Abstract—Many texture-segmentation schemes use an elaborate bank of filters to decompose a textured image into a joint space/spatial-frequency representation. Although these schemes show promise, and although some analytical work has been done, the relationship between texture differences and the filter configurations required to distinguish them remain largely unknown. This paper examines the issue of designing individual filters. Using a 2-D texture model, we show analytically that applying a properly configured bandpass filter to a textured image produces distinct output discontinuities at texture boundaries; the analysis is based on Gabor elementary functions, but it is the bandpass nature of the filter that is essential. Depending on the type of texture difference, these discontinuities form one of four characteristic signatures: a step, ridge, valley, or a step change in average local output variation. Accompanying experimental evidence indicates that these signatures are useful for segmenting an image. The analysis indicates those texture characteristics that are responsible for each signature type. Detailed criteria are provided for designing filters that can produce quality output signatures. We also illustrate occasions when asymmetric filters are beneficial, an issue not previously addressed.

Index Terms—Texture segmentation, texture discrimination, computer vision, Gabor functions, image segmentation

I. INTRODUCTION

Texture segmentation continues to be a challenging problem in computer vision. Examples of previously proposed approaches for segmenting textured images include those based on local geometric primitives [1]-[4], local statistical features [5]-[8], random field models [5], [9]-[11], and fractals [12], [13]. While these approaches can be applied successfully to many texture segmentation problems, any given approach is limited in the variety of textures that it can segment.

The human visual system, on the other hand, can preattentively segment textures robustly. This realization has motivated extensive studies and has led to a promising theory of human texture perception. This theory, supported by both psychophysical and neurophysiological data, holds that the human visual system is performing some form of local spatial-frequency analysis on the retinal image and that this analysis is done by a bank of tuned bandpass filters [14]-[20]. The concept of local spatial frequency, or local frequency, had been put forth in the context of communication systems many years earlier by Gabor [21]. Classically, images are viewed as either a collection of pixels (spatial domain) or the sum of sinusoids of infinite extent (spatial-frequency domain). Gabor, however, observed that the spatial representation and the spatial-frequency representation are just opposite extremes of a continuum of possible joint space/spatial-frequency representations. In a joint space/spatial-frequency representation for images, frequency is viewed as a local phenomenon (i.e., as a local frequency) that can vary with position throughout the image. Using this paradigm within the framework of human vision, perceptually significant texture differences presumably correspond to differences in local spatial-frequency content. Texture segmentation thus involves decomposing a retinal image into a joint space/spatial-frequency representation (by using a bank of bandpass filters) and then using this information to locate regions of similar local spatial-frequency content.

This paradigm has spurred researchers to devise a number of new texture-segmentation schemes for computer vision [22]-[31]. Two major issues arise, though, in constructing a successful scheme: the design of individual filters and the configuration of the filter bank. Regarding issue 1), several classes of functions have been proposed for the filters: Gabor elementary functions [22]-[26], [28], [29], [31], [32], a difference of offset Gaussians [27], [33], and Gaussian derivatives (Hermite polynomials) [33]. Regarding issue 2), Malik and Perona [27] took great pains to attempt to mimic the human visual system and have perhaps provided the most detailed justification for a particular filter-bank structure. Others have also used a complete bank of filters for texture segmentation [22], [24], [25], [26].

Although the filter-bank paradigm has shown much potential, and although some analytical work has been done to demonstrate the efficacy of certain types of filters [22], [23], [34], the relationships between texture differences and the filter configurations required to discriminate them remain largely unknown. We believe that an adequate understanding of how to design an individual filter is essential for understanding how to build a suitable filter bank. This paper addresses the issue of filter design.

Using a general 2-D texture model similar to one proposed by Clark and Bovik [23], we show analytically that applying a properly tuned bandpass filter to a textured image produces...
distinct output discontinuities at texture boundaries. Depending on the type of texture difference and the way the filter is tuned, the filter output can exhibit one of four characteristic discontinuities, or signatures, at the texture boundary: a step, valley, or ridge, or a step change in average local output variation. The analysis indicates those texture characteristics leading to the various signature types. It also provides parameter-selection guidelines for designing effective filters. Experimental results show that the signatures are useful for segmenting textured images. They also corroborate the quantitative analytical results. We further demonstrate instances where spatially asymmetric filters are beneficial, an issue not previously addressed.

Our analysis assumes that a Gabor elementary function (GEF) is used in a filter. This assumption is discussed and justified in Section II. The analysis, performed in Section IV and based on a texture model defined in Section III, reveals that it is the bandpass characteristic of a filter function that is essential in producing the various types of output signatures. Thus, other filter functions could conceivably be used, but our filter-design criteria, discussed in Section V, address only GEF's. The design criteria derived here are based on structural attributes of the textures of interest. Since such information is not readily available for arbitrary natural textures, we describe a technique for designing filters for such situations in a companion paper [35]. Other sections give experimental results (Section VI) and concluding remarks (Section VII).}

\[ g(x, y) = \frac{1}{2\pi \sigma_x \sigma_y} \exp \left\{ -\frac{1}{2} \left( \frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right) \right\}, \]

and \((\sigma_x, \sigma_y)\) characterize the spatial extent and bandwidth of \(h\). The aspect ratio of \(g(x, y)\) is given by \(\lambda \triangleq \sigma_y / \sigma_x\) and gives a measure of the filter’s asymmetry. The Fourier transform of \(h\) is the following equation:

\[ H(u, v) = \exp \left\{ -2\pi^2 \left( \sigma_x (u - U)^2 + \sigma_y (v - V)^2 \right) \right\}, \]

where \((u - U)' , (v - V)'\) = [(u - U) cos \(\theta\) + (v - V) sin \(\theta\), -(u - U) sin \(\theta\) + (v - V) cos \(\theta\)] are shifted and rotated frequency coordinates. \(H(u, v)\) is a Gaussian that is shifted \((U, V)\) frequency units along the frequency axes \((u, v)\) and rotated by an angle \(\theta\) relative to the positive \(u\)-axis. Thus, \(H\) acts as a bandpass filter with center frequency \((U, V)\) [relative to \((u, v)\)] and a bandwidth controlled by \(\sigma_x\) and \(\sigma_y\). Note that when the aspect ratio \(\lambda\) of \(g(x, y)\) differs from unity, the Gaussian is asymmetric with an orientation \(\theta\) that generally differs from the orientation \(\phi\) of the complex sinusoid.

As our analysis shows, it is the bandpass nature of the GEF that is most essential for effectively analyzing a textured image. Hence, since the aforementioned possibilities for filter functions—the difference of offset Gaussians [27], [33] and Gaussian derivatives [33]—also share this property, the choice of the GEF is not restrictive. Within the context of modeling human texture perception, Malik and Perona mentioned that the exact choice of a filter function was unimportant, and they chose various variants on the difference of offset Gaussians for computational simplicity and physiological plausibility [27]. Also, Bovik et al. have discussed the efficacy of bandpass filters for texture segmentation [22], [34].

We now discuss the magnitude operation used in the Gabor filter (1). Julesz has shown that purely linear mechanisms are inadequate to explain how humans perceive texture [40]. This point was further asserted by Malik and Perona [27]. Therefore, to simulate human texture perception, some form of nonlinearity is desirable. The magnitude operator introduces the desirable nonlinearity into the filter. The convolution of an image with a GEF results in a complex-valued subimage. Bovik et al. have shown that the amplitude envelope of this subimage can be recovered by computing its magnitude, and that the resulting amplitude envelope is useful for texture segmentation [22]. Also, note that the magnitude operation has previously been suggested extensively [22]-[26], [31].

Note that the magnitude operation is not without flaw. Aside from being implausible neurophysiologically, Malik and Perona have shown that computing the magnitude makes it impossible to discriminate certain texture pairs [27]. Appendix I analytically verifies this assertion, but then shows that if mimicking human perception is not essential, then a wide range of textures can be segmented without using a nonlinearity. In
spite of shortcomings, the magnitude computation provides a convenient analysis tool and serves as a benchmark for comparing alternatives.

III. TEXTURE MODEL

Although researchers have not agreed on a precise definition for texture, several descriptions have been proposed [5], [41]-[43]. Many textures can be described as a collection of similar, but not necessarily identical, primitive objects arranged in some repeating pattern. Based on this notion, we model texture as a collection of simple objects called texels. Groups of similar texels form regions of homogeneous texture. A textured image consists of two or more regions where texture differences between regions are induced by varying the type and/or organization of the texels. Using Rao’s terminology, this approach can represent a variety of textures, the texels within a region may vary randomly in orientation, and the position and shape of the texels may be perturbed (e.g., Fig. 8(a)). For convenience, we divide textured images into two levels of complexity: uniform and nonuniform. For uniform textures, all texels within a region are identical in shape and orientation and are spaced uniformly (e.g., Fig. 3(a)). For nonuniform textures, deferring to a real-world image setting, Clark and Bovik employed a similar model, but their analysis leads to somewhat more tractable results and also more easily leads to an understanding of specific filter-output behavior. Now,

\[ F[i(x,y)] = I(u,v) = \hat{I}_1(u,v) + \hat{I}_2(u,v) \]  

where (evaluating (7) and (11))

\[ \hat{I}_1(u,v) = \frac{2\pi}{\Delta x \Delta y} \sum_{k,l} S_{k,l} T_1 \left( \frac{2\pi k}{\Delta x}, \frac{2\pi l}{\Delta y} \right) \]  

\[ \hat{I}_2(u,v) = \frac{2\pi}{\Delta x \Delta y} \sum_{k,l} S_{k,l} e^{-j\pi (u-2\pi k/\Delta x)v} \Delta x \Delta y \]  

For uniform textured region \( i_2 \) of support \( s \times s \) and centered at \( (r,0) \) is given by the following:

\[ i_2(x,y) = \hat{I}_1(u,v) + \hat{I}_2(u,v) \]  

Define a real deterministic function that has a Fourier transform \( T_1(u,v) \), make analysis tractable. For example, well-defined sinc functions, such as (8), occur frequently during the subsequent analysis. Also, a spatially limited \( i \) conforms to a real-world image setting, Clark and Bovik employed a similar model, but their analysis led to somewhat more tractable results and also more easily leads to an understanding of specific filter-output behavior. Now,

\[ F[i(x,y)] = I(u,v) = \hat{I}_1(u,v) + \hat{I}_2(u,v) \]  

where (evaluating (7) and (11))

\[ \hat{I}_1(u,v) = \frac{2\pi}{\Delta x \Delta y} \sum_{k,l} S_{k,l} T_1 \left( \frac{2\pi k}{\Delta x}, \frac{2\pi l}{\Delta y} \right) \]  

\[ \hat{I}_2(u,v) = \frac{2\pi}{\Delta x \Delta y} \sum_{k,l} S_{k,l} e^{-j\pi (u-2\pi k/\Delta x)v} \Delta x \Delta y \]  

\[ S_{k,l} = S(u-2\pi k/\Delta x, v-2\pi l/\Delta y) \]  

and \( S(u,v) \) is given by (8). Alternatively, the following can be used:

\[ I(u,v) = \frac{2\pi}{\Delta x \Delta y} \sum_{k,l} S_{k,l} \left( T_1 \left( \frac{2\pi k}{\Delta x}, \frac{2\pi l}{\Delta y} \right) + T_2 \left( \frac{2\pi k}{\Delta x}, \frac{2\pi l}{\Delta y} \right) e^{-j\pi (u-2\pi k/\Delta x)v} \Delta x \Delta y \right) \]  

\[ + T_3 \left( \frac{2\pi k}{\Delta x}, \frac{2\pi l}{\Delta y} \right) e^{-j\pi (u-2\pi k/\Delta x)v} \Delta x \Delta y \]  

where

\[ \Pi_i(x,y) = \begin{cases} 1, & |x| < \frac{s}{2} \text{ and } |y| < \frac{s}{2}, \\ 0, & \text{otherwise} \end{cases} \]
Observe that \( \hat{I}_1 \) consists of a collection of scaled 2-D sinc functions centered at the harmonics \((2\pi k / \Delta x, 2\pi l / \Delta y)\). The amplitude of the sinc \( S_{k,l} \) at harmonic \((2\pi k / \Delta x, 2\pi l / \Delta y)\) is proportional to the value of the Fourier transform of the texel \( T_1 \) evaluated at that harmonic. \( \hat{I}_2 \) also consists of a collection of scaled 2-D sinc functions centered at the harmonics. The amplitudes of the sincs for \( I_2 \), however, are proportional to \( T_2 \) rather than \( T_1 \), and their phase components are influenced by a complex phase factor. Thus, by (17), \( I \) is a sum of scaled sincs \( S_{k,l} \). Each sinc consists of a component from each texture region, or, more colloquially, each \((k, l)\) component of \( I \) consists of \( \text{a pair of sincs} \), one for each texture.

Thus, the texture segmentation problem is to find the boundary separating regions \( t_1 \) and \( t_2 \) in image \( i \). Pursuant to the model’s construction, the boundary separating these two textures is the line segment given by \( x = r/2 \) and \( |y| < s/2 \).

We wish to understand how the Gabor filter (1) will help in locating this boundary.

IV. CHARACTERIZING GABOR FILTER OUTPUTS

We show analytically that the application of Gabor filters to textured images produces outputs that exhibit discontinuities in the neighborhood of texture boundaries. This is shown within the context of the texture model defined in the previous section. We begin by analyzing those texture configurations that produce a step signature. This is followed by an analysis of texture types that produce a valley or ridge signature. The section concludes with a qualitative discussion of nonuniform textures, which leads to the fourth signature type, a step change in average local output variation. The analysis in this section ultimately leads to the filter-design criteria presented in Section V.

A. Textures Made Up of Different Texels: Step Signature

This section derives conditions when the application of the Gabor filter (1) to a uniformly textured image produces a step signature. The step signature is characterized by a step change in the Gabor-filter output \( m \) at the boundary between two textured regions. This signature type occurs when a properly tuned Gabor filter is applied to a uniformly textured image that contains two textures whose constituent texels \( t_1 \) and \( t_2 \) differ.

To derive this result, consider the outcome of applying a Gabor filter (1) to the textured-image model \( I \) in (13) (or, equivalently, \( i \) in (12)). The goal is to design a filter that enables “easy” localization of the texture boundary. Analytically, the approach is to design a Gabor filter that passes the image energy centered about one harmonic \((k, l)\). This is equivalent to passing one and only one scaled sinc \( S_{k,l} \) in (17), where the sinc draws contributions from each texture, i.e., to design a filter that passes one sinc pair occurring at some harmonic \((k, l)\). Each sinc in the pair represents a gate function in the spatial domain. Each gate coincides with one of the two region boundaries, and the difference in gate amplitude is proportional to the amplitude difference between the two sints (i.e., \([T_1 - T_2]\)). By filtering out a sinc pair whose sints differ significantly in amplitude, a filter output is produced that is approximately constant within a region, but differs between regions, thus forming a step signature.

Designing a Gabor filter involves specifying the five parameters \((U, V, \sigma_x, \sigma_y, \theta)\) of the GEF \( H \) in (4). To pass the single sinc-pair at harmonic indices \((k, l)\), the center frequency \((U, V)\) of \( H \) is specified as \( U = 2\pi k / \Delta x, V = 2\pi l / \Delta y \). The bandwidth of \( H \), determined by \((\sigma_x, \sigma_y)\), is then selected so that \( H \) passes most of the image energy centered about harmonic \((k, l)\) while also largely rejecting the image energy at adjacent harmonics. Since harmonic spacing is proportional to texel spacing \((\Delta x, \Delta y)\), the ratios \((\sigma_x / \Delta x, \sigma_y / \Delta y)\) determine this filter characteristic. Clearly, the choice of \((\sigma_x / \Delta x, \sigma_y / \Delta y)\) is a trade-off between attenuation of the desired harmonic and a rejection of adjacent harmonics. The consequences of this trade-off are discussed in Section V.

Applying \( H \) to \( I \) gives the following:

\[ I_f(u, v) = H(u, v)I(u, v) \]

Since \( H \) has been designed to pass only those frequency components in the neighborhood of \((U, V)\), we can write the following equation:

\[
I_f(u, v) \approx \frac{2\pi}{\Delta x \Delta y} H(u, v)S(u - U, v - V) \\
\times \left\{ T_1 + T_2 e^{-j\theta(u-U)} \right\} 
\]

(18)

where \( T_1 \) and \( T_2 \) are abbreviations for \( T_1(U, V) \) and \( T_2(U, V) \). Observing that \( H \) in (4) is a function of \( u-U \) and \( v-V \), we define the function \( S_f \) as follows:

\[
S_f(u-U, v-V) = H(u, v)S(u-U, v-V) 
\]

(19)

where

\[
F^{-1}[S_f(u-U, v-V)] = s_f(x, y)e^{j(Ux+Vy)} 
\]

and

\[
s_f(x, y) = F^{-1}[S_f(u, v)]. 
\]

(20)
By substituting $S_f$ in (18), the inverse Fourier transform of $I_f$ can be expressed as follows:

$$i_f(x, y) = \frac{2\pi}{\Delta x \Delta y} e^{i(Ux + Vy)} [T_1 s_f(x, y) + T_2 s_f(x - r, y)].$$

(21)

Computing the magnitude of $i_f$ completes the application of the Gabor filter and gives the following equation:

$$m(x, y) = |i_f(x, y)| = \frac{2\pi}{\Delta x \Delta y} \sqrt{A + B + C}$$

(22)

where

$$A = |T_1|^2 s_f^2(x, y)$$
$$B = |T_2|^2 s_f^2(x - r, y)$$
$$C = (T_1^* T_2 + T_1 T_2^*) s_f(x, y) s_f(x - r, y).$$

(It can be seen in (4), (8), (19), and (20) that $s_f$ is real.)

To understand the behavior of $m$, we first need to determine $s_f$. $s_f$ equals a sinc multiplied by a Gaussian. Thus, in the spatial domain, $s_f$ can be expressed as the convolution of a Gaussian with a gate function:

$$s_f(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x - \alpha, y - \beta) d\alpha d\beta$$

(23)

$$s_f(x - r, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x - \alpha, y - \beta) d\alpha d\beta.$$

(24)

where $g$ is the Gaussian (3). The quantity $m$ can now be evaluated by examining its behavior at the texture boundary and at points far removed from the boundary (or, equivalently, at points within the interiors of each texture). Assume that the region width $r$ in the $x$ direction is large relative to $\sigma_x$, and that the region height $s$ in the $y$ direction is large relative to $\sigma_y$. Then, for points away from the textured image’s outer boundary and left of the texture boundary (i.e., $|y| < s/2$ and $|x| < r/2$ for points in region 1), $s_f(x, y) \approx 1$ and $s_f(x - r, y) \approx 0$, pursuant to (23) and (24). Then, $m \approx \frac{2\pi}{\Delta x \Delta y} |T_1|$. Similarly, for points to the right of the texture boundary in region 2, $m \approx \frac{2\pi}{\Delta x \Delta y} |T_2|$. Now, at the texture boundary ($x = r/2$), the filter output $m$ is as follows:

$$m(r/2, y) \approx \frac{2\pi}{\Delta x \Delta y} \sqrt{|T_1|^2/4 + |T_2|^2/4 + (T_1 T_2 + T_1 T_2^*)/4}$$

$$\approx \frac{2\pi}{\Delta x \Delta y} \sqrt{(|T_1| + |T_2|)^2}$$

$$= \frac{2\pi}{\Delta x \Delta y} |T_1 + T_2|,$$

(25)

because $s_f(x, y)$ and $s_f(x - r, y)$ both $\approx 1/2$ at $x = r/2$. Now suppose that $T_1$ and $T_2$ are both real and positive. Then, $m(r/2, y)$ becomes the average of values far to the right and left of the texture boundary, and (22) can be rewritten as follows:

$$m(x, y) = \frac{2\pi}{\Delta x \Delta y} \left( |T_1|^2 s_f^2(x, y) + |T_2|^2 s_f^2(x - r, y) + 2 T_1 T_2 s_f(x, y) s_f(x - r, y) \right)^{1/2}$$

$$= \frac{2\pi}{\Delta x \Delta y} \left( T_1 s_f(x, y) + T_2 s_f(x - r, y) \right).$$

Observing that $s_f(x - r, y) \approx 1 - s_f(x, y)$, we see that $m$ is a linear function of $s_f$. Since $s_f$ is the integral of a Gaussian, its shape is similar to a sigmoid function. Thus, $m$ is also shaped like a sigmoid in the neighborhood of the texture boundary. Assuming that $|T_1| \neq |T_2|$, $m$ is given by the following constant value:

$$A_1 = \frac{2\pi}{\Delta x \Delta y} |T_1|$$

(26)

over region 1 and by the following constant value:

$$A_2 = \frac{2\pi}{\Delta x \Delta y} |T_2|$$

(27)

over region 2 with a sigmoid transition between regions; i.e., the following:

$$m(x, y) \approx \begin{cases} A_1, & x < r/2, \\ \text{sigmoidal transition from } A_1 \text{ to } A_2, & x \text{ near } r/2 \\ A_2, & x > r/2. \\ \end{cases}$$

Thus, $m$ resembles a step function with the transition occurring near the texture boundary.

Suppose now that $T_1$ and $T_2$ are negative or complex. Then, $m$ can take on values $< \min(A_1, A_2)$ or $> \max(A_1, A_2)$ near the texture boundary. We refer to these possibilities as undershoot and overshoot. To see how undershoot can occur, (25) shows that near the texture boundary, $m$ is proportional to $(T_1 + T_2)$. Thus, if $T_1$ and $T_2$ are negative or complex, the magnitude of their sum can be less than the magnitude of either component. Overshoot can occur if the Gabor-filter center frequency $(U, V)$ is not equal to one of the harmonics of $I$. The phenomena of undershoot and overshoot need not overly complicate the detection of the texture boundary. They are illustrated in the results section of this paper and are discussed analytically in [44].

B. Textures Using Identical Texels, but Exhibiting a Texture Phase Difference: Valley and Ridge Signatures

This section shows that certain texture-phase differences can be detected without explicitly computing phase differences (cf. [22], [23], [34]). The approach is to design a suitable Gabor filter that detects discontinuities in the filter output $m$ caused by abrupt changes in the texture phase. Since the magnitude operation in the Gabor filter (1) discards the phase of the GEF-filtered image, information is lost. Appendix I discusses the issue of phase and alternatives to magnitude computation.

An example of a texture phase change is illustrated in Fig. 6(a). The two uniform regions are identical, but offset both horizontally and vertically. Thus, the Fourier transform magnitudes of the two regions are identical, but their respective phase characteristics differ. We refer to this type of texture-phase difference as a texture-phase difference. (This phenomena could equivalently be viewed as a collection of different texels near the texture boundary, but analysis suggests that a difference-in-phase interpretation is more appropriate.) The derivation to follow shows that a texture-phase difference produces a valley in the Gabor-filter output $m$ when the GEF
is properly tuned; if an improperly tuned GEF is used, a ridge occurs in m at the texture boundary.

Valley Signature: Again, the goal is to design a filter that enables easy localization of the texture boundary. Analytically, the procedure is to design a Gabor filter that passes the image energy centered about one harmonic \( (2\pi k/Ax, 2\pi l/\Delta y) \). This is equivalent to passing one and only one sinc pair centered about some harmonic \( (2\pi k/\Delta x, 2\pi l/\Delta y) \). In this case, the amplitudes of the sincs are identical. The offset regions, however, produce a phase shift \( \psi \) (given in (30)) between the sincs, resulting in a drop in filter output, given by (33), near the texture boundary.

We first modify the texture model of Section III to fit the texture-phase-difference scenario. Define a texel \( t_1 \) as before, and construct a uniform textured region \( i_1 \) as in (6). Define a second texel \( i_2 \) equal to \( i_1 \), but shifted \( \delta x \) in the \( x \) direction and \( \delta y \) in the \( y \) direction, where \( 0 < \delta x < \Delta x \) and \( 0 < \delta y < \Delta y \). Then,

\[
I_2(x, y) = I_1(x - \delta x, y - \delta y).
\]

A uniform texture \( i_2 \) whose texels are periodic in \( x \) and \( y \) can be constructed from this texel as shown in (9), and a uniform textured region \( i_2 \) of support \( \tau \times \tau \) and centered at \( (r, 0) \) can be formed from \( i_2 \) as shown in (10). Thus, a uniform textured image \( i \) that exhibits a texture-phase difference at \( x = r/2 \) can be formed similarly to (12):

\[
i(x, y) = i_1(x, y) + i_2(x, y)
\]

\( \mathcal{F}[i(x, y)] \) is then similar to (13):

\[
I(u, v) = I_1(u, v) + I_2(u, v)
\]

\( I_1(u, v) \) is given by (14), but \( I_2(u, v) \) differs from (15), because of the following condition:

\[
T_2(u, v) = T_1(u, v)e^{-j(\omega x + \omega y)}
\]

Thus,

\[
I(u, v) = \frac{2\pi}{\Delta x \Delta y} \sum_{k, l} S_{kl} T_1 \left( \frac{2\pi k}{\Delta x}, \frac{2\pi l}{\Delta y} \right) \left\{ 1 + e^{-j\omega(u-2\pi k/\Delta x)}e^{-j\omega(\delta y/\Delta y)} \right\}
\]

(28)

Let the GEF \( H \) have center frequency \( (U, V) \), where \( U = 2\pi k/\Delta x \) and \( V = 2\pi l/\Delta y \) for some \( (k, l) \), and select \( (\sigma_x/\Delta x, \sigma_y/\Delta y) \) as in Section IV.A. Applying \( H \) to \( I \) approximately passes only the sinc-pair centered at \( (U, V) \):

\[
I_f(u, v) = H(u, v)I(u, v)
\]

\[
\approx \frac{2\pi}{\Delta x \Delta y} T_1(U, V)H(u, v)S(u - U, v - V)
\]

\[
\left\{ 1 + e^{-j\omega(u-U)}e^{-j\omega(\delta y/\Delta y)} \right\}.
\]

Defining \( S_f \) as in (19), the inverse Fourier transform of \( I_f \) is as follows:

\[
s_f(x, y) = \frac{2\pi}{\Delta x \Delta y} T_1(U, V)e^{-j(\omega x + \omega y)}
\]

\[
\left[ s_f(x, y) + s_f(x - r, y)e^{-j2\pi(\delta x/\Delta x + \delta y/\Delta y)} \right].
\]

Let

\[
\psi = 2\pi(\delta x/\Delta x + \delta y/\Delta y).
\]

\( \psi \) represents the total relative phase shift between regions 1 and 2. Computing the magnitude of \( \psi \) completes the application of the Gabor filter and gives the following equation:

\[
m(x, y) = C|s_f(x, y) + s_f(x - r, y)e^{-j\psi}|
\]

\[
= C \sqrt{s_f^2(x, y) + s_f^2(x - r, y) + 2s_f(x, y)s_f(x - r, y) \cos \psi}.
\]

(31)

where

\[
C = \frac{2\pi}{\Delta x \Delta y}|T_1(U, V)|.
\]

(32)

Consider the behavior of \( m \). Assume that a phase shift occurs, i.e., \( \forall (k, l), \psi \neq \) a multiple of \( 2\pi \); or, equivalently, choose some \( (k, l) \) such that \( \cos \psi \neq 1 \). This holds because of the restrictions placed earlier on \( \delta x \) and \( \delta y \). The image does not exhibit a phase discontinuity in the \( y \) direction. So, in subsequent analyses, it is assumed that \( y \) is far removed from the image’s outer boundaries (i.e., \( |y| \ll s/2 \)). Consider \( m \) over three regions. The first region consists of those values of \( x \) such that \( |x| < r/2 \) (i.e., points in region 1 far from both the texture boundary and the image’s outer boundary; see Fig. 1). In this case, (23) and (24) indicate that for \( r \) large relative to \( \sigma_x, s_f(x, y) \approx 1 \), and \( s_f(x - r, y) \approx 0 \). Thus, from (31), \( m \approx C \). The second region consists of those values of \( x \) for \( r/2 < x < 3r/2 \) (points in region 2). In this case, \( s_f(x - r, y) \approx 0, s_f(x, y) \approx 0 \), and \( m \approx C \). The last region is in the neighborhood of the texture boundary, \( x = r/2 \). At \( x = r/2 \), for large \( r, s_f(x, y) \approx s_f(x - r, y) \approx 1/2 \), and from (31), the following is true:

\[
m(r/2, y) \approx C \sqrt{0.5(1 + \cos \psi)}
\]

(33)

Summarizing,

\[
m(x, \cdot) \approx \begin{cases} 
C, & x < r/2, \\
C \sqrt{0.5(1 + \cos \psi)}, & x \text{ near } r/2 \\
C, & x > r/2.
\end{cases}
\]

Note that near the texture boundary \( (x \approx r/2) \), \( m(x, \cdot) \leq C \) is required, because of the weighting factor \( \sqrt{0.5(1 + \cos \psi)} \). Thus, for this situation, a valley signature occurs near the texture boundary.

When no phase shift exists between the two regions, the transition takes on its maximum value \( C \), which is the same value as for points far removed from the transition. This is expected because without a texture-phase shift, the two regions are indistinguishable. If \( \psi = \pi \) (maximum texture-phase difference between two regions), the value of \( m \) at the texture boundary is 0, and a minimum valley results. Note from (30) that the depth of the valley depends on the ratios \( \delta x/\Delta x \) and \( \delta y/\Delta y \). These ratios represent the amount of texture-phase shift in \( x \) and \( y \) relative to the texel periods. Thus, the greater the texture-phase shift, the deeper the valley. Reference [44] develops a method for estimating valley depth.
Ridge Signature: With the Gabor filter designed as indicated above, a ridge signature cannot occur near the texture boundary. It can be shown, however, that if the Gabor filter is tuned to a frequency other than a harmonic, a ridge is produced.

To see this, consider the frequency domain representation of the following GEF-filtered image:

\[ I_f(u,v) = H(u,v)J(u,v). \]

In this case, let \( U = 2\pi k / \Delta x + \delta U, \ V = 2\pi l / \Delta y + \delta V, \) where \( \delta U \) and \( \delta V \) are chosen so that \( 2\pi k / \Delta x < U < 3\pi(k+1) / \Delta x \) and \( 2\pi l / \Delta y < V < 3\pi(l+1) / \Delta y. \) Thus, the inverse Fourier transform \( i_f \) has the same form as in (29), but now \( s_f \) represents a gate function convolved with a GEF having center frequency \( (\delta U, \delta V) \) rather than a gate function convolved with a simple Gaussian. Convolving a gate function with a GEF produces a complex quantity. So, \( s_f \) becomes complex. Thus, computing the magnitude of \( i_f \) as in (31) produces the following equation:

\[
m(x,y) = C \sqrt{P_x P_y + P_0 P_0 e^{-j\psi} + P_0 P_0 e^{j\psi}}
\]

(34)

where \( P_x = s_f(x,t,y) \) and \( \psi \) again is the total relative phase shift between regions 1 and 2, per (30). It can be shown that if \( \delta U \neq 0 \) or \( \delta V \neq 0, \) then the complex terms \( P_0 \) and \( P_x \) constructively interfere with each other to produce an increase in filter output near the texture boundary. This results in a ridge signature of the following form:

\[
m(x,\cdot) \approx \begin{cases} 
D_1, & x < r/2 \\
D_2(\psi), & x \near r/2, \quad \text{(texture boundary)} \\
D_1, & x > r/2 
\end{cases}
\]

where \( D \) is a constant and \( 2\psi > 1 \) for \( x \near r/2 \) [44].

Although generating a ridge is not the ideal result, it can still be useful for texture segmentation. To be able to perform texture segmentation, the ridge must be reasonably strong. The ridge’s strength depends upon its height, which depends on the texture phase shift \( \psi. \) Reference [44] gives a detailed development of the ridge signature and a method for estimating ridge height.

C. Nonuniform Textures

Unlike uniform textures, nonuniform textures consist of texels that can undergo perturbations in position, orientation, and shape. These perturbations make analysis more difficult. Although Clark and Bovik investigated the effect of texel orientation and shape perturbations is much more difficult, especially in the general case, because the results depend strongly on individual texel characteristics. Thus, we do not provide a quantitative analysis for nonuniform textures, but instead present qualitative arguments and experimental results (Section VI). Reference [35] gives more discussion on nonuniform textured images.

Our experimental results indicate that the signatures obtained for uniformly textured images also can occur for nonuniformly textured images. In contrast to uniform textures, though, a nonuniformly textured image cannot be represented in the frequency domain simply as a 2-D impulse train. The perturbations possible in the texels introduce more frequency components beyond simple harmonics. The net result is that the Gabor filter output signatures obtained for nonuniformly textured images typically exhibit many local output variations. (See, e.g., Fig. 8(b).) In spite of these output variations, the results of Section VI demonstrate that Gabor filter outputs derived from nonuniformly textured images can be useful for texture segmentation.

In addition to the three signature types derived for uniform textures, nonuniformly textured images can exhibit a fourth signature type. Consider, for example, two regions that have the same average spatial-frequency content, but with different spatial-frequency variation between regions. In this case, the local variation in the Gabor filter output \( m \) will differ in the two regions. We refer to this fourth form of discontinuity as a step change in the average local output variation of \( m. \) Fig. 9(b) gives an example of this type of signature.

With a step, valley, or ridge discontinuity, standard image segmentation techniques can be used to locate the discontinuity. This is not the case, however, with a change in average local output variation without some further filtering. One possible solution is to transform this quantity into a change in mean value. Turner [31] encountered this problem and suggested using a bandpass filter for detecting such local changes. For our situation, this would involve applying a second Gabor filter to the output \( m \) of the first. If the variations within two regions have similar frequency content, then the second filter output would be proportional to the magnitude of the variation. Thus, a difference in average output variation would translate into a difference in mean for the second filter’s output.

Another simple method for transforming a difference in average local output variation into a difference in mean is to do the following:

1) remove the dc component from the filter output \( m \) by subtracting the mean value of \( m, \)
2) rectify the result by computing the magnitude, and
3) smooth the rectified output by applying a lowpass filter.

This sequence of operations can be written concisely as follows:

\[ \operatorname{LPF}[|m(x,y) - \mu_m|], \]

(35)

where \( \mu_m \) is the mean value of \( m \) and \( \operatorname{LPF} \) is a lowpass filter. This method does not make any assumptions on the frequency content of the input. Other methods are possible.

Reference [35] examines nonuniformly textured images in depth. It gives a numerical technique for measuring the discriminability of particular nonuniform-texture pairs, which in turn leads to a method for selecting appropriate Gabor filter parameters for discriminating the textures.

V. PARAMETER SELECTION

Section IV showed that when a properly tuned Gabor filter is applied to a textured image, distinct output signatures arise
A. Texels in Two Textures Differ

Assume that the texels $t_1$ and $t_2$ differ. Thus, the discussion focuses on the design of a Gabor filter tuned for producing a step-signature output. Also, assume for the time being that the texel spacings for the two textures are identical, i.e., $\Delta x_1 = \Delta x_2 = \Delta x$ and $\Delta y_1 = \Delta y_2 = \Delta y$. This condition will later be relaxed.

The parameters to select for the Gabor filter are $(\sigma_x, \sigma_y, \lambda, (U, V), \theta)$. Fig. 2 illustrates the relationship between filter size and texel spacing. The ellipse represents the one-standard-deviation contour of the Gaussian envelope of a Gabor filter, i.e., $(x, y)$, such that $(x/\sigma_x)^2 + (y/\sigma_y)^2 = 1$. The positioning of the ellipse at point $(x, y)$ represents the position of the GEF when the Gabor filter output $m$ is computed at point $(x, y)$.

The choice of $\sigma_x$ and $\sigma_y$ is a trade-off between Gabor filter output variation and accurate boundary localization. When $\sigma_x > \Delta x$ and $\sigma_y > \Delta y$, the filter envelope encompasses multiple texels, regardless of its position in the image. Although the positions of the texels vary within the envelope as the filter progresses across the image, the Gabor filter output $m$...
remains approximately constant over a region. If \( \sigma_y \ll \Delta x \) or \( \sigma_x \ll \Delta y \), the filter output depends on whether a texel occurs within the GEF envelope. This results in periodic Gabor filter output variations throughout a region. To avoid significant output variation, \( \frac{\sigma_y}{\Delta x} \) and \( \frac{\sigma_x}{\Delta y} \) should both be large; i.e., the GEF’s spatial extent should cover a number of texels. If \( (\sigma_x, \sigma_y) \) are large, though, near the texture boundary, the filter envelope will extend into both regions. This region overlap is what produces the sigmoid output transition described in Section IV-A. As \( (\sigma_x, \sigma_y) \) become larger, the transition becomes more gradual, making it more difficult to locate the texture boundary. In practice, we have found that filter performance is relatively insensitive to these ratios and that setting them to unity is a good compromise, i.e., \( \sigma_x = \Delta x \), \( \sigma_y = \Delta y \).

If the texel spacings in the \( x \) and \( y \) directions differ (\( \Delta x \neq \Delta y \)), but the two textures still use the same spacing, then, using the aforementioned design criteria, \( \sigma_x \neq \sigma_y \). Thus, the filter’s aspect ratio \( \lambda = \sigma_y/\sigma_x \neq 1 \), resulting in an asymmetric filter. For asymmetric filters, the orientation \( \theta \) of the Gaussian in the GEF (2) becomes an issue. Based on the discussion above, the Gaussian should be oriented to encompass, on the average, as many texels as possible. If the texels are spaced over a rectangular lattice, the Gaussian should be oriented along the \( x \)-axis and \( y \)-axis (i.e., \( \theta = 0 \) or \( \pi/2 \)). If the texels are not spaced over such a lattice, but are situated relative to some rotated coordinate system \((x', y')\), then the Gaussian should be oriented along the rotated axes. The orientation of the complex sinusoid \( \phi \) is determined by the Gabor filter center frequency \((U, V)\), and thus by the analysis in Section IV, and depends on the spectral differences between texels. Therefore, the choice of \( \theta \) is, in general, independent of \( \phi \).

The choice of center frequency \((U, V)\) depends on the texel spacing (which determines the harmonics) and on the spectral differences between texels at the harmonics. As discussed in Section IV-A, \((U, V)\) should be set to the harmonic that differs most in power between the texels in the two regions. Although two texels might differ more at some nonharmonic frequency, using this frequency as a filter center frequency in general produces an output signature that exhibits overshoot and/or undershoot. Such signatures have lower values within the textures than do the values produced by a properly tuned filter.

B. Texel Spacings Differ Between Textures

When texel spacing is the same in both regions, each texture has spectral energy centered about the same harmonics (cf. (17)), and a Gabor filter can be designed to produce step-signature outputs. If the texel spacings of the two regions differ, the harmonics from the different textures do not coincide. Since a Gabor-filter can be tuned to only one harmonic, signature distortion will result. In analyzing this distortion, note that the Gabor filter operation (prior to computing the magnitude) is linear, allowing the study of each region independently. Assume that the Gabor filter center frequency \((U, V)\) equals a harmonic of region 1. Then, the analysis proceeds as it does for the step signature. The frequency coordinates for the nearest corresponding harmonic of region 2 can be written as \((U+6\delta U, V+6\delta V)\), where \((6\delta U, 6\delta V)\) is the frequency offset between the harmonics of the two regions. Thus, the analysis of region 2 becomes analogous to that for the ridge signature. Combining the results for the two regions and computing the magnitude results in the following equation:

\[
m(x, y) = |s_f(x, y)| = 2\pi \sqrt{P_1 P_2^* + P_0 P_0^* + P_1 P_3^* + P_2 P_4^*},
\]

where \( P_0 = T_r(U, V) s_f(x, y) \), \( P_r = T_r(U+6\delta U, V+6\delta V) s_f(x-r, y) \), and \( s_f(x, y) \) is given by (23), and \( s_f(x-r, y) \) is the result of convolving a gate function with a GEF. For points far removed from the texture boundary, an analysis similar to that in Section IV reveals that (36) produces a step signature. The analysis near the texture boundary, however, is complicated because of interactions between complex variables. Although a detailed analysis is impractical, the presence of the GEF integral (i.e., \( s_f(x-r, y) \)) in (36) suggests that overshoot and/or undershoot can be expected (44). An example in Section VI corroborates this observation.

VI. RESULTS

All images used in the examples to follow consist of two regions. With the exception of Fig. 10, each region is composed of a collection of synthesized texels. Each texel is formed from line segments 20 pixels long by 2 pixels wide. The average intensity difference between regions is minimized by using approximately the same number of pixels in each texel. The size of the images is \( 512 \times 512 \) pixels.

Except where noted, the Gabor filter parameters were determined as follows. Section V recommended that \( \sigma_x = \Delta x \) and \( \sigma_y = \Delta y \). For most of the examples, \( \Delta x = \Delta y \). Hence, \( \sigma_x = \sigma_y = 24 \) pixels. The aspect ratio \( \lambda = 1 \). The algorithm developed in [35] was used to determine the center frequency \((U, V)\). This algorithm finds the harmonic
(U = 2πk/Δx, V = 2πl/Δy) that produces the largest step (or the deepest valley, or the highest ridge for texture phase differences). In the figure captions, the Gabor filter center frequencies (U, V) are reported in polar coordinates (F, φ) so that the orientation of the filter’s sinusoid is explicit.

The input images are defined digitally. Thus, aliasing in the images is not an issue. Aliasing is an issue, however, for the GEF, because it must be sampled before applying it to the image. Since the GEF’s are not bandwidth-limited, some aliasing will occur, regardless of the sample rate. Bovik et al. derived the required sample rate for various percentages of alias energy [22]. In the examples, the GEF’s are sampled so that the energy due to aliasing is < 1%2. Since the GEF’s are not spatially limited, truncation is necessary. We truncate them to a width of 6σ, which represents an error of about 0.2%. Finally, all points within one-half the filter width from the boundary are discarded in the final output to eliminate the wraparound error that arises in discrete convolution.

A. Difference in Texel Type

Fig. 3(a) illustrates a uniformly textured image consisting of +’s and L’s. Fig. 3(b) gives a plot of a Gabor filter output m(x, y) versus x and y. The vertical axis gives m(x, y) (the maximum and minimum filter outputs are indicated on the axis), and the two axes approximately horizontal and into the page represent x and y. All Gabor filter outputs are depicted this way.

The shape of the profile is predominantly a step function, with some undershoot present. The output of a Canny edge detector [45] applied to the Gabor-filtered image is shown superimposed on the original image. See Fig. 3(c). As Fig. 3(b) and 3(c) indicate, the boundary between the two textured regions is well localized.

An estimate of step height in Fig. 3(b) can be found by using the equations for A1 (26) and A2 (27). These equations imply that the ratio |T1|/|T2| is a relative measure of the step height. Letting T1 correspond to a “+” and T2 correspond to an “L” for the image in Fig. 3(a) gives T1 = 22.38 and T2 = -1. Thus, the predicted step height is |T1|/|T2| = [22.38]/| -1| = 22.38. For Fig. 3(b), the ratio of left-region and right-region heights is 22.70, which is in good agreement with the predicted value.

In this example, undershoot occurs in the Gabor filter output, because T2 is negative. This phenomena is discussed in Section 4.A and in [44]. Although (25) demonstrates the possibility of undershoot, (22) must be evaluated to determine the position and extent of undershoot. Letting T1 = 22.38 and T2 = -1 in (22), an estimate of the signature in Fig. 3(b) was computed. Table II compares actual and predicted signature values at selected positions. (The actual values have been scaled by a constant factor for comparison.)

This example illustrates that even if the Gabor filter center frequency equals a harmonic, undershoot can occur, overshoot, however, cannot occur for this situation (see [44]).

Both overshoot and undershoot can occur if the Gabor filter is not tuned to a harmonic. Fig. 3(d) gives an example. This figure represents the output of a Gabor filter having center frequency (U = 0.0283, V = -0.0283) (input is Fig. 3(a)). The texel spacing is 24 pixels. Thus, the closest harmonic to the filter center frequency is k/Δx = l/Δy = 1/24 = 0.0417. The center frequency then is displaced 6U = 6V = 0.0134 cycles/pixel away from the nearest harmonic. Using these parameters, we computed estimates of the Gabor filter output m(x, y) at selected x [44]. As shown in Table III, the predicted values are in good agreement with the actual values of Fig. 3(d).

B. Difference in Texel Orientation

Fig. 4(a) shows a uniformly textured image consisting of texels that differ in orientation. Fig. 4(b) and 4(c) show the outputs of two Gabor filters that use the same center frequency, but use different values for σ. In Fig. 4(b), σ equals the texel spacing, and a smooth step signature is achieved. The region on the right produces the greatest Gabor filter output m, because the orientation of the Gabor filter sinusoid matches the texel orientations on the right. In Fig. 4(c), σ = 8, which is one-third of the texel spacing. For this small σ, ripple occurs in the output because the GEF does not cover multiple texels as it moves across the image. The resulting signature, though, exhibits a sharper step transition than the transition in Fig. 4(b).

C. Differences in Horizontal and Vertical Texel Spacing

Fig. 5(a) is equivalent to Fig. 4(a), with the exception that Δy = 2Δx for the two textures. Fig. 5(b) shows a corresponding Gabor filter output when its GEF uses a symmetrical Gaussian (i.e., σx = σy or λ = 1). Note the occurrence of significant ripple in the y direction. Fig. 5(c) gives the Gabor filter output for a filter with λ = 2 (σy = 2σx). For this case, the output is “smooth” in both x and y and resembles the smooth step of Fig. 3(b).

<table>
<thead>
<tr>
<th>Location</th>
<th>Actual</th>
<th>Predicted</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ region</td>
<td>20.00</td>
<td>22.40</td>
</tr>
<tr>
<td>Maximum undershoot</td>
<td>0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>L region</td>
<td>0.88</td>
<td>1.00</td>
</tr>
</tbody>
</table>

TABLE II COMPARISON OF ACTUAL AND PREDICTED GABOR FILTER OUTPUT VALUES FOR THE STEP SIGNATURE IN FIG. 3(b)

<table>
<thead>
<tr>
<th>Location</th>
<th>Actual</th>
<th>Predicted</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ region</td>
<td>33.50</td>
<td>38.20</td>
</tr>
<tr>
<td>Maximum overshoot</td>
<td>88.70</td>
<td>85.10</td>
</tr>
<tr>
<td>Maximum undershoot</td>
<td>0.11</td>
<td>0.80</td>
</tr>
<tr>
<td>L region</td>
<td>1.35</td>
<td>1.71</td>
</tr>
</tbody>
</table>

TABLE III COMPARISON OF ACTUAL AND PREDICTED GABOR FILTER OUTPUT VALUES FOR THE STEP SIGNATURE IN Fig. 3(d)
Fig. 3. Uniformly textured image consisting of '+'s and '-'s. $\Delta x = \Delta y = 24$ pixels. Fig. 3(b) is a 3-D plot of $m(x, y)$ versus $x$ and $y$. The vertical axis gives $m$. The axes that are approximately horizontal and into the page represent the $x$-axis and $y$-axis, respectively. Note that in the examples to follow, $\sigma_x = \sigma_y = \sigma$ unless otherwise stated. (a) Input image. (b) Gabor filter output $m$ exhibiting a step signature. Filter parameters: $F = 0.059$ cycles/pixel, $\phi = 135^\circ$, and $\sigma = 24$ pixels. (c) Superposition of the input image and the output (vertical white line) of a Canny edge detector applied to (b). (d) Gabor filter output $m$ exhibiting overshoot. Filter parameters: $F = 0.040$ cycles/pixel, $\phi = 135^\circ$, and $\sigma = 24$ pixels.

D. Texture Phase Differences

Fig. 6(a) consists of two identically textured regions, but the regions are shifted relative to each other in both the $x$ and $y$ directions ([36] gives a simpler example). The texel spacing is 24 pixels in both $x$ and $y$, and the region on the right is shifted $-8$ pixels in the $y$ direction and $-4$ pixels in the $x$ direction relative to the region on the left. Fig. 6(b) shows the output of a Gabor filter that is tuned to the frequency $(U = 0.0, V = -0.042)$. Note that a valley signature results.

Because the texel spacing is 24 pixels, the harmonics of the image are located at $(0.042k, 0.042l)$. Thus, the filter is tuned to a harmonic in both $u$ and $v$ with harmonic indices $k = 0$...
and $l = -1$. From (30), we get the following equation:

$$\psi = 2\pi [k_0 \delta x / \Delta x + i \delta y / \Delta y]$$
$$= 2\pi (-1)(-8/24) \text{ radians } \equiv 120^\circ.$$  

The expression for the relative depth of the valley is derived from (33), and results in $m(r/2, y) = 0.5C$, which represents a difference of only 7% from the actual valley depth.

Fig. 6(c) shows the result of processing Fig. 6(a) by using a Gabor filter that is tuned to a frequency significantly displaced from a harmonic. A ridge signature results. The harmonics are located at $(0.042, 0.042 - l)$, and the center frequency of the filter is $(U = 0.012, V = -0.041)$, so $k = 0$ and $l = 1$. $\sigma_x$ and $\sigma_y$ again equal 24. From (30), we get the following equation:

$$\psi = 2\pi [k_0 \delta x / \Delta x + i \delta y / \Delta y]$$
$$= 2\pi (1)(-8/24) \text{ radians } \equiv -120^\circ,$$

and the method developed in [44] predicts a ridge height of 2.973. The actual ridge height in Fig. 6(c) is 3.24, a 9% error.
E. Texel Spacing Difference Between Regions

Fig. 7(a) shows a uniformly textured image similar to Fig. 3(a), except that the texel spacing differs between the two regions. Fig. 7(b) shows a corresponding Gabor filter output, where the filter is tuned to a harmonic corresponding to the region of ‘+’s. Although the signature is predominantly a step, some undershoot is present near the texture boundary. Fig. 7(c) shows a similar filter output, but now the filter is tuned to a harmonic for the region of L’s. In this case, both overshoot and undershoot are present. Observe how in each case (Fig. 7(b) and 7(c)), the region producing the greatest filter response corresponds to the one whose harmonic
Fig. 6. Uniformly textured image with regions shifted both horizontally and vertically, thus producing a texture-phase discontinuity (see text). \( \Delta x = \Delta y = 24 \) pixels. (a) Input image. (b) Gabor filter output \( m \) exhibiting a valley signature. Filter tuned to a harmonic. Filter parameters: \( F = 0.042 \) cycles/pixel, \( \phi = -90.0^\circ \), and \( \sigma = 24 \) pixels. (c) Gabor filter output \( m \) exhibiting a ridge signature. Filter not tuned to a harmonic. Filter parameters: \( F = 0.0427 \) cycles/pixel, \( \phi = 74.05^\circ \), and \( \sigma = 24 \) pixels.

Fig. 8(a) depicts a nonuniformly textured image produced by introducing random orientations and positional perturbations into the texels (L’s and L’s) of Fig. 3(a). Fig. 8(b) shows a filter output. The random effects cause large fluctuations in the output. Fig. 8(c) shows the result of applying a Canny edge detector to Fig. 8(b). Because of the fluctuations, the detected boundary does not perfectly match the "actual" boundary. The predicted boundary is, for the most part, correct to within \( \pm 1/2 \) texel. For typical nonuniform textures (where the actual texture boundary is not well defined), matches the filter center frequency. Analysis presented in [44] verifies this empirical result.

F. Nonuniform Textures

Fig. 8(a) depicts a nonuniformly textured image produced by introducing random orientations and positional perturbations into the texels (L’s and L’s) of Fig. 3(a). Fig. 8(b) shows a filter output. The random effects cause large fluctuations in the output. Fig. 8(c) shows the result of applying a Canny edge detector to Fig. 8(b). Because of the fluctuations, the detected boundary does not perfectly match the "actual" boundary. The predicted boundary is, for the most part, correct to within \( \pm 1/2 \) texel. For typical nonuniform textures (where the actual texture boundary is not well defined),
Fig. 7. Uniformly textured image similar to Fig. 3, but with each textured region having a different texel spacing. \( \Delta r_1 = \Delta y_1 = 24 \) pixels, \( \Delta r_2 = \Delta y_2 = 32 \) pixels. (a) Input image. (b) Gabor filter output \( m \). Filter tuned to a harmonic corresponding to a region of ‘+’s (some undershoot present). Filter parameters: \( F = 0.059 \) cycles/pixel, \( \phi = 115^\circ \), and \( \sigma = 24 \) pixels. (c) Gabor filter output \( m \) (some undershoot and overshoot present). Filter tuned to a harmonic corresponding to region of ‘\( \backslash \)'s. Filter parameters: \( F = 0.095 \) cycles/pixel, \( \phi = 115^\circ \), and \( \sigma = 24 \) pixels.

such fluctuation in the computed texture boundary is expected.

Fig. 9(a) gives a nonuniformly textured image consisting of triangles and arrows. Fig. 9(b) shows a filter output exhibiting a step change in average local output variation. After applying (35), the change in average local output variation was transformed to the step signature shown in Fig. 9(c).

Fig. 10(a) shows an example of a natural texture pair taken from Brodatz [46]. The left region is “grass lawn” (D9), and the right region is “cotton canvas” (D77). Pursuant to Rao’s taxonomy [43], D9 is an example of a disordered texture, and
The examples above are meant to typify Gabor filter outputs, but there are exceptional cases. For example, if a filter is tuned to a frequency component that has similar magnitude in both textures, the envelope can be nondiscriminating; i.e., the filter is not appropriate for discriminating between these two textures. If the textures are uniform, then the envelope will be flat. If they are nonuniform, then the envelopes may exhibit many fluctuations and show no distinguishing characteristics between regions. Nonuniform textures can produce other exceptions. One common example occurs when a filter is tuned to a frequency band apparently not involved in determining the texture boundary. In this case, a discontinuity might occur at a location other than the texture boundary. This “problem” also exists for the human visual system in the form of optical illusions and the perception of structures within structures. Reference [35] gives an example that demonstrates this phenomenon.

VII. CONCLUSION

This paper provides mathematical and experimental evidence suggesting that the application of Gabor filters to textured images produces certain characteristic output signatures that are useful for segmenting the image. Detailed criteria were given for designing tuned Gabor filters that yield the characteristic signatures. We emphasize that the detailed analysis presented here is based on a mathematical model of a bipartite textured image. A similar analysis for natural textures is impractical because the large variability and imprecise definition of these textures. For arbitrary natural textures, though, considerable experimental evidence supports the validity of our model and the resulting predicted filter responses [35], [44]. Because of difficulties in analysis, we have developed an algorithm for designing filters to produce distinct signatures for arbitrary natural (or synthetic) texture pairs. This algorithm, presented in [35], uses a search strategy to find the filter whose output is most consistent with the design criteria presented here.

It is clear that in a truly autonomous texture-segmentation architecture (such as the human visual system), filters cannot be customized to individual textures. In principle, a bank of filters is required that spans the expected orientation and frequency domain of the textures of interest. Although we have not addressed the problem of filter bank configuration, it is interesting to contrast the popular “rosette” configuration to our findings [26], [32], [47]. The rosette configuration is basically an ad hoc means for selecting an array of filters that cover the 2-D (U, V) frequency plane. One formulation that directly leads to this pattern are Gabor wavelets [28], [32]. In the 2-D frequency plane, the rosette consists of overlapping filters whose center frequencies lie on concentric circles centered at the origin. This configuration spans 360 degrees of orientation and spans frequencies from dc upward to any desired resolution.

One limitation of the rosette pattern is that it does not allow independent selection of \((\sigma_x, \sigma_y)\) and \((U, V)\); in practice, however, this is not a problem. Recall that \((\sigma_x, \sigma_y)\) are
Fig. 9. Nonuniformly textured image consisting of arrows and triangles. \( \Delta x = \Delta y = 24 \) pixels. (a) Input image. (b) Gabor filter output \( m \) exhibiting a step change in average local output variation at the texture boundary. Filter parameters are \( F = 0.042 \) cycles/pixel, \( \phi = 90^\circ \) and \( \sigma = 16 \) pixels. (c) Resulting output after (35) is applied to (b). Note that the difference in average local output variation in (b) is transformed into a step signature.

related to the texel spacing. Since texel spacing changes with image scale and frequency content is proportional to scale, \((\sigma_x, \sigma_y)\) and \((U, V)\) are related. Thus, independent selection of center frequency and filter size might not be necessary. The rosette pattern also does not allow for varying filter asymmetry. We showed, however, that asymmetric filters can be beneficial when the texel spacings differ in \( x \) and \( y \). Often, in practice, though, smooth signatures can still be attained at some cost in boundary localization if the texel-spacing difference is disregarded. Thus, it appears that
the rosette configuration is reasonably consistent with our findings.

**APPENDIX I**

This appendix further discusses the choice of nonlinearity for the Gabor filter. (See (1).) As shown below, computing the magnitude after filtering results in a loss of information. In particular, the phase component of the filter output is discarded. In fact, some experiments have shown that the image phase component is more important in preserving image quality than is the amplitude component [48]. To avoid this information loss, some researchers have proposed methods for extracting phase information directly [22], [49].

The primary motivation for ignoring phase comes from psychophysical studies of the human visual system. Although neurophysiological evidence suggests that quadrature-pair filters might exist in the visual cortex [20] (thus enabling phase detection), certain psychophysical results suggest that humans do not encode phase information directly [50], at least not for texture segmentation. Consequently, some researchers have explicitly eliminated phase information from their texture-segmentation algorithms [27], [51]. Although, admittedly, information is lost by ignoring phase, some phase-related phenomena can be recovered directly from the amplitude envelope. (See Bovik [34] for a discussion of the effect of phase on the amplitude envelope.) This is because the phase and amplitude components are not independent; a change in one will produce a change in the other.

To further explore these points, suppose that instead of computing the magnitude of the GEF-filtered image, one simply demodulated this image and applied a low-pass filter. Demodulating the GEF-filtered image (21) essentially eliminates the complex exponential leaving a pair of offset gates with *complex* coefficients $T_1$ and $T_2$. Contrast this with (22), where the magnitude operation has been applied and only the magnitudes of $T_1$ and $T_2$ determine the gate amplitudes. If $T_1$ and $T_2$ differ only in sign, the Gabor filter output $m$ in (22) is nondiscriminating. This implies that if textures differ only in the sign of contrast, the textures cannot be discriminated. This is precisely the argument used by Malik and Perona in criticizing the magnitude computation [27]. Using the demodulation approach, though, the coefficients $T_1$ and $T_2$ after demodulation are complex. Hence, the filter output will reflect not only differences in sign but also differences in phase. Although this approach is more discriminating, the sensitivity to phase will lead to segmentations that are not consistent with human performance.

There is a method for retaining sign differences between $T_1$ and $T_2$ while ignoring phase information. The method involves convolving the image with only the real portion of a GEF. More precisely, define a new filter $h_r$ as a Gaussian modulated sinusoid (cf. (2)), as follows:

$$h_r(x, y) = g(x', y') \cos(2\pi(Ux + Vy) + \phi)$$

where $\phi$ is some arbitrary constant phase angle. Let $H_r$ be the Fourier transform of $h_r$:

$$H_r(u, v) = 1/2 (H_+ e^{-j\phi} + H_- e^{j\phi})$$

Fig. 10. Natural textures “grass lawn” (D9) and “cotton canvas” (D77) from Brodatz. Filter parameters are $F = 0.003$ cycles/pixel, $\sigma = 60.0^\circ$ and $\alpha = 30$ pixels. (a) Input image. (b) Gabor filter output exhibiting a step signature. (c) Canny edge-detector output superimposed on input image.
where
\[ H_+ = \exp \left\{ -\frac{1}{2} \left[ (\sigma_x[u + U])^2 + (\sigma_y[v + V])^2 \right] \right\} \]
\[ H_- = \exp \left\{ -\frac{1}{2} \left[ (\sigma_x[u - U])^2 + (\sigma_y[v - V])^2 \right] \right\}. \]

Note that \( H_+ \) is similar to (4), except that it is symmetric about the frequency origin. Applying \( H_+ \) to \( I \) in (13) (cf. (18–19)) gives the following:
\[ I_+(u, v) = \frac{1}{2\Delta x \Delta y} \left\{ I_+ + I_- \right\} \quad (39) \]

where
\[ I_+ = H_+ e^{-j\delta S(u + U, V + V)} \left\{ T_1(U, V) + T_2(U, V) e^{-j\phi(u + U)} \right\} \]
\[ I_- = H_- e^{j\delta S(u - U, V - V)} \left\{ T_1(-U, -V) + T_2(-U, -V) e^{-j\phi(u - U)} \right\}. \]

Note that \( I_+ \) is equivalent to the sum of \( I_f \) in (18) and the mirror image of \( I_f \).

We now demodulate and low-pass filter \( I_+ \). This is equivalent to shifting both \( I_f \) and the mirror image of \( I_f \) to the origin and summing them. This results in the following equation:
\[ I_0(u, v) \approx K \left\{ T_1(U, V) + T_2(U, V) e^{-j\phi(u + U)} \right\} \]
\[ + K \left\{ T_1(-U, -V) + T_2(-U, -V) e^{-j\phi(u - U)} \right\} \quad (40) \]

where
\[ K = \frac{1}{2\Delta x \Delta y} \exp \left\{ -\frac{1}{2} \left[ (\sigma_x)^2 + (\sigma_y)^2 \right] \right\} S(u, v). \]

Real images \( T_1(U, V) \) and \( T_2(-U, -V) \) are complex conjugates, as are \( T_2(U, V) \) and \( T_2(-U, -V) \). Therefore, (40) reduces to the following expression:
\[ I_0(u, v) \approx 2K \left\{ \text{Re}[T_1] + \text{Re}[T_2] e^{-j\phi(u)} \right\}. \quad (41) \]

By arguments similar to those used for deriving the step signature in Section IV.A, the final output of the alternate filter \( I_0(x, y) \), which is given by the inverse Fourier transform of (41), approximates two offset gate functions coincident with \( I_f \). The amplitudes of the gates are proportional to the real parts of \( T_1 \) and \( T_2 \) rather than to their magnitude. Thus, the sign is preserved without retaining phase information. This approach, then, can discriminate textures whose textels differ only in the sign of contrast.

It is important to point out that all of the operations used here are linear, and, as mentioned in Section II, that some form of nonlinearity is essential in simulating human performance (assuming that this is desirable). If a suitable nonlinearity could be found and imposed after demodulation, this method could provide an alternative to the more elaborate architecture proposed by Malik and Perona [27].

Farrokhnia and Jain used \( h_r \) in their texture-segmentation work, but still employed the magnitude computation [24], [26]. As we have just shown, though, the sign information is still lost with this approach. Our contribution here is in analytically showing the potential of demodulating the filter output and contrasting this approach with other methods. It is important to realize that simply using \( h_r \) as a filter does not guarantee that sign information will be preserved.

REFERENCES


[40] B. Julesz, personal communication cited in Malik and Perona [27].


[51] J. Wakeley received the B.S.E.E. degree from the University of Pennsylvania in 1959, and the M.S.E.E. degree from the Polytechnic Institute of Brooklyn, Brooklyn, NY, in 1964. He is currently a Research Associate at the Applied Research Laboratory at The Pennsylvania State University, University Park, PA. His research interest is in the application of human vision to the processing of underwater acoustic reverberation.