Modulation Models for Image Processing and Wavelet-Based Image Demodulation[1]

Joseph P. Havlicek†, Alan C. Bovik†, and Petros Maragos‡

†Laboratory for Vision Systems, University of Texas, Austin, TX 78712-1084
‡Division of Applied Sciences, Harvard University, Cambridge, MA 02138

Abstract

This paper studies AM-FM modulation models for two- and higher-dimensional complex-valued nonstationary signals and images that are locally coherent but globally wideband. Demodulation techniques are studied for demodulating both one-component and multi-component signal. For the case of multi-component signals immersed in noise, the individual components must be isolated by spectrally localized multiband filtering prior to demodulation. An algorithm for demodulating the filtered signal components is developed using a quasi-eigenfunction approximation, and bounds for the approximation error are given. After multiband filtering with Gabor wavelets, the filtered demodulation algorithm is used to estimate the dominant emergent frequencies of multi-component images.

1 Introduction

We explore methods for the accurate and efficient extraction of amplitude modulation (AM) and frequency modulation (FM) information in nonstationary, yet locally coherent single-component n-dimensional complex valued signals and images of the form

$$t(\mathbf{x}) = a(\mathbf{x}) \exp[j\varphi(\mathbf{x})] \tag{1}$$

where $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, $t : \mathbb{R}^n \to \mathbb{C}$, $a : \mathbb{R}^n \to [0, 1]$, and $\varphi : \mathbb{R}^n \to \mathbb{R}$, and we also consider,for the first time, multi-component signals of the form

$$t(\mathbf{x}) = \sum_{k=1}^{K} a_k(\mathbf{x}) \exp[j\varphi_k(\mathbf{x})], \tag{2}$$

which are sums of components of the form (1). These models are inherently ambiguous representations; indeed, given an observed signal $t(\mathbf{x})$ there generally exist an infinite number of function pairs $a(\mathbf{x})$, $\varphi(\mathbf{x})$ for which the model (1) is exact. The models are most useful when the signal (or *all* of the components in the case of the model (2)) is "locally narrowband" in the sense that the amplitude $a(\mathbf{x})$ and instantaneous frequency $\nabla\varphi(\mathbf{x})$ do not vary too rapidly [3]. The model (1) has been successfully used in the analysis of textured images when combined with Gabor wavelet image decompositions [1] - [3] and certain nonlinear energy operators [4]. One-dimensional AM-FM models of the form (1) have also been extensively applied to the analysis of speech signals [5] - [9]. Our objective is to estimate the amplitude envelope(s) $a(\mathbf{x})$ and the instantaneous frequencies $\nabla\varphi(\mathbf{x})$ that may characterize the local image structure.

We use the standard orthonormal basis for $\mathbb{R}^n$, and denote the unit vector in the x_i direction by $\mathbf{e}_i$. We write $f^{(i)}$ for $\frac{\partial}{\partial x_i} f$ and $f^{(i,k)}$ for $\frac{\partial^2}{\partial x_i \partial x_k} f$.

2 Single-Component AM-FM Signal

In this section we assume that we are given a signal $t(\mathbf{x})$ of the form (1). For images, the model (1) is best regarded as the result of applying a point logarithmic operation to the analytic function of a positive real signal, wherein the AM envelope $a(\mathbf{x})$ may be interpreted as the image contrast function [4], hence $0 \leq a(\mathbf{x}) \leq 1$ (the analytic function of the real signal is uniquely determined up to an additive constant).

To solve the amplitude estimation problem, we note that

$$|t(\mathbf{x})| = |a(\mathbf{x}) \exp[j\varphi(\mathbf{x})]| = |a(\mathbf{x})| = a(\mathbf{x}). \tag{3}$$

[1]A.C. Bovik and J.P. Havlicek are supported in part by a grant from Texas Instruments. P. Maragos was supported by the National Science Foundation under Grant NSF MIP-9120624 and by a NSF Presidential Young Investigator Award under grant NSF MIP-86-58150 with matching funds from Xerox.

Observing that

$$\begin{aligned}\nabla t(\mathbf{x}) &= \sum_k t^{(k)}(\mathbf{x})\mathbf{e}_k \\ &= \sum_k \left\{a^{(k)}(\mathbf{x})\exp[j\varphi(\mathbf{x})] + jt(\mathbf{x})\varphi^{(k)}(\mathbf{x})\right\}\mathbf{e}_k \\ &= jt(\mathbf{x})\nabla\varphi(\mathbf{x}) + \exp[j\varphi(\mathbf{x})]\nabla a(\mathbf{x}), \qquad (4)\end{aligned}$$

we have that

$$\begin{aligned}\nabla\varphi(\mathbf{x}) &\approx \frac{\nabla t(\mathbf{x})}{jt(\mathbf{x})} \qquad (5) \\ &= \nabla\varphi(\mathbf{x}) - j\frac{\nabla a(\mathbf{x})}{a(\mathbf{x})}. \qquad (6)\end{aligned}$$

In the absence of noise and other signal components, the squared-magnitude error associated with (5),

$$B(\mathbf{x}) = \frac{|\nabla a(\mathbf{x})|^2}{|a(\mathbf{x})|^2}, \qquad (7)$$

is precisely the *instantaneous bandwidth* (albeit in n dimensions) as defined by Cohen in one dimension [10],[11]. This quantity provides an indication of the degree to which $t(\mathbf{x})$ is locally coherent. If there is no amplitude modulation, then $B(\mathbf{x})$ is everywhere zero and there is a single emergent frequency at each point in the domain. Thus, (5) is exact for monochromatic signals of the form

$$t(\mathbf{x}) = A\exp[j\Omega^T\mathbf{x}], \qquad (8)$$

where A and $\Omega = (\omega_1, \omega_2, \ldots, \omega_n)$ are constant. When amplitude modulation is present, however, multiple local frequencies are in general emergent at each point in the domain; $B(\mathbf{x})$ is related to their concentration about $\nabla\varphi(\mathbf{x})$. But even for signals that are not monochromatic, the error in (5) is restricted to the imaginary component of the frequency estimate, and hence we have the amicable result that

$$\nabla\varphi(\mathbf{x}) = \mathrm{Re}\left[\frac{\nabla t(\mathbf{x})}{jt(\mathbf{x})}\right]. \qquad (9)$$

Fig. 1 shows a reasonably sophisticated example of a single-component image with constant amplitude and Gaussian phase $\varphi(x_1, x_2) = (\pi K\sigma^2/N)\exp\left[-\left(x_1^2 + x_2^2\right)/\sigma^2\right]$, where $K = 10$, $\sigma \approx 190$, and $N = 256$. The local estimates of $\nabla\varphi(\mathbf{x})$ given by (5) are shown in Fig. 2, where needle length is proportional to period (the reciprocal of $|\nabla\varphi(\mathbf{x})|$). Close examination reveals that the estimated instantaneous frequencies agree remarkably well with perception.

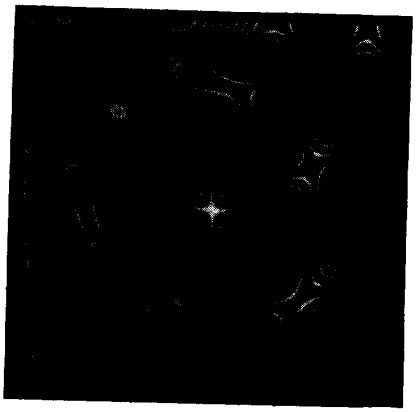


Figure 1: Image with Gaussian phase.

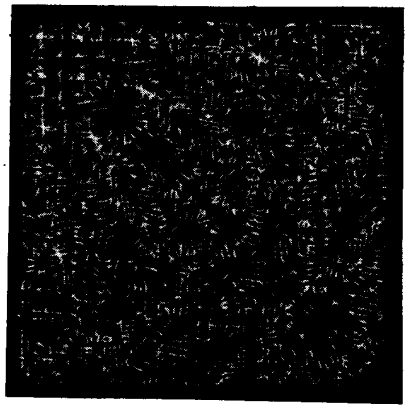


Figure 2: Estimates of $\nabla\varphi(\mathbf{x})$ for the image in Fig. 1.

3 Multi-Component Signals and Noise

The single-component locally narrowband model (1) breaks down in the presence of out-of-band additive noise, when multiple locally narrowband signal components are present, and when both multiple components and broadband noise are present. In each of these cases, the nonlinear demodulation algorithms (3),(9) fail as a result of interference between the signal and the noise, or in the case of multiple components as a result of interference between the signal components themselves.

Therefore, in these situations it becomes necessary to isolate the various locally narrowband components prior to demodulation. Our approach to accomplishing this is to pass the image through multiple bandpass (multiband) prefilters, while ensuring that the filters

are sufficiently narrowband to prevent interference between signal components and that the frequency domain is adequately sampled by the filter set. This scheme offers the advantage of isolating the AM-FM components on a spatially local basis, which is required since the models (1),(2) are nonstationary and may in fact be *globally wideband* yet *locally narrowband.*

3.1 Filtered signal approximations

In developing the demodulation algorithm for the filtered signal components, we use the following important approximation: if a signal component of the form (1) is input to a linear system $g_\sigma(\mathbf{x})$ with n-dimensional Fourier transform $G_\sigma[\Omega]$, then the response $t_\sigma(\mathbf{x}) = t(\mathbf{x}) * g_\sigma(\mathbf{x})$ can be approximated by

$$\widehat{t}_\sigma(\mathbf{x}) = t(\mathbf{x}) G_\sigma[\nabla\varphi(\mathbf{x})]. \tag{10}$$

The approximation (10) is a quasi-extension of the eigenfunction concept of linear system theory, and is exact when $t(\mathbf{x})$ is monochromatic. Theorem 1, which follows, provides a tight bound on the approximation error under very reasonable constraints on $t(\mathbf{x})$. The bound is expressed in terms of the p-energy variances $\bar{\Delta}_i^{(p)}(g_\sigma)$, $\bar{\Delta}_{i,j}^{(p)}(g_\sigma)$ of the filter $g_\sigma(\mathbf{x})$ [3] defined by

$$\bar{\Delta}_i^{(p)}(f) = \left[\sum_i \left|\Delta_i^{(p)}(f)\right|^p\right]^{\frac{1}{p}} = \left\|\{\Delta_i^{(p)}(f)\}\right\|_{\ell^p} \tag{11}$$

and

$$\bar{\Delta}_{i,j}^{(p)}(f) = \left[\sum_{i,j} \left|\Delta_{i,j}^{(p)}(f)\right|^p\right]^{\frac{1}{p}} = \left\|\{\Delta_{i,j}^{(p)}(f)\}\right\|_{\ell^p}, \tag{12}$$

where

$$\Delta_i^{(p)}(f) = \left[\int_{\mathbb{R}^n} |x_i|^p |f(\mathbf{x})|^p \, d\mathbf{x}\right]^{\frac{1}{p}} \tag{13}$$

and

$$\Delta_{i,j}^{(p)}(f) = \left[\int_{\mathbb{R}^n} |x_i x_j|^p |f(\mathbf{x})|^p \, d\mathbf{x}\right]^{\frac{1}{p}}, \tag{14}$$

as well as the Sobolev p-norms [3] $\bar{\delta}_i^{(p)}(a)$ and $\bar{\delta}_{i,j}^{(p)}(\varphi)$ of the signal amplitude and phase defined by

$$\bar{\delta}_i^{(p)}(f) = \left[\sum_i \left|\delta_i^{(p)}(f)\right|^p\right]^{\frac{1}{p}} = \left\|\{\delta_i^{(p)}(f)\}\right\|_{\ell^p} \tag{15}$$

and

$$\bar{\delta}_{i,j}^{(p)}(f) = \left[\sum_{i,j} \left|\delta_{i,j}^{(p)}(f)\right|^p\right]^{\frac{1}{p}} = \left\|\{\delta_{i,j}^{(p)}(f)\}\right\|_{\ell^p}, \tag{16}$$

where

$$\delta_i^{(p)}(f) = \left[\int_{\mathbb{R}^n} \left|f^{(i)}(\mathbf{x})\right|^p d\mathbf{x}\right]^{\frac{1}{p}} \tag{17}$$

and

$$\delta_{i,j}^{(p)}(f) = \left[\int_{\mathbb{R}^n} \left|f^{(i,j)}(\mathbf{x})\right|^p d\mathbf{x}\right]^{\frac{1}{p}}. \tag{18}$$

The statement of Theorem 1 follows, where reals $p, q > 1$ are *conjugate exponents* if $p + q = pq$.

Theorem 1: Let $g_\sigma : \mathbb{R}^n \to \mathbb{C}$, $a : \mathbb{R}^n \to \mathbb{R}$, and $\varphi : \mathbb{R}^n \to \mathbb{R}$ be such that a is once-differentiable and φ is twice-differentiable. Then for conjugate exponent pairs (p, q), and (p', q') with $q, q' > n$ and $1 < p, p' < \frac{n}{n-1}$ for which $|x_i g_\sigma(\mathbf{x})|^p$, $|a^{(i)}(\mathbf{x})|^q$, $|x_i x_j g_\sigma(\mathbf{x})|^{p'}$, and $|\varphi^{(i,j)}(\mathbf{x})|^{q'}$ are integrable on $\mathbb{R}^n$ for $1 \le i, j \le n$,

$$\begin{aligned} \varepsilon_t(\mathbf{x}) &= \left|t_\sigma(\mathbf{x}) - \hat{t}_\sigma(\mathbf{x})\right| \\ &\le \left(\frac{q'}{q'-n}\right) \bar{\Delta}_i^{(p')}(g_\sigma) \bar{\delta}_i^{(q')}(a) \\ &+ \frac{q^2}{(q-n)(2q-n)} \bar{\Delta}_{i,j}^{(p)}(g_\sigma) \bar{\delta}_{i,j}^{(q)}(\varphi) \end{aligned} \tag{22}$$

■

The proof of the theorem proceeds via techniques similar to those developed in [3] and is too involved for presentation here.

Corollary: For conjugate exponent pairs $(\widetilde{p},\widetilde{q})$ and $(\widehat{p},\widehat{q})$ such that $\left|a^{(i)}(\mathbf{x})\right|^{\widetilde{pq}}$, $\left|\varphi^{(i)}(\mathbf{x})\right|^{\widetilde{qq}}$, $|x_i g_\sigma(\mathbf{x})|^{\widehat{p}}$, and $\left|a^{(i,j)}(\mathbf{x})\right|^{\widehat{q}}$ are integrable on $\mathbb{R}^n$ for $1 \le i, j \le n$, Theorem 1 also bounds the error in the approximation

$$\begin{aligned} \widehat{t}_\sigma^{(i)}(\mathbf{x}) &= a^{(i)}(\mathbf{x}) \exp[j\varphi(\mathbf{x})] G_\sigma[\nabla\varphi(\mathbf{x})] \\ &+ j a(\mathbf{x}) \varphi^{(i)}(\mathbf{x}) \exp[j\varphi(\mathbf{x})] G_\sigma[\nabla\varphi(\mathbf{x})]. \end{aligned} \tag{23}$$

■

The corollary is proved by applying Minkowski's inequality followed by Hölder's integral inequality [12], and bounds the error in the quasi-eigenfunction approximation $t_\sigma^{(i)}(\mathbf{x})$ for the partial derivative of the filtered signal. Since the partials of $\widehat{t}_\sigma(\mathbf{x})$ are the components of $\nabla\widehat{t}_\sigma(\mathbf{x})$, upon substituting (10) and (23) into (5) the filtered frequency demodulation algorithm becomes

$$\nabla\varphi(\mathbf{x}) \approx \mathrm{Re}\left[\frac{\nabla\widehat{t}_\sigma(\mathbf{x})}{j\widehat{t}_\sigma(\mathbf{x})}\right], \tag{24}$$

which is useful only if the error bound (22) can be made small. Given the frequency estimate in (24), the filtered amplitude estimate is obtained by post-normalizing (10) by $G_\sigma[\nabla\varphi(\mathbf{x})]$ and applying (3).

3.2 Gabor wavelets

For the multiband demodulation algorithm to be applicable in a wide variety of situations, we wish to minimize the approximation error (22) without placing unnecessary constraints on the signal $t(\mathbf{x})$. Both terms in the error can be minimized by choosing filters $g_\sigma(\mathbf{x})$ with minimal p-energy variances. As we observed above, however, these filters must also be spectrally localized so that they can effectively isolate the multiple locally narrowband signal components that may be present. Further stochastic arguments have been presented elsewhere [3],[9] which suggest that the filters $g_\sigma(\mathbf{x})$ should also have small Sobolev p-norms. Simultaneously optimizing these criteria inevitably leads to minimum uncertainty filters of the form

$$g_\sigma(\mathbf{x}) = [2\pi\sigma^2]^{-\frac{n}{4}} \exp[-\mathbf{x}^T\mathbf{x}/(4\sigma^2)], \qquad (25)$$

which are unit-energy Gaussians. Frequency shifting and scaling these filters to adequately sample the frequency plane results in channel filters that are Gabor wavelets. Fig. 3 shows the frequency responses of the set of 40 spatio-spectrally localized octave-bandwidth Gabor wavelets we used for multiband filtering in the examples of the next section. Fig. 4 shows a (single-component) radial chirp image. The estimates of $\nabla\varphi(x)$ obtained using (24) after processing with the wavelet filter set are shown in Fig. 4(b).

4 Multi-Component Examples

The problem of isolating and estimating the the individual amplitude envelopes $a_k(\mathbf{x})$ and instantaneous frequencies $\nabla\varphi_k(\mathbf{x})$ for a signal of the form (2) is in general a formidable one, since the components may merge, vanish, or develop sudden transitions in frequency or contrast. Furthermore, *a priori* information concerning the number of components that are present in a given signal is not usually available. Nevertheless, one can always estimate the dominant emergent frequency component on a local basis by applying the filtered demodulation algorithm (24) at each point in the domain to the output $t_*(\mathbf{x})$ of the channel filter whose magnitude response is largest at that point. If the number of components is known to be K, then all

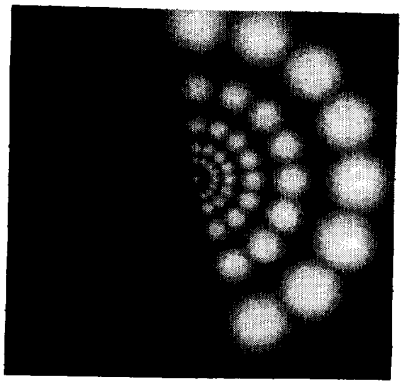


Figure 3: Frequency responses of multiband Gabor wavelets

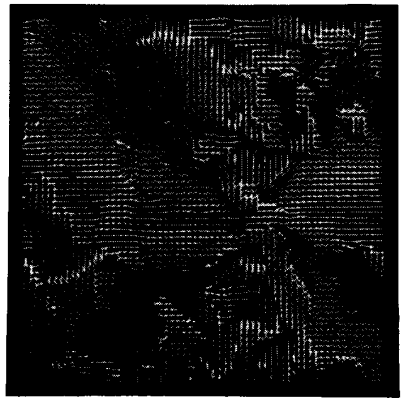
K amplitudes and instantaneous frequencies can be estimated at every point in the domain by applying the filtered demodulation algorithm to the K outputs of the multiband channel filters with the K strongest magnitude responses on a point-by-point basis.

Fig. 5(a) and Fig. 6(a) show images of a tree and of burlap, respectively, both of which seem well suited for treatment with the multi-component model (2). In the case of the burlap image, there are clearly at least one dominant vertical and one dominant horizontal component existing throughout the image. Fig. 5(b) is the dominant emergent frequency for the tree image estimated at each point from the channel filter with the largest magnitude response, where needle length is proportional to the period. Estimates for the dominant emergent frequency for the burlap image are shown in Fig. 6(b). Like the single-component instantaneous frequency estimates in Fig. 2, the estimates in Fig. 5(b) and Fig. 6(b) agree remarkably well with perception. Discontinuous boundaries between regions where different signal components are dominant are also clearly visible in the latter two figures. Finally, frequency estimates based on the outputs of the channel filters with second largest magnitude response at each point are shown for the Tree and Burlap images in Fig. 5(c) and Fig. 6(c), respectively.

5 Future Work

Future work will include treatment of the multidimensional modulation algorithms in the digital domain and further development of techniques for esti-

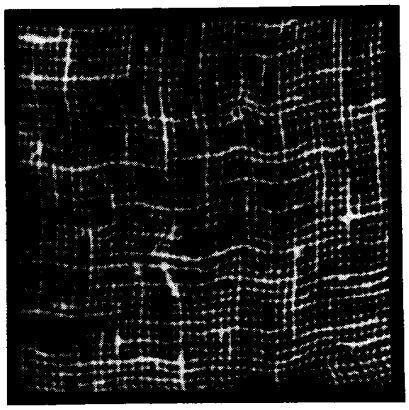
(*a*)



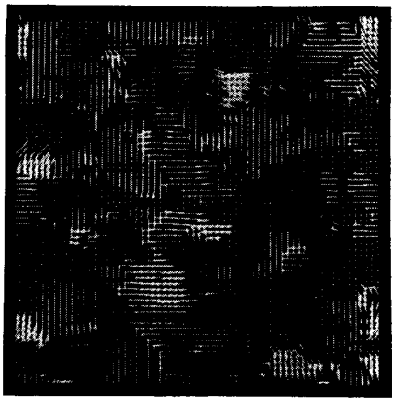
(*b*)

(*c*)

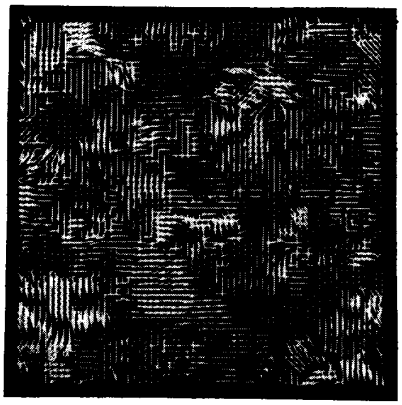
Figure 5: Tree. (*a*) Image. (*b*) $\nabla\varphi(\mathbf{x})$ computed from filter with largest magnitude response. (*c*) $\nabla\varphi(\mathbf{x})$ computed from filter with 2nd largest magnitude response.



(*a*)

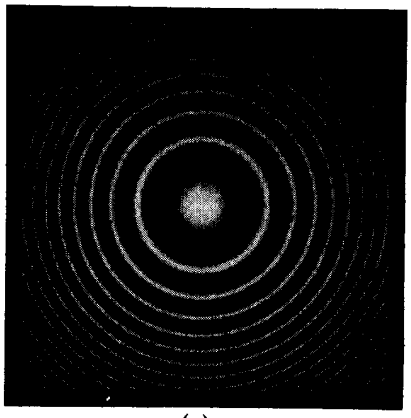


(*b*)

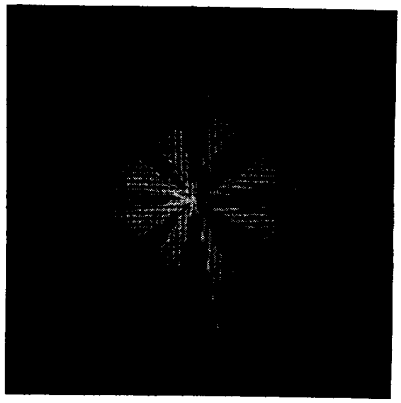


(*c*)

Figure 6: Burlap. (*a*) Image. (*b*) $\nabla\varphi(\mathbf{x})$ computed from filter with largest magnitude response. (*c*) $\nabla\varphi(\mathbf{x})$ computed from filter with 2nd largest magnitude response.



(*a*)



(*b*)

Figure 4: Radial chirp. (*a*) Image. (*b*) $\nabla\varphi(\mathbf{x})$ computed from filter with largest magnitude response.

mating the amplitudes and instantaneous frequencies of the individual components of multi-component signals and images. We believe that the model (2) will find widespread application.

Acknowledgement

The authors wish to express their gratitude to Boaz Super who implemented the multiband Gabor wavelet filters.

References

[1] A.C. Bovik, M. Clark and W.S. Geisler, "Multichannel texture analysis using localized spatial filters," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-12, pp. 55-73, Jan. 1990.

[2] A.C. Bovik "Analysis of multichannel narrowband filters for image texture segmentation," *IEEE Trans. Signal Process.*, vol. SP-39, pp. 2025-2043, Sept. 1991.

[3] A.C. Bovik, N. Gopal, T. Emmoth and A. Restrepo, "Localized measurement of emergent image frequencies by Gabor wavelets," *IEEE Trans. Info. Theory*, Special Issue on Wavelet Transforms and Multiresolution Signal Analysis, vol. IT-38, pp. 691-712, Mar. 1992.

[4] P. Maragos, A.C. Bovik and T.F. Quatieri , "A multidimensional energy operator for image processing," in *Proc. SPIE Symp. Visual Commun. Image Process.*, Boston, MA, Nov. 1992.

[5] P. Maragos, T.F. Quatieri and J.F. Kaiser, "Speech nonlinearities, modulations, and energy operators," in *Proc. IEEE Int'l. Conf. Acoust., Speech, Signal Process.*, Toronto, Ontario, Canada, May 1991.

[6] P. Maragos, T.F. Quatieri and J.F. Kaiser, "On separating amplitude from frequency modulations using energy operators," in *Proc. IEEE Int'l. Conf. Acoust., Speech, Signal Process.*, San Francisco, CA, May 1992.

[7] P. Maragos, T.F. Quatieri and J.F. Kaiser, "On amplitude and frequency demodulation using energy operators," *IEEE Trans. Acoust., Speech, Signal Process.*, to appear, 1993.

[8] A.C. Bovik, P. Maragos and T.F. Quatieri, "Measuring amplitude and frequency modulations in noise using multiband energy operators," *IEEE Int'l. Symp. Time-Freq. and Time-Scale Anal.*, Victoria, British Columbia, Canada, Oct. 1992.

[9] A.C. Bovik, P. Maragos and T.F. Quatieri, "AM-FM energy detection and separation in noise using multiband energy operators," *IEEE Trans. Signal Proc.*, Special Issue on Wavelets and Signal Processing, submitted.

[10] L. Cohen and C. Lee, "Standard deviation of instantaneous frequency," in *Proc. IEEE Int'l. Conf. Acoust., Speech, Signal Process.*, Glasgow, Scotland, May 23-26, 1989.

[11] L. Cohen and C. Lee, "Instantaneous bandwidth for signals and spectrogram," in *Proc. IEEE Int'l. Conf. Acoust., Speech, Signal Process.*, Albuquerque, NM, Apr. 1990.

[12] W. Rudin, *Real and complex analysis.* New York: McGraw-Hill, 1987.