak) frequency bandwidths $B = (0.7, 1.0, \text{ and } 1.3 \text{ octaves})$ were used in

the segmentation examples in Sections III-E and IV-C; the corresponding (half-peak) orientation bandwidths are $\Omega \approx (27^\circ, 37^\circ, \text{ and } 46^\circ)$. By comparison, orientation bandwidths in the simple cells average about 45° (with considerable variation), while frequency bandwidths range from the very narrow (~ 0.5 octaves) to very fairly broad (~ 2.5 octaves). Thus, the filters used in the simulations are roughly in accordance with experimental evidence. While the complex form of the Gabor function is a convenient representation, biological implementation obviously proceeds by separate computation of the (real-valued) component functions. Representation of opposite response polarities apparently involves separate computation of ON/OFF components corresponding to excitatory/inhibitory responses. The manner in which the responses of receptive fields having different phase characteristics are combined remains unclear. While the quadrature or Hilbert transform interpretation is computationally convenient, there is only unconfirmed evidence that neighboring simple cells tuned to the same orientation and spatial frequency bands occur in quadrature pairs [54], [57]. Moreover, the measured receptive field profiles are not always symmetrical (or antisymmetrical): asymmetric responses possibly being modeled by Gabor functions with an additional phase shift [13]: $h(x, y) = g_\sigma(x, y) \cdot \exp[2\pi j(Ux + Vy + \varphi)]$ with respective real and imaginary components

$$g_\sigma(x, y) \cdot \left\{ \cos 2\pi\varphi \cos [2\pi j(Ux + Vy)] - \sin 2\pi\varphi \sin [2\pi j(Ux + Vy)] \right\}$$

and

$$g_\sigma(x, y) \cdot \left\{ \sin 2\pi\varphi \cos [2\pi j(Ux + Vy)] + \cos 2\pi\varphi \sin [2\pi j(Ux + Vy)] \right\}.$$

The purpose of receptive fields having this shape remains unclear, as does whether or not a preponderance of them occur in quadrature pairs. Computational advantages of using filters with components expressible as linear combinations of symmetrical Gabor filters are also not apparent.

Texton Detection versus Texture Filtering: Models of preattentive texture perception often parallel experimental findings of neural function. One important class of models is exemplified by the work of Julesz and others [1]–[3], [58], [59] who contend that texture processing involves the computation of the densities of fundamental texture features labelled “textons.” The candidate (gray-level, monocular, static) textons identified thus far include elongated blobs (lines segments, ellipses, etc.) of various orientations, sizes, and aspect ratios, line terminations, and crossings of line segments. Texton detection is consistent with the feature detection model of early visual processing, presumably involving the responses of the bar- and edge-sensitive receptive fields. Segregation of textures then proceeds by comparison of relative texton densities [44], [45], [58], [59]. There is certainly considerable evidence that the visual system does regard such simple fea-

tures as important, and that visual discrimination of textures coincides with differences in texton density. However, the notion that textons are resolved first, and their densities afterwards is questionable: there is no apparent mechanism for computing measures of texton density within the confines of the feature detection model. Clearly, what is needed is a model allowing for the simultaneous local extraction of textons and global computation of their densities.

The view taken here is that the Gabor multiple channel model fills this need in optimal fashion. The spatial localization properties of the channel filters clearly satisfy the apparent requirement of sensitivity to local features, whereas narrow channel bandwidths ensure that global properties are simultaneously obtained: texton densities are embedded in the channel center frequencies and in the channel envelopes. Localized spatial phenomena such as textural phase discontinuities are easily computed since the spatial responses are localized. Despite the similarities of the multiple channel model to concepts of spectrum analysis, textural differences are still resolved on a localized spatial basis. Hence the fact that preattentively discriminable pseudo-random textures having identical energy spectra can be generated [60], [61] does not preclude the multiple channel model. The Gabor implementation effectively unifies the solution of the conflicting problems of determining local textural structures (features, texture boundaries) and identifying the spatial extents of textures contributing significant spectral information, e.g., the densities of oriented and/or elongated textons.

VI. CONCLUDING REMARKS

In this paper a computational approach for the analysis of visible surface texture has been demonstrated. Although the concepts put forth are based on various evidences for similar processing in biological vision systems, the approach has been presented with little reference to biological arguments. It was found that the approach is feasible for segmenting both artificial and natural textures in a predictable manner.

Rather than giving a lengthy summary of the results obtained in this paper, we instead choose to suggest other avenues of related research. The Gabor model of visual information processing may be applicable to many other computer vision tasks than the analysis of monochromatic, static, monocular irradiance patterns. The increasingly rapid development of models of the “front-end” processing of visual information in advanced biological systems has been accompanied by attempts to develop similar algorithms for accomplishing a large variety of computer vision tasks; the many computational approaches based on the ZC’s of the LoG edge detector serves as a good example. It is natural that the feasibility of the Gabor model be likewise tested by applying it to other low-level vision tasks under appropriate mathematical frameworks. For example, it may be interesting to use the segmentation boundaries computed from the channel envelopes as matching primitives in feature-based ste-

reo algorithms; the boundaries can be expected to have properties similar to those of LoG-based ZC's (continuity, thinness), but to be less densely distributed. The segmentation boundaries could then be matched to give crude 3-D shape, depth, and orientation information, or in conjunction with ZC's to produce a denser depth map. In a parallel implementation, the computational complexity of the ensuing correspondence problem could be reduced by only computing matches from features taken from corresponding channels, although some means for ensuring the consistency of the matches across channels would be needed.

There is recent evidence for Gabor-shaped receptive fields which are directionally sensitive to local image velocities [62], [63]; many of the "optimal" properties associated with the simultaneous localization of spatial patterns in space and frequency might be extended to include temporal localization. This approach has recently been considered by Heeger [64].

Although there is no evidence indicating that Gabor functions are involved in color processing or color segmentation, the fact that color is usually regarded as a region attribute suggests that region segmentation might be effected in a localized manner by applying a set of filters to the color components. Perhaps the most difficult problem, however, is not the development of new algorithms for accomplishing visual tasks using Gabor filters, but rather, the development of fast algorithms and architectures for filtering an image by a large (complete) set of Gabor filters. While the Gaussian pyramid [64] and the proposed Gabor pyramid [42] effectively reduce the computation at each scale, the number of filters required at each scale of processing is still formidable.

ACKNOWLEDGMENT

The authors gratefully acknowledge the comments of one of the reviewers who helped to improve the paper.

REFERENCES

- [1] B. Julesz and J. R. Bergen, "Textons, the fundamental elements in preattentive vision and perception of textures," *Bell Syst. Tech. J.*, vol. 62, pp. 1619-1645, 1983.
- [2] B. Julesz, "Visual pattern discrimination," *IRE Trans. Inform. Theory*, vol. 8, pp. 84-92, 1962.
- [3] —, "Experiments in the visual perception of texture," *Sci. Amer.*, vol. 232, pp. 34-43, 1975.
- [4] M. Clark and A. C. Bovik, "Texture discrimination using a model of visual cortex," in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, Atlanta, GA, 1986.
- [5] M. R. Turner, "Texture discrimination by Gabor functions," *Biol. Cybern.*, vol. 55, pp. 71-82, 1986.
- [6] M. Clark, A. C. Bovik, and W. S. Geisler, "Texture segmentation using a class of narrowband filters," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, Dallas, TX, 1987.
- [7] A. C. Bovik, M. Clark, and W. S. Geisler, "Computational texture analysis using localized spatial filtering," in *Proc. IEEE Comput. Soc. Workshop Comput. Vision*, Miami Beach, FL, Dec. 1987.
- [8] D. Gabor, "Theory of communication," *J. Inst. Elect. Eng.*, vol. 93, pp. 429-457, 1946.
- [9] J. G. Daugman, "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.
- [10] R. N. Bracewell, *The Fourier Transform and Its Applications*. New York: McGraw-Hill, 1978.
- [11] 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.
- [12] —, "Gabor's signal expansion and degrees of freedom of a signal," *Opt. Acta*, vol. 29, pp. 1223-1229, 1982.
- [13] S. Marcelja, "Mathematical description of the responses of simple cortical cells," *J. Opt. Soc. Amer.*, vol. 70, pp. 1297-1300, 1980.
- [14] W. K. Pratt, *Digital Image Processing*. New York: Wiley, 1978.
- [15] W. H. H. J. Lunscher and M. P. Beddoes, "Optimal edge detector design I: Parameter selection and noise effects," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-8, pp. 164-177, 1986.
- [16] D. Marr, *Vision*. San Francisco, CA: W. H. Freeman, 1982.
- [17] W. S. Geisler and D. B. Hamilton, "Sampling theory analysis of spatial vision," *J. Opt. Soc. Amer. A*, vol. 3, pp. 62-70, 1986.
- [18] P. Brodatz, *Textures: A Photographic Album for Artists and Designers*. New York: Dover, 1966.
- [19] H. C. Nothdurft, "Orientation sensitivity and texture segmentation," *Vision Res.*, vol. 25, pp. 551-560, 1985.
- [20] B. K. P. Horn, *Robot Vision*. Cambridge, MA: M.I.T. Press, 1986.
- [21] T. Poggio and V. Torre, "On edge detection," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-8, pp. 147-163, 1986.
- [22] D. Marr and E. Hildreth, "Theory of edge detection," *Proc. Roy. Soc. London B*, vol. 207, pp. 187-217, 1980.
- [23] V. Berzins, "Accuracy of Laplacian edge detectors," *Comput. Vision, Graphics, Image Processing*, vol. 27, pp. 195-210, 1984.
- [24] J. F. Canny, "Finding edges and lines in images," M.I.T. Artificial Intell. Lab., Cambridge, MA, AI Memo 720, 1983.
- [25] M. D. Levine, *Vision in Man and Machine*. New York: McGraw-Hill, 1985.
- [26] R. M. Haralick, K. Shanmugan, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-3, pp. 610-621, Nov. 1973.
- [27] R. M. Haralick, "Statistical and structural approaches to texture," *Proc. IEEE*, vol. 67, pp. 786-804, 1979.
- [28] S. W. Zucker and D. Terzopoulos, "Finding structure in co-occurrence matrices for texture analysis," *Comput. Vision, Graphics, Image Processing*, vol. 12, pp. 286-308, Mar. 1980.
- [29] C. H. Chen, "A study of texture classification using spectral features," in *Proc. Int. Conf. Pattern Recognition*, Munich, West Germany, Oct. 19-22, 1982.
- [30] H. Wechsler, "Texture analysis—A survey," *Signal Processing*, vol. 2, pp. 271-282, 1982.
- [31] L. Van Gool, P. Dewaele, and A. Oosterlinck, "Texture analysis anno 1983," *Comput. Vision, Graphics, Image Processing*, vol. 29, pp. 336-357, 1985.
- [32] R. Bajcsy, "Computer description of textured surfaces," in *Proc. Int. Joint Conf. Artif. Intell.*, Stanford, CA, Aug. 20-23, 1973.
- [33] R. Bajcsy and L. Lieberman, "Texture gradient as a depth cue," *Comput. Graphics Image Processing*, vol. 5, pp. 52-67, Mar. 1976.
- [34] O. D. Faugeras, "Texture analysis and classification using a human visual model," in *Proc. Int. Conf. Pattern Recogn.*, Kyoto, Japan, Nov. 7-10, 1978.
- [35] K. I. Laws, "Textured image segmentation," Ph.D. dissertation, Dep. Eng., Univ. Southern California, 1980.
- [36] M. Pietikainen, "On the use of hierarchically computed 'Mexican hat' features for texture discrimination," *Comput. Vision Lab., Univ. Maryland, College Park, Tech. Rep. TR-968*, Nov. 1980.
- [37] A. Ikonopoloulos and M. Unser, "A directional filtering approach to texture discrimination," in *Proc. 7th Int. Conf. Pattern Recogn.*, 1984.
- [38] F. D'Astous and M. E. Jernigan, "Texture discrimination based on detailed measures for the power spectrum," in *Proc. 7th Int. Conf. Pattern Recogn.*, 1984.
- [39] H. Knutsson and G. H. Granlund, "Texture analysis using two-dimensional quadrature filters," in *Proc. IEEE Workshop Comput. Architect. Pattern Anal. Image Database Management*, Oct. 12-14, 1983.
- [40] J. M. Coggins and A. K. Jain, "A spatial filtering approach to texture analysis," *Pattern Recogn. Lett.*, vol. 3, pp. 195-203, 1985.
- [41] R. Wilson and M. Spann, "Finite prolate spheroidal sequences and their applications II: Image feature description and segmentation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 10, pp. 193-203, Mar. 1988.
- [42] M. Porat and Y. Y. Zeevi, "The generalized Gabor scheme of image representation in biological and machine vision," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 10, pp. 452-468, July 1988.
- [43] 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.
- [44] T. C. Rearick, "A texture analysis algorithm inspired by a theory of preattentive vision," in *Proc. Int. Conf. Comput. Vision Patt. Recogn.*, San Francisco, CA, 1985.
- [45] H. Voorhees and T. Poggio, "Detecting textons and texture boundaries in natural images," in *Proc. First Int. Conf. Comput. Vision*, London, England, June 8-11, 1987.
- [46] D. H. Hubel and T. N. Wiesel, "Receptive fields of single neurons in the cat's striate cortex," *J. Physiol. London*, vol. 148, pp. 574-591, 1959.
- [47] D. H. Hubel and T. N. Wiesel, "Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex," *J. Physiol. London*, vol. 160, pp. 106-154, 1962.
- [48] F. W. Campbell and J. G. Robson, "Applications of Fourier analysis of the visibility of gratings," *J. Physiol. London*, vol. 197, pp. 551-556, 1968.
- [49] A. Pantle and R. Sekular, "Contrast response of human visual mechanisms sensitive to orientation and direction of motion," *Vision Res.*, vol. 9, pp. 397-406, 1970.
- [50] J. A. Movshon, I. D. Thompson, and D. J. Tolhurst, "Spatial summation in the receptive fields of simple cells in the cat's striate cortex," *J. Physiol. London*, vol. 283, pp. 53-77, 1978.
- [51] R. L. De Valois, D. G. Albrecht, and L. G. Thorell, "Cortical cells: Bar and edge detectors, or spatial frequency filters," in *Frontiers in Visual Science*, S. J. Cool and E. L. Smith, Eds. New York: Springer-Verlag, 1978.
- [52] J. G. Daugman, "Two-dimensional spectral analysis of cortical receptive field profiles," *Vision Res.*, vol. 20, pp. 847-856, 1980.
- [53] J. J. Kulikowski, S. Marcelja, and P. O. Bishop, "Theory of spatial position and spatial frequency relations in the receptive fields of simple cells in the visual cortex," *Biol. Cybern.*, vol. 43, pp. 187-198, 1982.
- [54] D. A. Pollen and S. F. Ronner, "Visual cortical neurons as localized spatial frequency filters," *IEEE Trans. Syst., Man, Cybern.*, vol. 13, pp. 907-916, 1983.
- [55] M. A. Webster and R. L. De Valois, "Relationship between spatial-frequency and orientation tuning of striate-cortex cells," *J. Opt. Soc. Amer.*, vol. 2, pp. 1124-1132, 1985.
- [56] D. G. Stork and H. R. Wilson, "Analysis of Gabor function descriptions of visual receptive fields," *Assoc. Res. Vision Ophthal. Ann. Spring Meeting*, Sarasota, FL, May 1988.
- [57] D. A. Pollen and S. F. Ronner, "Phase relationships between adjacent simple cells in the visual cortex," *Science*, vol. 212, pp. 1409-1411, 1981.
- [58] B. Julesz, "Textons, the elements of texture perception, and their interactions," *Nature*, vol. 290, pp. 91-97, 1981.
- [59] J. Beck, "Textural segmentation, second-order statistics, and textural elements," *Biol. Cybern.*, vol. 48, pp. 125-130, 1983.
- [60] B. Julesz and T. Caelli, "On the limits of Fourier decompositions in visual texture perception," *Perception*, vol. 8, pp. 69-73, 1979.
- [61] B. Julesz, "Spatial nonlinearities in the instantaneous perception of textures with identical power spectra," *Phil. Trans. Roy. Soc. London B*, vol. 290, pp. 83-94, 1980.
- [62] E. H. Adelson and J. R. Bergen, "Spatiotemporal energy models for the perception of motion," *J. Opt. Soc. Amer. A*, vol. 2, pp. 284-299, 1985.
- [63] A. B. Watson and A. J. Ahumada, Jr., "Model of human visual-motion sensing," *J. Opt. Soc. Amer. A*, vol. 2, pp. 322-341, 1985.
- [64] D. J. Heeger, "Optical flow from spatiotemporal filters," in *Proc. First Int. Conf. Comput. Vision*, 1987, pp. 181-190.



Alan Conrad Bovik (S'80-M'84) was born in Kirkwood, MO, on June 25, 1958. He received the B.S. degree in computer engineering in 1980, and the M.S. and Ph.D. degrees in electrical and computer engineering in 1982 and 1984, respectively, all from the University of Illinois, Urbana-Champaign.

He is currently Associate Professor in the Department of Electrical and Computer Engineering, the Department of Computer Sciences, and the Biomedical Engineering Program at the University of Texas at Austin, where he is the Director of the Laboratory for Vision Systems. His current research interests include computer vision, nonlinear statistical methods in digital signal and image processing, biomedical image processing, and computational aspects of biological visual perception.

Dr. Bovik is the Stark Centennial Endowed Fellow in Engineering at the University of Texas and a Registered Professional Engineer in the State of Texas. He is an Honorable Mention winner of the international Pattern Recognition Society Award, and is Associate Editor of the international journal *Pattern Recognition*. He is Local Arrangements Chairman for the IEEE Computer Society Workshop on the Interpretation of 3-D Scenes in Austin, TX, November 1989, a member of the Program Committee of the Tenth International Conference on Pattern Recognition in Atlantic City, NJ, June 1990, Program Co-Chairman of the SPIE Conference on Image Processing, SPIE/SPSE Symposium on Electronic Imaging, Santa Clara, CA, February 1990, and Conference Chairman of the SPIE Conference on Medical Image Processing, Santa Clara, CA, February 1990. He has recently been named Associate Editor of the IEEE TRANSACTIONS ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING.



Marianna Clark (S'84-M'86) was born in Kansas City, MO, on March 9, 1953. She received the B.S. and M.S. degrees in electrical engineering from Texas Tech University, Lubbock, in 1976 and 1978, respectively.

She was a Member of the Technical Staff at Sandia National Laboratory in Albuquerque, NM, from 1978 to 1983. She is currently employed at Lockheed Missiles and Space Corporation, Austin Division, while pursuing the Ph.D. degree at the University of Texas at Austin.

Ms. Clark is a member of Sigma Xi.

Wilson S. Geisler received the A.B. degree in psychology from Stanford University, Stanford, CA, in 1971, and the Ph.D. degree in experimental and mathematical psychology from Indiana University, Bloomington, in 1975.

He joined the Department of Psychology at the University of Texas in 1975 where he is currently a Full Professor. Dr. Geisler's current research interests include the psychophysics, physiology, and quantitative modeling of spatial vision, visual adaptation, and color vision. He also has an active interest in texture, stereo, and motion perception, and in image processing and the theory of ideal observers.

Dr. Geisler is currently a member of the Visual Sciences B Study Section of the National Eye Institute of NIH.