Image demodulation using multidimensional energy separation

Petros Maragos

School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332

Alan C. Bovik

Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, Texas 78712

Received January 3, 1995; revised manuscript received April 18, 1995; accepted April 19, 1995

Locally narrow-band images can be modeled as two-dimensional (2D) spatial AM–FM signals with several applications in image texture analysis and computer vision. We formulate an image-demodulation problem and present a solution based on the multidimensional energy operator $\Phi(f) = \|\nabla f\|^2 - f\nabla^2 f$. This nonlinear operator is a multidimensional extension of the one-dimensional (1D) energy-tracking operator $\Psi(f) = (f')^2 - ff''$, which has been found useful for demodulating 1D AM–FM and speech signals. We discuss some interesting properties of the multidimensional operator and develop a multidimensional energy-separation algorithm to estimate the amplitude envelope and instantaneous frequencies of 2D spatially varying AM–FM signals. Experiments are also presented on applying this 2D energy-demodulation algorithm to estimate the instantaneous amplitude contrast and spatial frequencies of image textures bandpass filtered by means of Gabor filters. The attractive features of the multidimensional energy operator and the 2D energy-separation algorithm are their simplicity, efficiency, and ability to track instantaneously varying spatial-modulation patterns.

1. INTRODUCTION

Image textures of the locally narrow-band type can be modeled as two-dimensional (2D) spatial AM–FM signals,

$$f(x, y) = a(x, y)\cos[\phi(x, y)], \tag{1}$$

that are 2D sines containing both amplitude modulation (AM) and frequency modulation (FM). That is, they have a spatially varying amplitude $a(x, y)$ and a spatially varying instantaneous frequency vector $\boldsymbol{\omega}(x, y) = \nabla\phi(x, y)$. Of course, given a signal f, the amplitude and phase signals a and ϕ can be defined in an infinite number of ways. However, only certain interpretations of them are meaningful in modeling locally narrow-band images. In particular, the amplitude is used to model local image contrast, and the frequency vector contains rich information about the locally emergent spatial frequencies. Thus it is reasonable to assume that the amplitude $a(x, y)$ and the frequency vector $\boldsymbol{\omega}(x, y)$ are locally narrow-band signals and hence locally smooth. Such modulation models have been proposed by Bovik *et al.*[1] and have been applied to a variety of image processing and vision problems. In Refs. 1 and 2 these models are not applied directly to the whole (possibly wideband) image. Instead they are used on its bandpass-filtered versions that are outputs from a filter bank consisting of 2D Gabor filters. Some useful consequences of the bandpass filtering are an increased noise tolerance and the enforcement of some smoothness on the amplitude and frequency signals. The motivation for using Gabor filters is their attaining the lower limit of joint space-frequency resolution uncertainty and their ability to model early filtering stages of human vision.[3]

An important problem in modeling image modulations with spatial AM–FM signals is to estimate the 2D amplitude and frequency signals with computational vision algorithms that have low complexity and small estimation error. In this paper we develop such an efficient approach for demodulation of 2D AM–FM signals based on a multidimensional energy operator introduced by Maragos *et al.*[4] and on a related 2D energy-separation algorithm. Our work has been inspired by similar work for one-dimensional (1D) signal and speech processing, in which a 1D energy-tracking operator[5,6] was used to develop a 1D energy-separation algorithm[7] with applications to speech and AM–FM signal demodulation.[8–10]

We begin in Section 2 by summarizing the basic ideas behind the 1D energy operator and energy separation. Then in Section 3 we introduce an energy operator for multidimensional continuous-domain signals, derive many of its useful properties, and show how it can be used to estimate the parameters of multidimensional sinusoids. Extending the theoretical analysis to multidimensional AM–FM signals, we derive an algorithm that can demodulate a multidimensional AM–FM signal and estimate its amplitude envelope and instantaneous frequencies. By using the multidimensional framework, we also generalize the energy operator to vector-valued image signals. For 2D signals, if we replace the partial derivatives in the 2D energy operator with one-sample differences, we obtain a discrete-domain 2D energy operator, which is identical to the one developed by Yu *et al.*[11] for digital image edge detection and was also used in Ref. 12 for image enhancement. In Section 4 we repeat the theoretical analysis for the 2D discrete-energy operator and derive a related 2D discrete AM–FM demodulation algorithm. Finally, Section 5 discusses our experiments in using the 2D discrete-energy operator and energy-separation algorithm to demodulate image textures into spatially varying amplitude and frequency image components.

0740-3232/95/091867-10$06.00

2. ONE-DIMENSIONAL ENERGY OPERATOR AND ENERGY SEPARATION

In their work on nonlinear modeling of speech production[13] Teager and Teager developed a nonlinear differential operator Ψ for 1D continuous-time signals $f(t)$, defined as

$$\Psi(f)(t) \triangleq [f'(t)]^2 - f(t)f''(t), \tag{2}$$

where $f' = \mathrm{d}f/\mathrm{d}t$ and $f'' = \mathrm{d}^2f/\mathrm{d}t^2$. The discrete-time counterpart of Ψ is the operator

$$\Psi_d(f)(n) \triangleq f^2(n) - f(n-1)f(n+1) \tag{3}$$

for discrete-time signals $f(n)$. Both operators were first introduced systematically by Kaiser.[5,6] Ψ is an energy-tracking operator because it can track the energy of simple harmonic oscillators that produce sinusoidal oscillatory signals; this energy is proportional to both the amplitude squared and the frequency squared of the oscillation. Hence we shall refer to Ψ and Ψ_d as the 1D *energy operators*.

The energy operators are very efficient in instantaneously estimating the modulating signals of 1D AM–FM signals. Specifically, Maragos *et al.*[8,9] showed that the energy operators can approximately estimate the envelope of AM signals and the instantaneous frequency of FM signals. For 1D AM–FM signals,

$$f(t) = a(t)\cos[\phi(t)], \tag{4}$$

they have also found that the energy operator tracks the energy product

$$\Psi\{a(t)\cos[\phi(t)]\} \approx a^2(t)\omega^2(t), \tag{5}$$

where $\omega(t) = \mathrm{d}\phi(t)/\mathrm{d}t$ is the instantaneous (angular) frequency. This approximate result is valid (i.e., the approximation error is negligible) if the time-varying amplitude $a(t)$ and frequency $\omega(t)$ do not vary too fast in time or too greatly in value compared with the carrier. A usual way that this can happen is when the amount of modulation is small and the bandwidths of the amplitude and the frequency modulating signals are much smaller than the carrier frequency. Further, by applying Ψ to the derivative $f'(t)$ and combining the energy output with Eq. (4), Maragos *et al.* also developed an *energy-separation algorithm*[7,10] (ESA) that separates the energy product [Eq. (4)] into amplitude and frequency components. Thus the ESA fully demodulates the AM–FM signal by estimating its amplitude envelope $|a(t)|$ and instantaneous frequency $\omega(t)$. Similar results and algorithms have been derived for discrete-time AM–FM signals.[10]

The energy operator and the ESA are very efficient for AM–FM demodulation. So far their major application has been the instantaneous tracking of modulations in speech resonances, which are modeled as AM–FM signals[8–10]; such a model was motivated by several nonlinear fluid-dynamic phenomena during speech production.[13] The effects of noise and Gabor or other signal prefiltering on the performance of the energy operator and the ESA have been studied in detail in Ref. 14.

3. CONTINUOUS-DOMAIN MULTIDIMENSIONAL ENERGY OPERATOR

Let $f(\mathbf{x})$ be a d-dimensional real-valued signal with a continuous argument $\mathbf{x} = (x_1, \ldots, x_d) \in \mathbf{R}^d$, $d = 2, 3, \ldots$. Then we define the d-dimensional energy operator by

$$\Phi(f)(\mathbf{x}) \triangleq \|\nabla f(\mathbf{x})\|^2 - f(\mathbf{x})\nabla^2 f(\mathbf{x}), \tag{6}$$

where

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_d}\right)$$

is the gradient of f,

$$\|\nabla f\|^2 = \left(\frac{\partial f}{\partial x_1}\right)^2 + \ldots + \left(\frac{\partial f}{\partial x_d}\right)^2$$

is the Euclidean norm squared of the gradient, and

$$\nabla^2 f = \sum_{k=1}^{d} \frac{\partial^2 f}{\partial {x_k}^2}$$

is the Laplacian of f. From its definition it follows directly that we can express $\Phi(f)$ as

$$\Phi(f) = \sum_{k=1}^{d} \left(\frac{\partial f}{\partial x_k}\right)^2 - f\left(\frac{\partial^2 f}{\partial {x_k}^2}\right) = \sum_{k=1}^{d} \Psi_k(f), \tag{7}$$

where

$$\Psi_k(f) \triangleq \left(\frac{\partial f}{\partial x_k}\right)^2 - f\,\frac{\partial^2 f}{\partial {x_k}^2}. \tag{8}$$

Thus the output of the Φ is a sum of energy components. Each energy component is the output of the 1D energy operator Ψ applied along each of the d directions x_k. Hence, in analogy with the 1D case, we shall refer to Φ as the multidimensional energy operator. Next we derive a few of its properties.

Let us apply Φ to the product and the sum of two d-dimensional signals $f(\mathbf{x})$ and $g(\mathbf{x})$. Expressing $\Phi(f)$ as $\Phi(f) = f \cdot f - f\nabla^2 f$, where $(\cdot)$ denotes inner product, and using the general facts

$$\nabla(fg) = g\nabla f + f\nabla g,$$
$$\nabla^2(fg) = g\nabla^2 f + f\nabla^2 g + 2(\nabla f)\cdot(\nabla g)$$

yields the product rule

$$\Phi(fg) = f^2\Phi(g) + g^2\Phi(f). \tag{9}$$

The sum rule is also simple to derive:

$$\Phi(f+g) = \Phi(f) + \Phi(g) + 2(\nabla f)\cdot(\nabla g) - f\nabla^2 g - g\nabla^2 f. \tag{10}$$

As a special case, applying Φ to a signal $f(\mathbf{x})$ plus constant c yields

$$\Phi[f(\mathbf{x}) + c] = \Phi(f)(\mathbf{x}) - c\nabla^2 f(\mathbf{x}). \tag{11}$$

For a multidimensional exponential signal the output of the energy operator is identically zero:

$$\Psi[\exp(\mathbf{c} \cdot \mathbf{x})] = 0\,, \tag{12}$$

where $\mathbf{c} = (c_1, \ldots, c_d)$ is an arbitrary constant vector and $\mathbf{c} \cdot \mathbf{x} = \sum_{k=1}^{d} c_k x_k$. Combining Eq. (12) with Eq. (9) implies that we can extend all the results in this paper to signals that contain an arbitrary constant scaling factor A and/or a multiplicative exponential component, because

$$\Psi[A \exp(\mathbf{c} \cdot \mathbf{x}) f(\mathbf{x})] = A^2 \exp(2\mathbf{c} \cdot \mathbf{x}) \Psi[f(\mathbf{x})]. \tag{13}$$

A. Cosines with Constant Amplitude and Constant Frequencies

Applying Ψ to a d-dimensional cosine

$$f(\mathbf{x}) = A \cos(\boldsymbol{\omega}_c \cdot \mathbf{x} + \theta) \tag{14}$$

with constant phase offset θ, constant amplitude A, and constant-frequency vector

$$\boldsymbol{\omega}_c = (\omega_{c,1}, \ldots, \omega_{c,d}) \tag{15}$$

yields

$$\Psi[A \cos(\boldsymbol{\omega}_c \cdot \mathbf{x} + \theta)] = A^2 \left(\sum_{k=1}^{d} {\omega_{c,k}}^2 \right) = A^2 \|\boldsymbol{\omega}_c\|^2. \tag{16}$$

Thus when Ψ is applied to a multidimensional cosine it yields the product of the amplitude squared and the frequency-vector norm squared.

Now to estimate the individual $d + 1$ parameters $|A|$, $\omega_{c,1}, \ldots, \omega_{c,d}$ we also apply Ψ to the d partial derivatives

$$\frac{\partial f}{\partial x_k}(\mathbf{x}) = \frac{\partial [A \cos(\boldsymbol{\omega}_c \cdot \mathbf{x} + \theta)]}{\partial x_k} = -A\omega_{c,k} \sin(\boldsymbol{\omega}_c \cdot \mathbf{x} + \theta) \tag{17}$$

of cosine f. Then, by Eq. (16),

$$\Psi\left(\frac{\partial f}{\partial x_k} \right)(\mathbf{x}) = (A\omega_{c,k})^2 \|\boldsymbol{\omega}_c\|^2 \tag{18}$$

for all $k = 1, \ldots, d$. By combining Eqs. (16) and (18) we obtain the following $d + 1$ equations for exact estimation of the absolute amplitude and d absolute frequencies:

$$|\omega_{c,k}| = \left[\Psi\left(\frac{\partial f}{\partial x_k} \right) \Big/ \Psi(f) \right]^{1/2}, \qquad k = 1, 2, \ldots, d\,, \tag{19}$$

$$|A| = \Psi(f) \Big/ \left[\sum_{k=1}^{d} \Psi\left(\frac{\partial f}{\partial x_k} \right) \right]^{1/2}. \tag{20}$$

We call the above equations the multidimensional *continuous energy-separation algorithm (CESA)*. They are an extension of the 1D CESA developed in Refs. 7 and 10.

In 1D cosines $x(t) = \cos(\omega_1 t + \theta)$ it does not matter whether ω_1 is positive or negative; we can always assume that it is positive, because even if $\omega_1 < 0$ we can write $x(t) = \cos(|\omega_1| t - \theta)$ and hence absorb the negative sign in the phase offset θ. However, in 2D cosines $f(x, y) = \cos(\omega_1 x + \omega_2 y + \theta)$ the relative signs of ω_1 and ω_2 matter, because they control the direction of the frequency vector, which can lie in any of the four quadrants. Without loss of generality we can always assume that ω_1 is positive. But we need to know whether ω_2 is positive or negative, and the 2D ESA can give us only its absolute value.

B. AM–FM Signals

Consider the real-valued d-dimensional AM–FM signal

$$f(\mathbf{x}) = a(\mathbf{x}) \cos[\phi(\mathbf{x})], \tag{21}$$

where $a(\mathbf{x})$ is the spatially varying amplitude, $\phi(\mathbf{x})$ is the phase signal,

$$\boldsymbol{\omega}(\mathbf{x}) \stackrel{\triangle}{=} \nabla \phi(\mathbf{x}) = [\omega_1(\mathbf{x}), \ldots, \omega_d(\mathbf{x})] \tag{22}$$

is the spatially varying d-dimensional instantaneous-frequency vector, and

$$\omega_k(\mathbf{x}) \stackrel{\triangle}{=} \frac{\partial \phi}{\partial x_k}(\mathbf{x}) \tag{23}$$

is the kth instantaneous angular-frequency signal. Assuming for each k that $\omega_k(\mathbf{x})$ is either nonnegative or nonpositive (but not of mixed sign) for all $\mathbf{x}$, we can express it as

$$\omega_k(\mathbf{x}) = \omega_{c,k} + \omega_{m,k} q_k(\mathbf{x})\,, \tag{24}$$

where $\omega_{c,k}$ is a constant-center frequency (usually referred to as the carrier frequency), $q_k(\mathbf{x}) \in [-1, 1]$ is the kth normalized frequency-modulating signal, and $\omega_{m,k}$ is the maximum deviation of ω_k from its center value. Henceforth we assume that $0 \le \omega_{m,k} \le |\omega_{c,k}|$.

Applying Ψ to f yields

$$\Psi[a \cos(\phi)] = a^2 \|\boldsymbol{\omega}\|^2 - \tfrac{1}{2} a^2 \sin(2\phi) \nabla^2 \phi + \cos^2(\phi) \Psi(a)\,. \tag{25}$$

For demodulation the desired term in Eq. (25) is $a^2 \|\boldsymbol{\omega}\|^2$. We view the rest of the terms as approximation error and show next that they are negligible under realistic assumptions.

Assume that $a(\mathbf{x})$ is band limited in a circular frequency sphere of radius ω_a. That is, if $A(\mathbf{u})$ is its d-dimensional Fourier transform, then $A(\mathbf{u}) = 0$ for $\|\mathbf{u}\| > \omega_a$. Then if we define the mean spectral absolute value of a as

$$\mu_a = \frac{1}{(2\pi)^d} \int_{-\omega_a}^{\omega_a} \cdots \int_{-\omega_a}^{\omega_a} |A(\mathbf{u})| \mathrm{d}u_1 \ldots \mathrm{d}u_d\,, \tag{26}$$

it can be shown that for each k

$$|a(\mathbf{x})| \le a_{\max} \le \mu_a\,, \tag{27}$$

$$\left| \frac{\partial a}{\partial x_k} \right| \le \omega_a \mu_a\,, \tag{28}$$

$$\left| \frac{\partial^2 a}{\partial {x_k}^2} \right| \le {\omega_a}^2 \mu_a\,, \tag{29}$$

$$|\Psi(a)(\mathbf{x})| \le 2d {\omega_a}^2 {\mu_a}^2\,, \tag{30}$$

where $a_{\max} = \sup_{\mathbf{x}} |a(\mathbf{x})|$. Assume also that each fre-

quency signal $\omega_k(\mathbf{x})$ is band limited with bandwidth $\omega_{f,k} < |\omega_{c,k}|$. Then we can consider the approximation

$$\Phi[a\cos(\phi)] \approx a^2\|\boldsymbol{\omega}\|^2, \tag{31}$$

with an approximation error

$$E(\mathbf{x}) = \Phi[a\cos(\phi)] - a^2\|\boldsymbol{\omega}\|^2 \tag{32}$$

that is bounded by

$$|E(\mathbf{x})| \le \left(2d\omega_a{}^2 + \frac{1}{2}\sum_{k=1}^{d}\omega_{m,k}\omega_{f,k}\mu_{q_k}\right)\mu_a{}^2, \tag{33}$$

given that

$$\frac{\partial^2\phi}{\partial x_k{}^2} = \omega_{m,k}\frac{\partial q_k}{\partial x_k}. \tag{34}$$

Assuming that $a_{\max} \approx \mu_a$ (which is true with equality if a is a cosine or has linear Fourier phase), the realistic conditions

$$\omega_a \ll \min_k|\omega_{c,k}|, \qquad \sum_{k=1}^{d}\omega_{m,k}\omega_{f,k} \ll \|\omega_c\|^2 \tag{35}$$

make the maximum absolute value of the error E much smaller than the maximum absolute value of the desired term. Thus, under such conditions, approximation (31) is valid in the sense that the relative error is $\ll 1$. Note that conditions (35) imply that the amplitude and frequency signals do not vary too fast in space or too greatly in value compared with the carriers.

Now let us apply Φ to the partial derivatives

$$\frac{\partial f}{\partial x_k} = \frac{\partial a}{\partial x_k}\cos(\phi) - a\omega_k\sin(\phi). \tag{36}$$

Owing to conditions (35), the maximum absolute value of the second term in $\partial f/\partial x_k$ has a much larger order of magnitude than does that of the first term. Thus we approximate $\partial f/\partial x_k \approx -a\omega_k\sin(\phi)$ and apply approximation (31) to obtain

$$\Phi\left(\frac{\partial f}{\partial x_k}\right) \approx \Phi[a\omega_k\sin(\phi)] \approx a^2\omega_k{}^2\|\boldsymbol{\omega}\|^2 \tag{37}$$

for each k. Combining approximations (31) and (37) yields the following CESA:

$$\left[\Phi\left(\frac{\partial f}{\partial x_k}\right)\Big/\Phi(f)\right]^{1/2} \approx |\omega_k(\mathbf{x})|, \qquad k = 1, 2, \ldots, d, \tag{38}$$

$$\Phi(f)\Big/\left[\sum_{k=1}^{d}\Phi\left(\frac{\partial f}{\partial x_k}\right)\right]^{1/2} \approx |a(\mathbf{x})|. \tag{39}$$

This algorithm can estimate at each location $\mathbf{x}$ the amplitude envelope and the magnitude of instantaneous frequencies of the spatially varying AM–FM signal.

To recover the sign of the instantaneous-frequency signals $\omega_k(\mathbf{x})$, we consider their decomposition [Eq. (24)] into a constant term $\omega_{c,k}$, which we usually assume to be equal to the local mean of $\omega_k(\mathbf{x})$, and a spatially varying term $\omega_{m,k}q_k(\mathbf{x})$ whose magnitude does not exceed $|\omega_{c,k}|$. Thus, if we *a priori* know the carrier frequencies $\omega_{c,k}$, their signs reveal the signs of the instantaneous-frequency signals since the latter are assumed to possess a constant sign equal to that of their mean values (i.e., the carriers). Such an *a priori* knowledge of the signs of the carrier frequencies $\omega_{c,k}$ can be obtained, for example, in applying the 2D ESA to bandpass-filtered versions of the original signal, in which case the center-frequency vector of the bandpass filter is taken to be approximately equal to the carrier-frequency vector (assuming a symmetry of the bandpass filter around its center frequency).

C. Energy Operators for Vector-Valued Signals

The framework of the multidimensional energy operator is also useful in developing energy operators for vector-valued signals.

Consider a 1D vector-valued signal

$$\mathbf{f}(t) = [f_1(t), f_2(t), \ldots, f_n(t)], \tag{40}$$

where all n components are real valued. Define its vector derivative

$$\mathbf{f}' = (f_1', f_2', \ldots, f_n') \tag{41}$$

and its second derivative $\mathbf{f}'' = (\mathbf{f}')'$. Then we define an energy operator for vector-valued signals:

$$\Theta(\mathbf{f})(t) \triangleq \|\mathbf{f}'(t)\|^2 - \mathbf{f}(t)\cdot\mathbf{f}''(t). \tag{42}$$

It is simple to show that

$$\Theta(\mathbf{f}) = \sum_{k=1}^{n}\Psi(f_k). \tag{43}$$

Hence the energy of the vector-valued signal is the sum of the energies of its scalar component signals.

Next we outline several applications of energy operators for vector-valued signals.

1. Complex Signals

Let $\mathbf{f}(t)$ be a 1D complex-valued signal. We can define an energy operator for complex-valued signals $\mathbf{f}$ as

$$C(\mathbf{f})(t) \triangleq \|\mathbf{f}'(t)\|^2 - \mathrm{Re}[\mathbf{f}^*(t)\mathbf{f}''(t)], \tag{44}$$

where $()^*$ denotes complex conjugate. This complex operator has interesting properties and applications to image texture analysis. Now if we form a vector-valued signal with $n = 2$, $f_1 = \mathrm{Re}(\mathbf{f})$, and $f_2 = \mathrm{Im}(\mathbf{f})$, it follows that

$$\begin{aligned} C(\mathbf{f}) &= \Theta\{[\mathrm{Re}(\mathbf{f}), \mathrm{Im}(\mathbf{f})]\} \\ &= \Psi[\mathrm{Re}(\mathbf{f})] + \Psi[\mathrm{Im}(\mathbf{f})]. \end{aligned} \tag{45}$$

Thus the analysis of the application of energy operators to complex-valued signals can be reduced to simply analyzing separately their real and imaginary parts.

2. System of n Oscillators

Let $\mathbf{f}(t)$ be the position vector tracing a continuous smooth-motion curve in n-dimensional space, where each component motion is due to one from a system of n oscillators. Then $\mathbf{f}'(t)$ is the velocity vector tangent to the curve and $\mathbf{f}''(\mathbf{t})$ is the acceleration vector. Then $\|\mathbf{f}'\|^2$ is

kinetic energy (per unit mass) of the system, and $-\mathbf{f} \cdot \mathbf{f}''$ is potential energy.

3. Multispectral Images

Multispectral images with n spectral channels can be represented by 2D vector-valued signals $\mathbf{f}(x, y)$. The case $n = 2$ can be used for complex-valued 2D images. The energy operator Θ can be extended directly to 2D vector-valued signals by replacement of the vector derivative $\mathbf{f}'$ with the $n \times 2$ matrix derivative $[\partial f_i/\partial x_j]$, $i = 1, \ldots, n$, where $x_1 = x$ and $x_2 = y$, the Euclidean norm for vectors with the Frobenious norm for matrices, and the vector 2nd derivative $\mathbf{f}''$ with the vector Laplacian $(\nabla^2 f_1, \ldots, \nabla^2 f_n)$.

4. DISCRETE-SPACE ENERGY OPERATOR FOR IMAGE SIGNALS

In general, if we replace derivatives in Φ with one-sample differences we obtain a discrete-space energy operator. For notational simplicity we restrict our discussion to 2D signals, e.g., still images.

The alternative interpretation, Eq. (7), of Φ as a sum of energy components along different directions allows us to extend it to discrete-space signals $f(m, n)$. Specifically, replacing each of these energy components with outputs from 1D discrete-time energy operators Ψ_d yields the following 2D discrete-space energy operator:

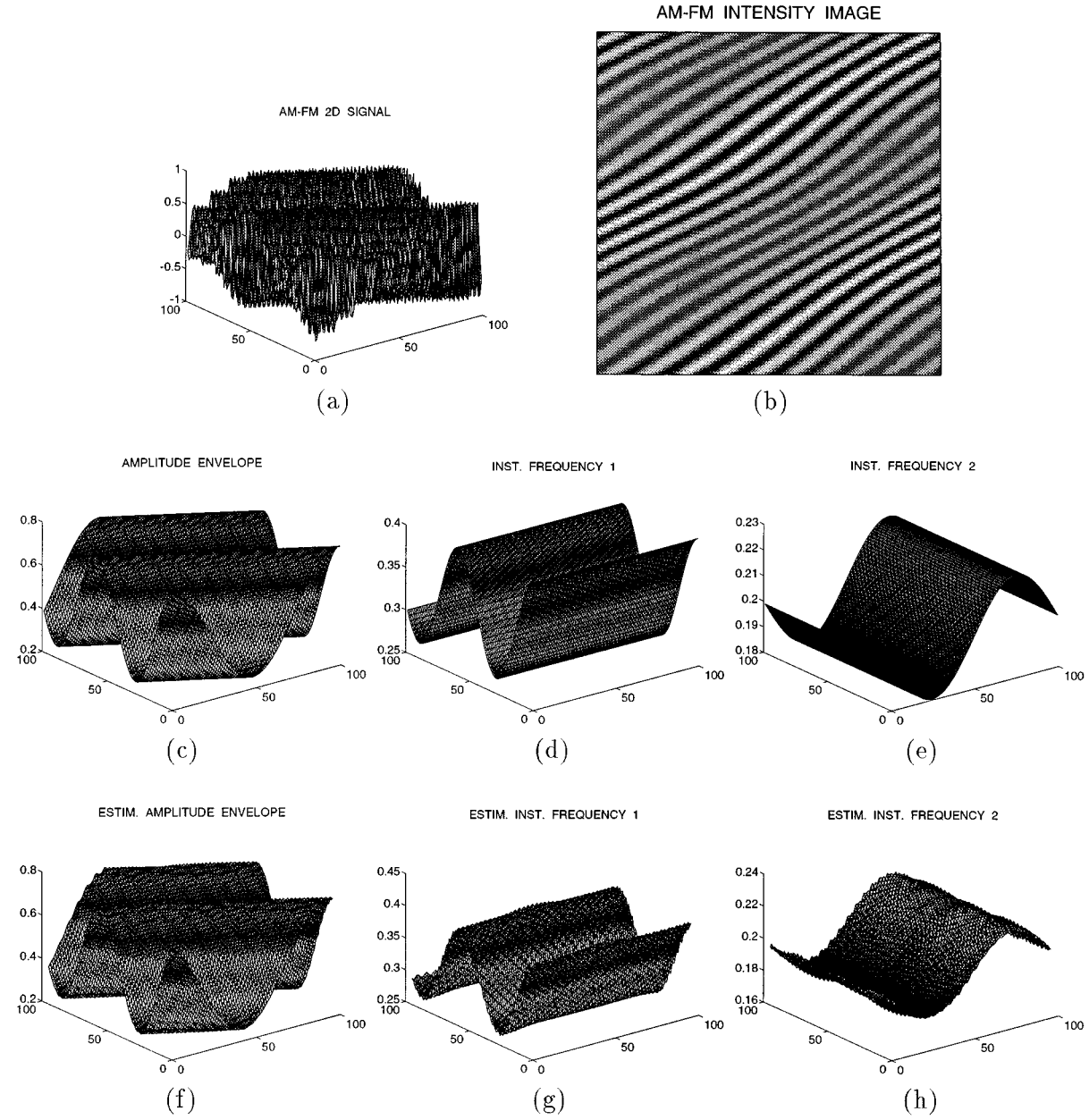

Fig. 1. (a) Perspective plot of original 2D AM–FM signal

$$f(m, n) = 0.5\left[1 + 0.5\cos\left(\frac{\pi}{30}m + \frac{\pi}{50}n\right)\right]\cos\left[\frac{\pi}{3}m + \frac{\pi}{5}n + 2\sin\left(\frac{\pi}{30}m\right) + \sin\left(\frac{\pi}{50}n + \frac{\pi}{2}\right)\right], \qquad m, n = 1, \ldots, 100;$$

(b) intensity image of the AM–FM signal f; (c) original amplitude envelope $a(m, n)$; (d) original instantaneous frequency $\Omega_1(m, n)/\pi$; (e) original instantaneous frequency $\Omega_2(m, n)/\pi$; (f) amplitude envelope estimated with the DESA; (g) frequency Ω_1/π estimated with the DESA; (h) frequency Ω_2/π estimated with the DESA.

$$\Phi_d(f)(m, n) \triangleq \Psi_{d,1}(f)(m, n) + \Psi_{d,2}(f)(m, n), \quad (46)$$

$$\Phi_d(f)(m, n) = 2f^2(m, n) - f(m-1, n)f(m+1, n) - f(m, n-1)f(m, n+1), \quad (47)$$

where

$$\Psi_{d,1}(f)(m, n) \triangleq f^2(m, n) - f(m-1, n)f(m+1, n) \quad (48)$$

applies the 1D energy operator vertically on all columns of f, whereas $\Psi_{d,2}$ operates on the rows. Expression (47) is identical to the discrete operator developed in Ref. 11 for digital image edge detection. Note that this is only one among many possible approaches to discretizing Φ. Replacing spatial derivatives in Φ with various 2D difference schemes yields a variety of 2D discrete-energy operators.

Applying Φ_d to a 2D sinusoid with constant amplitude and frequencies yields

$$\Phi_d[A\cos(\Omega_1 m + \Omega_2 n + \theta)] = A^2[\sin^2(\Omega_1) + \sin^2(\Omega_2)]. \quad (49)$$

Consider now a discrete AM–FM signal

$$f(m, n) = a(m, n)\cos[\phi(m, n)]. \quad (50)$$

Its vertical instantaneous frequency (in radians per sample)

$$\Omega_1(m, n) \triangleq \frac{\partial \phi}{\partial m} = \Omega_{c,1} + \Omega_{m,1} q_1(m, n) \quad (51)$$

has center frequency $\Omega_{c,1}$ and maximum frequency deviation $\Omega_{m,1} \le |\Omega_{c,1}|$. The frequency-modulating signal $q_1(m, n)$ is assumed to be a mathematical function with a known computable integral. The same assumption is made for the horizontal frequency signal $\Omega_2 = \partial\phi/\partial n$. All discrete-space frequencies are assumed to be in $[-\pi, \pi]$. Further, all frequency signals are assumed to be of constant sign, i.e., either nonnegative or nonpositive for all (m, n). We henceforth assume that a is band limited with bandwidth Ω_a and that both frequency signals are finite weighted sums of sinusoids and band limited with bandwidth Ω_f. (Our results also apply to the case in which the frequency signals are linear ramps; see the 1D case.[10]) Then under the realistic assumptions

$$\Omega_a \ll \min_k |\Omega_{c,k}|, \qquad \Omega_f \ll 1, \qquad \Omega_{m,k} \ll |\Omega_{c,k}|, \quad (52)$$

it follows from relation (46) and with procedures as in the 1D case in Ref. 9 that

$$\Phi_d\{a(m, n)\cos[\phi(m, n)]\} \approx a^2(m, n)\{\sin^2[\Omega_1(m, n)] + \sin^2[\Omega_2(m, n)]\}. \quad (53)$$

Now replacing the partial derivatives of Section 3 with symmetric three-sample differences in each direction yields the 2D signals,

$$g_1(m, n) = [f(m+1, n) - f(m-1, n)]/2, \quad (54)$$

$$g_2(m, n) = [f(m, n+1) - f(m, n-1)]/2, \quad (55)$$

which are also 2D AM–FM signals with amplitude and instantaneous frequencies that do not vary too fast or too much compared with the carriers $\Omega_{c,k}$. Hence (see also Ref. 10 for the 1D case)

$$\Phi_d[g_1] \approx a^2 \sin^2[\Omega_1](\sin^2[\Omega_1] + \sin^2[\Omega_2]), \quad (56)$$

$$\Phi_d[g_2] \approx a^2 \sin^2[\Omega_2](\sin^2[\Omega_1] + \sin^2[\Omega_2]), \quad (57)$$

where a, Ω_1, Ω_2 are spatially varying signals. Combining approximations (53), (56), and (57) yields a *discrete energy-separation algorithm (DESA):*

$$\arcsin\left(\left\{\frac{\Phi_d[f(m+1, n) - f(m-1, n)]}{4\Phi_d[f(m, n)]}\right\}^{1/2}\right) \approx |\Omega_1(m, n)|, \quad (58)$$

$$\arcsin\left(\left\{\frac{\Phi_d[f(m, n+1) - f(m, n-1)]}{4\Phi_d[f(m, n)]}\right\}^{1/2}\right) \approx |\Omega_2(m, n)|, \quad (59)$$

$$\frac{2\Phi_d[f(m, n)]}{\{\Phi_d[f(m+1, n) - f(m-1, n)] + \Phi_d[f(m, n+1) - f(m, n-1)]\}^{1/2}} \approx |a(m, n)|. \quad (60)$$

The DESA can estimate at each location the amplitude envelope and the magnitude of the two instantaneous-frequency signals of a spatial AM–FM signal. Its constraint is that $0 \le |\Omega_1|, |\Omega_2| \le \pi/2$; i.e., it can estimate frequencies up to one fourth of the sampling frequency. The signs of Ω_1 and Ω_2 can be found from the signs of their corresponding carrier frequencies. The DESA in this paper is a 2D extension of the 1D algorithm called DESA-2 in Ref. 10.

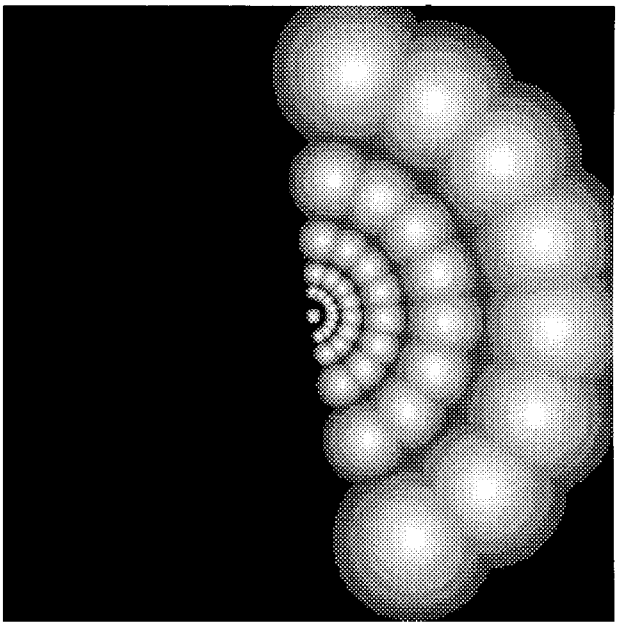

Fig. 2. Frequency responses (represented as intensities) of the 2D Gabor filters used in the filter bank. There are 40 filters arranged in a polar waveletlike tesselation on eight rays, with five filters per ray, plus one filter centered at $(\Omega_1, \Omega_2) = (0, 0)$. Each of the 41 filter responses in the figure has been independently scaled for maximum dynamic range in the available gray levels (from Ref. 1).

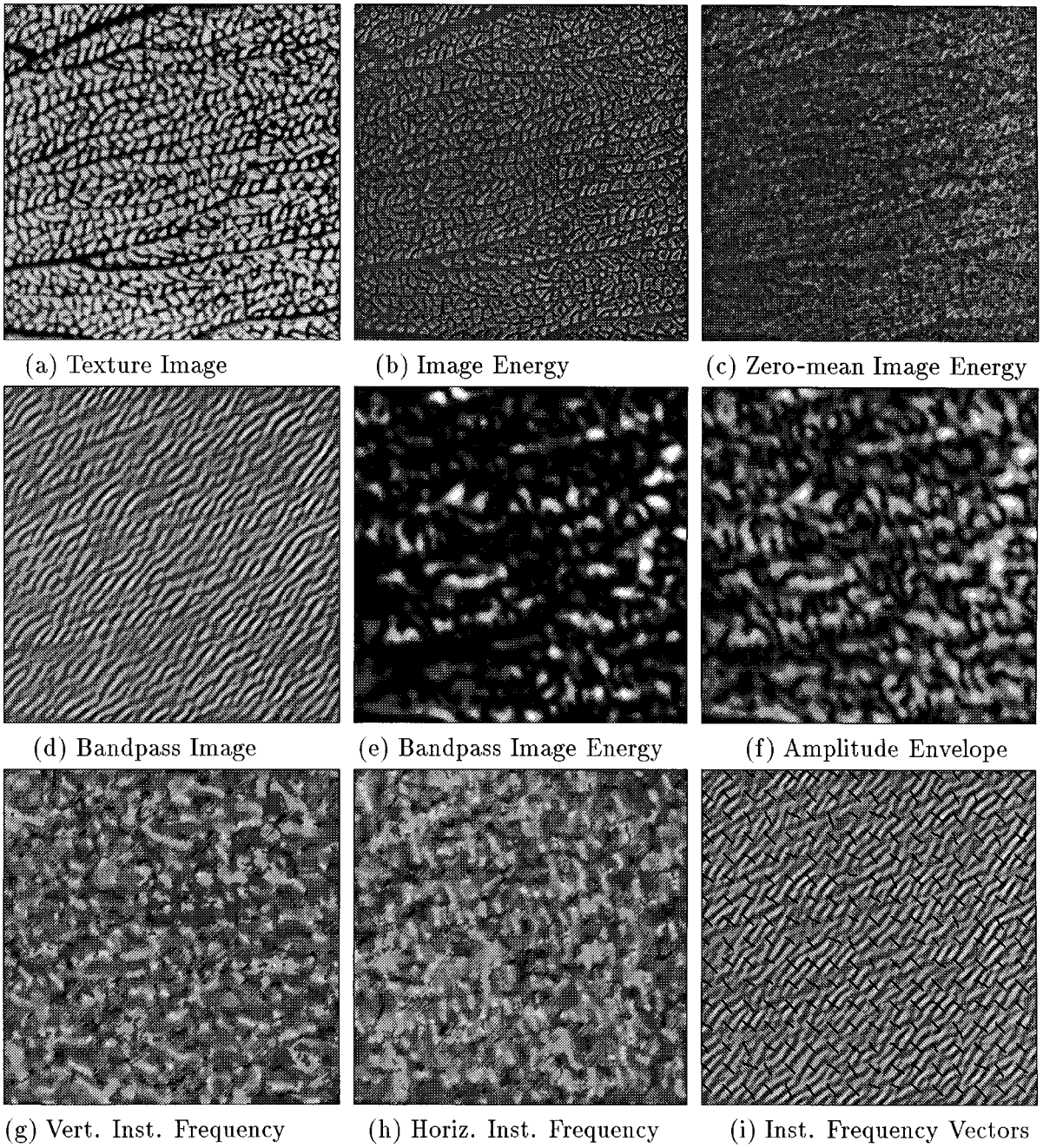

Fig. 3. (a) Intensity image I of a 256×256 pixel texture (sea fan); (b) energy $\Phi(I)$ of the intensity image; (c) energy $\Phi(I - \bar{I})$ of the zero-mean image; (d) bandpass-filtered image $f = I * g$, where g is the impulse response of a Gabor filter with the passband centered at horizontal and vertical frequencies of 27.2 and 27.2 cycles per image, respectively; (e) bandpass image energy $\Phi(f)$; (f) amplitude envelope of f estimated with the DESA; (g) instantaneous frequency Ω_1 of f estimated with the DESA; (h) instantaneous frequency Ω_2 of f estimated with the DESA; (i) frequency vectors (Ω_1, Ω_2), decimated and scaled, superimposed on the bandpass image. {Images in (f)–(h) have been filtered by means of a 3×3 median. All image plots are normalized so that intensities are in $[0, 255]$.}

If the AM–FM signal has constant amplitude A and constant frequencies $\Omega_{c,1}$ and $\Omega_{c,2}$, then the DESA equations provide an exact estimate of the absolute amplitude $|a(m, n)| = |A|$ and the absolute frequencies $|\Omega_1(m, n)| = |\Omega_{c,1}|$ and $|\Omega_2(m, n)| = |\Omega_{c,2}|$.

Figure 1 shows the application of the 2D energy operator and the DESA on a synthetic 2D AM–FM signal. The rms estimation error for the amplitude signal estimation, normalized by the rms of the true amplitude, was 2.5%. The relative rms errors for the estimated frequency signals were 1.5% and 1%. We see that, as the theory predicts, the DESA can estimate the instantaneous amplitude and frequency signals of a spatial AM–FM signal that had relatively large amounts of modulation (50% AM, 20% FM in Ω_1, and 10% FM in Ω_2) with very small estimation error and with very low computational complexity.

5. IMAGE-DEMODULATION EXPERIMENTS

The basic assumption behind the ability of the DESA or the CESA to demodulate 2D AM–FM signals is that

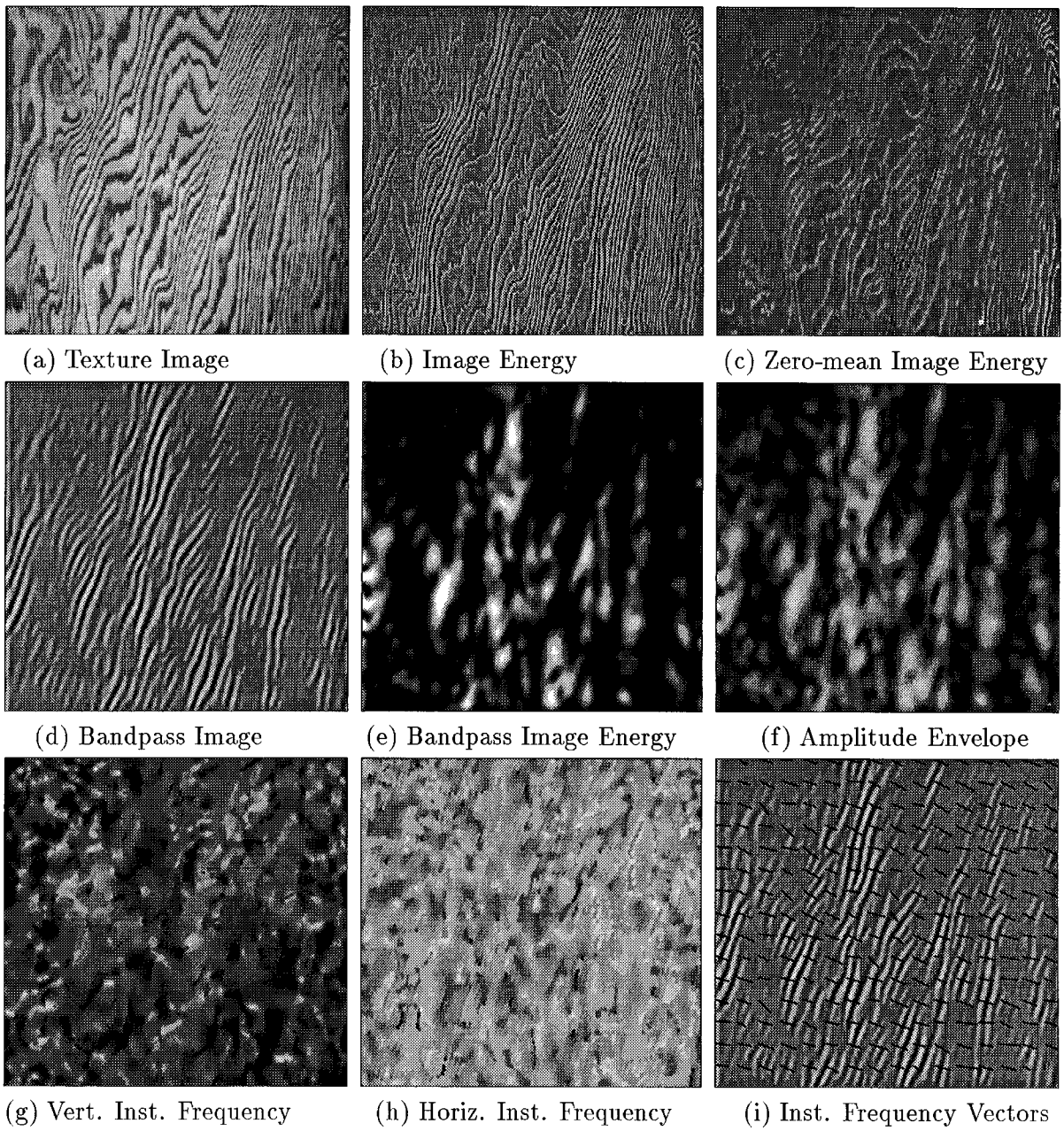

Fig. 4. (a) Intensity image I of a 240×250 pixel texture (wood); (b) energy $\Psi(I)$ of the intensity image; (c) energy $\Psi(I - \bar{I})$ of the zero-mean image; (d) bandpass-filtered image $f = I * g$ where g is the impulse response of a Gabor filter with the passband centered at horizontal and vertical frequencies of 35.5 and 14.7 cycles per image, respectively; (e) bandpass image energy $\Psi(f)$; (f) amplitude envelope of f estimated with the DESA; (g) instantaneous frequency Ω_1 of f estimated with the DESA; (h) instantaneous frequency Ω_2 of f estimated with the DESA; (i) frequency vectors (Ω_1, Ω_2), decimated and scaled, superimposed upon the bandpass image. {Images in (f)–(h) have been filtered by means of a 3×3 median. All image plots are normalized so that intensities are in $[0, 255]$.}

its input signal is narrow band. This then prohibits its direct application to wideband images. A good strategy in the latter case is to bandpass filter the image and apply the DESA to its narrow-band components, assuming these are well modeled by spatial AM–FM signals. This strategy also applies to globally wideband images that are locally narrow band. The bandpass prefiltering also yields two useful by-products: (1) some noise immunity (as explained in Ref. 14) and (2) the signs of the instantaneous-frequency signals from the signs of their corresponding filters' center frequencies (as explained in Subsection 3.B), since the 2D DESA provides absolute frequencies. For bandpass filters we use 2D Gabor filters of the wavelet type, designed as reported in Ref. 1 as a 2D radially symmetric filter bank, whose frequency responses are shown in Fig. 2. Each Gabor filter used in our experiments had a one-octave bandwidth measured radially between the half-peak points. Figures 3 and 4 show the application of the 2D energy operator and the DESA on two digitized texture images after they have been filtered by means of Gabor bandpass filters. From this and other similar real-image experiments we have observed the following:

The energy operator acting on the original nonnegative image I appears to enhance its contrast. However, if the energy operator is applied to the zero-mean image $I - \bar{I}$, where $\bar{I}$ is the global image mean, then the result usually has a much lower contrast and shows some low-pass spatial-energy activity. For example, compare the images in Figs. 3(b) and 4(b) with those in Figs. 3(c) and 4(c), respectively. This happens because

$$\Psi(I) = \Psi(I - \bar{I}) - \bar{I}\nabla^2 I . \tag{61}$$

The result is similar in the discrete case:

$$\Psi_d(I)(m, n) = \Psi_d(I - \bar{I})(m, n) - \bar{I}\left[\sum_{(i,j)\in B} I(m + i, n + j) - 4I(m, n)\right], \tag{62}$$

where $B = \{(1, 0), (-1, 0), (0, 1), (0, -1)\}$. Thus applying Ψ to the image I yields the energy of its zero-mean part minus its Laplacian amplified by the mean. The Laplacian is a high-pass operator, and hence, if the mean is large, the second term dominates and makes $\Psi(I)$ appear mainly as a contrast-enhanced version of I. This also supports the observations in Refs. 11 and 12 about the contrast-enhancing abilities of the 2D discrete operator Ψ_d.

Note that, if the energy operator is applied to a wide-band image as in Figs. 3(a) and 4(a), it may often yield unstable outputs and many negative values. However, when applied to the narrow-band (Gabor filtered) images of Figs. 3(d) and 4(d) it yields fewer negative values. Since the DESA requires nonnegative energies from the narrow-band signal and its derivatives, in our computer simulations we set all negative energy values equal to zero. The estimated energy and amplitude of the narrow-band images seem to convey low-pass information mainly of the contrast type. Of greater interest seem to be the spatial instantaneous-frequency signals, which can be more easily observed when shown as frequency vectors (scaled by some magnification factor to enhance visibility); in Figs. 3(i) and 4(i) we see that the direction of the frequency vectors is most often perpendicular to the local waves in the image. In general, the 2D DESA can yield realistic estimates of the instantaneous amplitude and frequencies in a locally narrow-band image. However, there are also several outlier estimates, which are caused primarily by instantaneous numerical singularities of the DESA, e.g., when it is dividing by a very low energy value. These cause spikes in the estimated amplitude and frequency signals, which can be effectively filtered out by a 2D median filter.

6. CONCLUSION

Locally narrow-band image textures can be modeled by spatial-modulation models of the AM–FM type with slowly varying amplitude and frequency signals. The amplitude signal models intensity contrast variations, whereas the instantaneous-frequency signals convey information about the locally emergent frequencies. We have shown that, if the amplitude and the instantaneous frequency signals do not vary too fast in space or too much in value compared with their mean values, then we can use the 2D energy operator and the 2D DESA to estimate the parameters of these models. Given the importance and applicability of these AM–FM image models, the 2D energy operator and the 2D DESA become important tools for image analysis and computational vision. The advantages of the DESA are simplicity, efficiency, low computational complexity, instantaneous adaptation owing to the differential nature of the energy operators, and ability to track spatial-modulation patterns.

ACKNOWLEDGMENTS

We are grateful to Joe Havlicek of the University of Texas at Austin for producing sets of bandpass-filtered texture images through a computer-simulated 2D Gabor filter bank, for providing the picture for Fig. 2, and for observing the importance of the relative signs of the horizontal and vertical frequencies. P. Maragos's research was supported by the National Science Foundation under grant MIP-9396301. A. Bovik's research was sponsored in part by the U.S. Air Force Office of Scientific Research, Air Force Systems Command, U.S. Air Force, under grant F49620-93-1-0307.

REFERENCES

1. A. C. Bovik, N. Gopal, T. Emmoth, and A. Restrepo, "Localized measurement of emergent image frequencies by Gabor wavelets," IEEE Trans. Inf. Theory **38,** 691–712 (1992).
2. J. P. Havlicek, A. C. Bovik, and P. Maragos, "Modulation models for image processing and wavelet-based image demodulation," presented at the 26th Annual Asilomar Conference on Signals, Systems and Computers, Monterey, California, October 1992.
3. J. E. Daugman, "Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters," J. Opt. Soc. Am. A **2,** 1160–1169 (1985).
4. P. Maragos, A. C. Bovik, and T. F. Quatieri, "A multidimensional energy operator for image processing," in *Visual Communications and Image Processing '92,* P. Maragos, ed., Proc. Soc. Photo-Opt. Instrum. Eng. **1818,** 177–186 (1993).
5. J. F. Kaiser, "On a simple algorithm to calculate the 'energy' of a signal," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Albuquerque, N. Mex., April 1990* (Institute of Electrical and Electronics Engineers, New York, 1990), pp. 381–384.
6. J. F. Kaiser, "On Teager's energy algorithm and its generalization to continuous signals," in *Proceedings of the IEEE Digital Signal Processing Workshop, Mohonk (New Paltz), N. Y., September 1990* (Institute of Electrical and Electronics Engineers, New York, 1990).
7. P. Maragos, J. F. Kaiser, and T. F. Quatieri, "On separating amplitude from frequency modulations using energy operators," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, San Francisco, Calif., March 1992* (Institute of Electrical and Electronics Engineers, New York, 1992), pp. II: 1–4.
8. P. Maragos, T. F. Quatieri, and J. F. Kaiser, "Speech nonlinearities, modulations, and energy operators," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Toronto, Canada, May 1991* (Institute of Electrical and Electronics Engineers, New York, 1991), pp. 421–424.
9. P. Maragos, J. F. Kaiser, and T. F. Quatieri, "On amplitude and frequency demodulation using energy operators," IEEE Trans. Signal Process. **41,** 1532–1550 (1993).
10. P. Maragos, J. F. Kaiser, and T. F. Quatieri, "Energy separation in signal modulations with application to speech analysis," IEEE Trans. Signal Process. **41,** 3024–3051 (1993).

11. T.-H. Yu, S. K. Mitra, and J. F. Kaiser, "A novel nonlinear filter for image enhancement," in *Image Processing Algorithms and Techniques II*, M. R. Civanlar, S. K. Mitra, and R. J. Moorhead, eds., Proc. Soc. Photo-Opt. Instrum. Eng. **1452,** 303–309 (1991).
12. S. K. Mitra, H. Li, I.-S. Lin, and T.-H. Yu, "A new class of nonlinear filters for image enhancement," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Toronto, Canada, May 1991* (Institute of Electrical and Electronics Engineers, New York, 1991), pp. 2525–2528.
13. H. M. Teager and S. M. Teager, "Evidence for nonlinear production mechanisms in the vocal tract," *NATO Advanced Study Institute on Speech Production and Speech Modelling, Bonas, France, July 1989* (Kluwer, Boston, Mass., 1990), pp. 241–261.
14. A. C. Bovik, P. Maragos, and T. F. Quatieri, "AM–FM energy detection and separation in noise using multiband energy operators," IEEE Trans. Signal Process. **41,** 3245–3265 (1993).