

4.4

AM-FM Image Models: Fundamental Techniques and Emerging Trends

Joseph P. Havlicek
University of Oklahoma

Peter C. Tay
University of Virginia

Alan C. Bovik
*The University of
Texas at Austin*

1	Introduction.....	377
2	Fundamentals of AM-FM Image Modeling	380
2.1	Two-Dimensional Energy Separation • 2.2 Two-Dimensional Analytic Image	
3	Practical Techniques for AM-FM Image Modeling	382
3.1	Filtered Demodulation • 3.2 Filterbank Design Considerations •	
3.3	Dominant Component Analysis • 3.4 Channelized Components Analysis •	
3.5	Reconstruction • 3.6 Postfiltering and Postprocessing	
4	Emerging Trends in AM-FM Image Modeling.....	390
4.1	Perfect Reconstruction Filterbanks • 4.2 Perfect Reconstruction FM Algorithm	
5	Conclusion	393
	Acknowledgment.....	393
	References	393

1 Introduction

AM-FM models describe an image in terms of highly localized two-dimensional (2D) frequency modulation (FM) and amplitude modulation (AM). These models are remarkably powerful for characterizing and processing textured images. This is because of their inherent ability to quantify the most important and basic elements of texture: The FM part of the model comprises instantaneous frequency vectors that capture the local texture orientation and granularity (i.e., the local directionality and the local coarseness or pattern spacing), whereas the AM part captures the local texture contrast (i.e., the disparity in intensity between the brightest and darkest elements of the local texture pattern). We begin with some illustrative examples which we discuss in relation to the discrete Fourier transform (DFT) described in Chapter 2.3.

Consider first the discrete-time one-dimensional (1D) chirp signal

$$f(k) = \cos\left(\frac{0.4\pi}{512}k^2\right) \quad (1)$$

defined for $0 \leq k < 512$ and graphed in Fig. 1A. Intuitively, this signal is a discrete sinusoid that oscillates faster and faster as the time index k advances (it may be thought of as $\cos(\omega k)$, where $\omega = \frac{0.4\pi}{512}k$). With the DFT, the signal can be represented as a sum of 512 discrete-time complex sinusoidal Fourier components. We refer to these components as *stationary* because they each have an amplitude and a frequency that are fixed for all k . However, the signal (1) is decidedly *nonstationary* in the sense that the local rate of oscillation (or *instantaneous frequency*) varies with k . In the DFT model of the signal, this nonstationary structure is created by subtle and sometimes complicated constructive and destructive interference between the stationary Fourier components. The DFT magnitude spectrum of the signal appears in Fig. 1B. Although an expert well-versed in chirp analysis can immediately identify the characteristic spectrum shape in Fig. 1B, we are going to argue in this chapter that the DFT is not always the most natural way to interpret the spatial and spectral structure of a nonstationary signal.

Suppose that we relax the constraints of the DFT model and permit instead a representation in terms of quasisinusoidal components having amplitudes and frequencies that

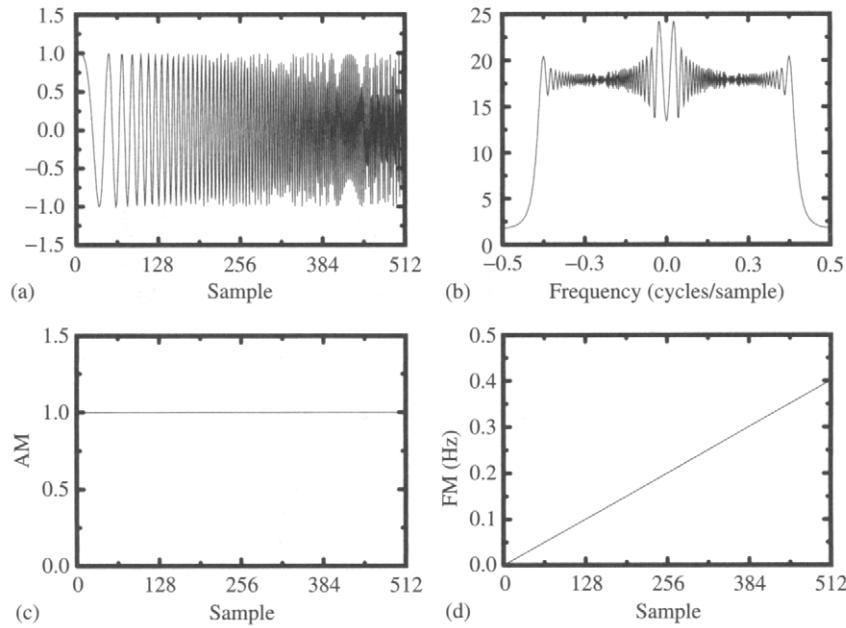


FIGURE 1 One-dimensional chirp example. **A:** Signal. **B:** Discrete Fourier transform magnitude. **C:** Computed amplitude modulation function. **D:** Computed frequency modulation function.

are allowed to vary arbitrarily but smoothly throughout the domain. Components of this type are called *AM-FM components* or *AM-FM functions*; they are like sinusoids with “locally adjustable” amplitudes and frequencies. The nonstationary amplitude envelope of such a component is called the *AM function*, while the derivative of the instantaneous phase is called the *FM function*.¹ In fact, the signal (1) may be precisely represented using only a single AM-FM component. The AM and FM functions of this component, which were computed using the demodulation algorithm given in [1], are shown in Fig. 1C and 1D. They could alternatively have been obtained using the discrete Teager-Kaiser Energy Operator (TKEO) and energy separation algorithm (ESA) described in [2]. The computed modulations indicate that the signal is a single nonstationary sinusoid with a constant amplitude and an instantaneous frequency that increases linearly. This AM-FM model characterizes the signal in a way that is at once both parsimonious and appealingly consistent with our intuition about how our senses (hearing in this case) would perceive the signal: When played through an audio speaker, the sound of the signal has a constant volume (AM) and a pitch that increases linearly with time (FM). Moreover, the computed AM-FM model naturally facilitates analysis and processing of the signal directly in terms of the nonstationary modulated structure. We refer to this as *modulation domain* signal processing. In 1D, AM-FM sig-

nal models have been used successfully for time-frequency analysis, speech analysis, and in the study of nonlinear air flow during human phonation [2–9].

With the 2D DFT, an image is represented as a sum of stationary sinusoidal Fourier components like the ones shown in Fig. 1 of Chapter 2.3. Each of these discrete-space sinusoids has an amplitude and a 2D frequency vector that are constant throughout the entire image domain. As in the 1D case, the 2D Fourier representation can only create nonstationary image structure through constructive and destructive interference between stationary Fourier components. The main idea behind AM-FM image modeling is to relax the constraints of the DFT and formulate a representation that captures the essential structure of an image using a relatively small number of quasi-sinusoidal 2D AM-FM components. These components are like discrete-space sinusoids with amplitudes and frequency vectors that are permitted to vary arbitrarily but smoothly throughout the image domain. Compared to the pure sinusoids in Fig. 1 of Chapter 2.3, the AM-FM components typically exhibit ridges and troughs that bend in orientation, bifurcate, and may be bright in some regions while dim in others.

Two basic examples are presented in Fig. 2. The wood grain image of Fig. 2A would be interpreted as a single homogeneous texture in many applications. Using the techniques described in Section 3, we computed a 43-component AM-FM model for this image. Adding these 43 components together, we obtained the approximate AM-FM image reconstruction shown in Fig. 2B. Whereas the DFT represents this 256×256 image as a sum of 256^2 sinusoidal Fourier components, it is clear from the figure that the essential structure has

¹For a discrete signal, we consider that the (discrete) instantaneous phase function contains samples of a corresponding continuous-domain phase function that may be differentiated. The discrete FM function contains the samples of this derivative.

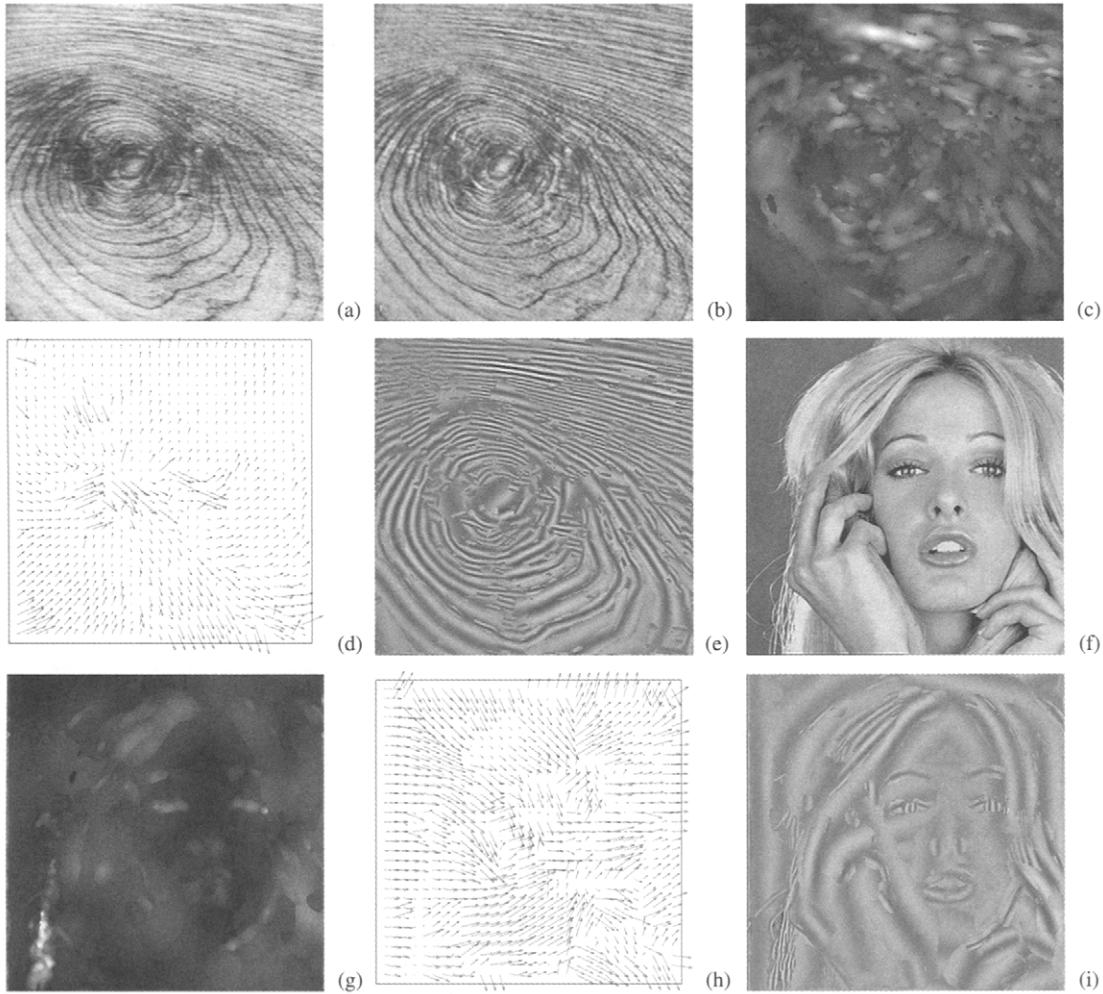


FIGURE 2 Basic two-dimensional AM-FM [amplitude modulation (AM); frequency modulation (FM)] modeling examples. **A:** Woodgrain image. **B:** Approximate reconstruction obtained by summing 43 computed AM-FM image components. **C:** Dominant AM function. **D:** Dominant FM function. **E:** Reconstruction of the dominant AM-FM component from (C) and (D). **F:** *Tiffany* image. **G:** Dominant AM function. **H:** Dominant FM function. **I:** Reconstruction of the dominant AM-FM component from (G) and (H).

been captured using only 43 AM-FM components. Next, we selected at each pixel one of the 43 AM-FM components as the dominant one (the algorithm for making this selection is explained in Section 3.3). By extracting the value of the AM function from the dominant component on a pixel-by-pixel basis, we obtained the dominant AM image function given in Fig. 2C. Generally, the dominant AM function is bright where there is high contrast in the local texture pattern and dim in regions where there is less contrast.²

²There is a technical issue that deserves some clarification at this point. We have stated that the amplitudes and frequencies of the AM-FM components should vary smoothly, yet the dominant AM function in Fig. 2C clearly is not smooth everywhere. This arises because, even though the AM function (and FM function) of each constituent AM-FM component might be smooth everywhere, the dominant AM function (and FM function) may still exhibit discontinuities along the boundaries between regions where different components are dominant. Also, in practice the componentwise AM and FM functions computed from a nontrivial image will generally exhibit some discontinuities.

We also extracted the FM frequency vector from the dominant component on a pixel-by-pixel basis to obtain the dominant FM image function shown in Fig. 2D. We refer to this figure as a *needle diagram*. Each directed needle in the diagram originates from a pixel and has an arrowhead at the terminating end to indicate the dominant texture orientation at the pixel. For clarity, only one needle is plotted for each block of 8×8 pixels. The needle lengths are inversely proportional to the magnitude of the FM frequency vector, so that short needles indicate high frequencies corresponding to image structures of small spatial extent, whereas long needles indicate low frequencies corresponding to larger structures. By combining the dominant AM and FM functions, we synthesized the dominant AM-FM image component. The result, which is a single AM-FM function depicting the dominant structure of the image, is shown in Fig. 2E. Notice that the needles in Fig. 2D are orthogonal to the salient ridges and troughs in Fig. 2A and Fig. 2E. With relation

to the stationary sinusoids in Fig. 1 of Chapter 2.3, the ridges and troughs in Fig. 2A, B, and E may be interpreted as nonstationary AM-FM wavefronts with directions of propagation given by the orientations of the needles in Fig. 2D.

A similar example is given in Fig. 2F–I. The reconstruction in Fig. 2I represents the dominant structure of the image using only a single AM-FM component. On local neighborhoods, this dominant component can be thought of as a “bendable” sinusoid with variable contrast. Note that the subject’s hair in Fig. 2F exhibits structure on at least three different spatial scales (or granularities). It is interesting to observe the various regions in Fig. 2I where the AM-FM model interprets the fine-, medium-, or coarse-scale structure as being dominant.

While the AM-FM dominant component reconstructions in Fig. 2E and Fig. 2I are interesting and informative for human consumption, the true power of the AM-FM approach for texture processing is manifest in the computed modulations of Fig. 2C, D, G, and H. Modulations of this type have been used successfully for jointly localized analysis [10–13], texture segmentation and classification [14–18], edge detection and image enhancement [19], estimating three-dimensional (3D) shape from texture [14, 20, 21], performing texture-based computational stereopsis [22], fingerprint analysis and classification [23, 24], content-based image retrieval [25], and for regenerating occluded and damaged textures [26]. In Section 2 of this chapter we explain the fundamental concepts of AM-FM image modeling. Practical techniques for computing AM-FM models like the ones shown in Fig. 2 are presented in Section 3. We summarize several important recent developments and emerging trends in AM-FM image modeling in Section 4. Finally, in Section 5 we conclude and briefly examine the significance of emerging trends in AM-FM modeling in terms of future research directions.

2 Fundamentals of AM-FM Image Modeling

Suppose we are given a real-valued 2D continuous-domain image $f(x, y)$. We assume that the mean has been subtracted out, so that the average value of $f(x, y)$ is zero. What we want is to find a vector-valued FM function $\nabla\varphi(x, y) = [U(x, y) \ V(x, y)]^T$, where

$$U(x, y) = \frac{\partial}{\partial x} \varphi(x, y), \quad (2)$$

$$V(x, y) = \frac{\partial}{\partial y} \varphi(x, y), \quad (3)$$

and a positive semidefinite AM function $a(x, y)$ so that

$$f(x, y) = a(x, y) \cos[\varphi(x, y)]. \quad (4)$$

This is an ill-posed problem because, for any given $f(x, y)$, there exist infinitely many pairs of functions $a(x, y), \nabla\varphi(x, y)$ that give equality in (4). The question of how one should select a particular pair of AM and FM functions to associate with the image $f(x, y)$ has been vigorously debated for decades [27–32]. Here, we will consider two main types of approaches to the problem.

Methods of the first type operate directly on the image values $f(x, y)$ to obtain solutions for $a(x, y)$ and $\nabla\varphi(x, y)$ (or, equivalently, for $a(x, y)$ and $\varphi(x, y)$). The 2D TKEO and its associated ESA [10], which are discussed in Section 2.1, are an example of a method of the first type. Methods of the second type start by adding an imaginary part to the image to generate a complex-valued extension $z(x, y)$ with $\text{Re}[z(x, y)] = f(x, y)$. As we shall see, a unique pair of AM and FM functions can then be computed from $z(x, y)$. The analytic image approach discussed in Section 2.2 is one example of a method of the second type. It is important to realize that the two types of approaches are actually equivalent. When a method of the first type computes a particular $a(x, y)$ and $\nabla\varphi(x, y)$, it implicitly specifies a complex image $z(x, y) = a(x, y) \exp[j\varphi(x, y)]$. Similarly, when a method of the second type computes a complex image $z(x, y)$, the complex image can always be written as $z(x, y) = a(x, y) \exp[j\varphi(x, y)]$. This implicitly specifies a particular AM function $a(x, y)$ and FM function $\nabla\varphi(x, y)$ to be associated with the original real-valued image $f(x, y)$. Thus, we see that selecting a particular pair of modulating functions is precisely equivalent to adding an imaginary part to $f(x, y)$.

It may seem strange that we would want to consider two different types of approaches that generally yield different answers to the same problem. The explanation hinges on two key points. First, the two different methods frequently yield solutions for the modulating functions that are quite similar. Second, it must be kept in mind that associating amplitude and frequency modulating functions with a given $f(x, y)$ is a problem that does not admit a unique solution. While the solutions obtained from the two types of methods are generally different, it is often the case that both achieve our goals of formulating a consistent model and of delivering AM and FM functions that are useful for solving real-world engineering problems.

2.1 Two-Dimensional Energy Separation

To understand the fundamentals of energy separation, consider a 1D continuous-time signal $f(t)$ that we want to model as $f(t) = a(t) \cos[\varphi(t)]$. The 1D continuous TKEO is given by $\Psi_c[f(t)] = \{f'(t)\}^2 - f''(t)f(t)$ [2]. If $f(t)$ is reasonably locally narrowband, then, for a particular pair of modulating functions $a(t)$ and $\varphi'(t)$ that are generally intuitively appealing and physically meaningful, we have that $\Psi_c[f(t)] \approx \{a(t)\varphi'(t)\}^2$. The quantity $\{a(t)\varphi'(t)\}^2$ is known as

the *Teager energy* of the signal. Moreover, applying the TKEO to approximate the Teager energy of the derivative of the signal yields $\Psi_c[f'(t)] \approx a^2(t)\{\varphi'(t)\}^4$. Thus, individual solutions for the AM and FM functions of $f(t)$ can be approximated by applying the ESA [2]

$$a(t) \approx \frac{\Psi_c[f(t)]}{\sqrt{\Psi_c[f'(t)]}}, \quad (5)$$

$$|\varphi'(t)| \approx \sqrt{\frac{\Psi_c[f'(t)]}{\Psi_c[f(t)]}}. \quad (6)$$

Note that the ESA estimates the FM function from the Teager energy, which involves the square of the instantaneous frequency $\varphi'(t)$. As a result, the *sign* of $\varphi'(t)$ is lost in (6). This presents no real difficulty since, once the AM function $a(t)$ has been estimated using (5), the unwrapped phase can be tracked according to $\varphi(t) = \arccos[f(t)/a(t)]$. The sign of $\varphi(t)$ should be set positive where the unwrapped phase is locally increasing and negative where the unwrapped phase is locally decreasing. For a 1D discrete-time signal $f(k)$, the 1D discrete TKEO is given by [2]

$$\Psi_d[f(k)] = f^2(k) - f(k+1)f(k-1). \quad (7)$$

For a locally narrowband signal $f(k)$, the discrete TKEO closely approximates the discrete Teager energy $\{a(k)\varphi'(k)\}^2$.

By applying (7) independently to the rows and columns of a discrete 2D image $f(n_1, n_2)$ and summing the results, one obtains the discrete 2D TKEO [10]

$$\begin{aligned} \Phi_d[f(n_1, n_2)] &= 2f^2(n_1, n_2) - f(n_1 + 1, n_2)f(n_1 - 1, n_2) \\ &\quad - f(n_1, n_2 + 1)f(n_1, n_2 - 1). \end{aligned} \quad (8)$$

The operator Φ_d implicitly selects a particular pair of modulating functions to associate with $f(n_1, n_2)$. For this pair, it approximately tracks the 2D discrete Teager energy $\Phi_d[f(n_1, n_2)] \approx a^2(n_1, n_2)|\nabla\varphi(n_1, n_2)|^2$. For images that are reasonably locally smooth, the modulating functions selected by the 2D TKEO are generally consistent with intuitive expectations [10]. With the TKEO, the magnitudes of the individual amplitude and frequency modulations can be estimated using the 2D discrete ESA [10]

$$|U(n_1, n_2)| \approx \arcsin \sqrt{\frac{\Phi_d[f(n_1 + 1, n_2) - f(n_1 - 1, n_2)]}{4\Phi_d[f(n_1, n_2)]}}, \quad (9)$$

$$|V(n_1, n_2)| \approx \arcsin \sqrt{\frac{\Phi_d[f(n_1, n_2 + 1) - f(n_1, n_2 - 1)]}{4\Phi_d[f(n_1, n_2)]}}, \quad (10)$$

$$a(n_1, n_2) \approx \sqrt{\frac{\Phi_d[f(n_1, n_2)]}{\sin^2(|U(n_1, n_2)|) + \sin^2(|V(n_1, n_2)|)}}. \quad (11)$$

Equations (9)–(11) are straightforward to implement digitally, either in software or in hardware. Furthermore, these operations are spatially localized, which makes them particularly suitable for implementation in sections or on a parallel computing engine. Note that Eqs. (9)–(11) all three involve square roots that are subject to failure if the corresponding TKEO outputs are negative at some pixels. Estimates of the AM and FM functions at such points can be obtained by simple spatial interpolation. Conditions for positivity of the TKEO were studied in [33]. The main advantages to using the TKEO and ESA described in this section are that they are computationally efficient and that they estimate the AM and FM functions from the values of the real image alone. The main disadvantage is that an auxiliary algorithm must be used to estimate the signs of the components of the FM function.

2.2 Two-Dimensional Analytic Image

The analytic image is a multidimensional analogue of the 1D *analytic signal* [29, 34–36]. For a real-valued 1D continuous-time signal $f(t)$, the analytic signal is given by $z(t) = f(t) + j\mathcal{H}[f(t)]$, where $\mathcal{H}[f(t)] = f(t) * \frac{1}{\pi t}$ is the 1D Hilbert transform. The analytic signal has many important properties, including the following:

1. It associates a unique pair of AM and FM functions with $f(t)$.
2. The Fourier spectrum of $f(t)$ is redundant in the sense that it is conjugate symmetric (i.e., $F(\omega)$ is completely determined by its values on either the positive half-line or the negative half-line). The analytic signal $z(t)$ reduces this redundancy in a particularly elegant way: $Z(\omega)$ is equal to $2F(\omega)$ on the positive frequencies and it is identically zero on the negative frequencies.
3. If $z(t)$ is written as a power series and the real variable t is replaced by a complex variable v , then the signal $z(v)$ is a *holomorphic* or *analytic* function: In the upper half of the complex v -plane, it satisfies the Cauchy-Riemann equations and it is differentiable to arbitrary order.

The Hilbert transform has been extended into multiple dimensions in several different ways for various applications [37–41]. For AM-FM image modeling, what is needed is a multidimensional Hilbert transform that yields a complex image capable of representing arbitrary orientations. In other words, the arrows appearing in needle diagrams like the ones of Fig. 2D and 2H must be able to assume orientations in two quadrants: one where the signs of the components of the frequency vector $\nabla\varphi(x, y)$ are the same and one where they are different. Accomplishing this requires that the Fourier

transform of the complex image be identically zero over half of the 2D frequency plane and generally nonzero over the other half. For a real-valued image $f(x, y)$, this requirement is satisfied by the analytic image

$$z(x, y) = f(x, y) + j\mathcal{H}_2[f(x, y)], \quad (12)$$

where $\mathcal{H}_2[f(x, y)]$ is the 2D directional Hilbert transform given in [42, 43]. Like the 1D analytic signal, the analytic image (12) associates a unique pair of AM and FM functions with $f(x, y)$ and elegantly reduces the spectral redundancy inherent in the Fourier transform $F(\omega_x, \omega_y)$. However, unlike the 1D case, the analytic image generally fails to satisfy the multidimensional Cauchy-Riemann equations.

Once the analytic image $z(x, y)$ has been constructed, the AM and FM functions of $f(x, y)$ may be obtained according to [13]

$$\nabla\varphi(x, y) = \operatorname{Re}\left[\frac{\nabla z(x, y)}{jz(x, y)}\right], \quad (13)$$

$$a(x, y) = |z(x, y)|. \quad (14)$$

These algorithms are easy to verify by direct calculation if one substitutes $z(x, y) = a(x, y) \exp[j\varphi(x, y)]$ into the right-hand sides of (13) and (14). Intuitively, the FM algorithm (13) works because the exponential is invariant under differentiation (i.e., $\frac{d}{dx}e^u = e^u \frac{du}{dx}$). As an interesting aside, it should be noted that the imaginary part of the quotient in (13) also gives precisely the *instantaneous bandwidth* of $z(x, y)$ [36].

For a discrete $N \times M$ real-valued image $f(n_1, n_2)$, the analytic image is given by $z(n_1, n_2) = f(n_1, n_2) + jq(n_1, n_2)$, where $q(n_1, n_2)$ is the discrete 2D directional Hilbert transform of $f(n_1, n_2)$. If $\tilde{F}(u, v)$ is the DFT of $f(n_1, n_2)$, then the DFT of $q(n_1, n_2)$ is given by $\tilde{Q}(u, v) = \mathcal{H}_2(u, v)\tilde{F}(u, v)$, where [42]

$$\tilde{\mathcal{H}}_2(u, v) = \begin{cases} -j, & u = 1, 2, \dots, \frac{N}{2} - 1, \\ j, & u = \frac{N}{2} + 1, \frac{N}{2} + 2, \dots, N - 1, \\ -j, & u = 0, v = 1, 2, \dots, \frac{M}{2} - 1, \\ -j, & u = \frac{N}{2}, v = 1, 2, \dots, \frac{M}{2} - 1, \\ j, & u = 0, v = \frac{M}{2} + 1, \frac{M}{2} + 2, \dots, M - 1, \\ j, & u = \frac{N}{2}, v = \frac{M}{2} + 1, \frac{M}{2} + 2, \dots, M - 1, \\ 0, & \text{otherwise.} \end{cases} \quad (15)$$

Thus, $z(n_1, n_2)$ may be computed by the following straightforward procedure. First, the DFT is used to obtain $\tilde{F}(u, v)$ from $f(n_1, n_2)$. Second, $\tilde{Q}(u, v)$ is computed by taking the

pointwise product of $\tilde{F}(u, v)$ with $\tilde{\mathcal{H}}_2(u, v)$ as given in (15). Third, the DFT of $z(n_1, n_2)$ is computed according to $\tilde{Z}(u, v) = \tilde{F}(u, v) + j\tilde{Q}(u, v)$. Finally, $z(n_1, n_2)$ is obtained by taking the inverse DFT of $\tilde{Z}(u, v)$. A more efficient algorithm for calculating $\tilde{Z}(u, v)$ may be derived by realizing that, for each u and each v , $\tilde{Z}(u, v)$ assumes one of only three possible values: zero, $2\tilde{F}(u, v)$, or $\tilde{F}(u, v)$.

The demodulation algorithms (13) and (14) may be discretized by carefully applying the quasi-eigenfunction approximation (QEA) given in [13]. Once the analytic image $z(n_1, n_2)$ has been constructed, the AM function $a(n_1, n_2)$ can be obtained by [13]

$$a(n_1, n_2) = |z(n_1, n_2)|. \quad (16)$$

The magnitudes and signs of the components of the FM function $\nabla\varphi(n_1, n_2)$ can then be approximated according to [13]

$$|U(n_1, n_2)| \approx \arccos\left[\frac{z(n_1 + 1, n_2) + z(n_1 - 1, n_2)}{2z(n_1, n_2)}\right], \quad (17)$$

$$|\operatorname{sgn}U(n_1, n_2)| \approx \operatorname{sgn} \arcsin\left[\frac{z(n_1 + 1, n_2) - z(n_1 - 1, n_2)}{2jz(n_1, n_2)}\right], \quad (18)$$

$$|V(n_1, n_2)| \approx \arccos\left[\frac{z(n_1, n_2 + 1) + z(n_1, n_2 - 1)}{2z(n_1, n_2)}\right], \quad (19)$$

$$|\operatorname{sgn}V(n_1, n_2)| \approx \operatorname{sgn} \arcsin\left[\frac{z(n_1, n_2 + 1) - z(n_1, n_2 - 1)}{2jz(n_1, n_2)}\right]. \quad (20)$$

Like the ESA of Section 2.1, the demodulation algorithms (16)–(20) are easily implemented in hardware or software and are well suited for implementation in sections or on a parallel processor. Two comments are in order concerning the frequency algorithms (17)–(20). First, these algorithms cannot be applied at pixels where $z(n_1, n_2) = 0$. At such pixels, frequency estimates may be obtained by interpolating the frequency estimates from neighboring pixels. Second, the arguments of the transcendentals in (17)–(20) are guaranteed to be real up to approximation errors; any nonzero imaginary component should be discarded prior to evaluating the arccos and arcsin functions. The main advantages of using the analytic image-based technique described in this section are that it is computationally efficient and that it estimates signed frequency. The main disadvantage is that the complex analytic image must be calculated explicitly.

3 Practical Techniques for AM-FM Image Modeling

The fundamental AM-FM modeling techniques described in Section 2 work well when applied to single-component images

like those shown in Fig. 2E and 2I. Each local neighborhood in a single-component image can be described by a single nonstationary AM-FM component (i.e., a single “bendable” sinusoid). However, most images encountered in practical applications are inherently multicomponent. The characteristics that distinguish a multicomponent signal or image are described technically in [36, 44, 45]. The intuitive notion is quite straightforward: multicomponent images are made up of local neighborhoods that are best characterized as sums of multiple nonstationary sinusoidal components, each having AM and FM functions that vary smoothly.

If applied directly to a multicomponent image, the discrete demodulation algorithms (9)–(11) and (16)–(20) will generally suffer from large approximation errors and fail. The most obvious means of circumventing this problem is to apply multiband linear filtering to the image to isolate the multiple AM-FM components from one another on a spatially local basis. Since all of the demodulation algorithms described in Section 2 are small in spatial extent, the filters need not isolate components on a global scale. But they must be spatially localized to avoid averaging out fine-scale nonstationary structure and also spectrally localized to reject cross-component interference during demodulation. What is required is that the response of each filter be dominated by at most one AM-FM image component at each pixel. In Section 3.1, we explain how (9)–(11) and (16)–(20) may be modified so that they can be applied to each filter response to estimate the AM and FM functions of the component that dominates the filter on a pixel-by-pixel basis.

3.1 Filtered Demodulation

Suppose that we are given a real-valued discrete image $f(n_1, n_2)$ and that the mean has been subtracted out so that the spatial average of the image is zero. Our goals are to model this image as a multicomponent AM-FM function according to

$$f(n_1, n_2) = \sum_{k=1}^K a_k(n_1, n_2) \cos[\varphi_k(n_1, n_2)] \quad (21)$$

and to compute estimates of the AM and FM functions for each of the K image components

$$f_k(n_1, n_2) = a_k(n_1, n_2) \cos[\varphi_k(n_1, n_2)]. \quad (22)$$

To achieve this, we process the image with a bank of linear translation invariant (LTI) bandpass filters. Design considerations for the filterbank are discussed in Section 3.2. Each individual filter in the bank is referred to as a *channel*. Our assumptions are that the filters are of the bandpass LTI type and that, at each pixel, the response of each channel is dominated by at most one AM-FM image component $f_k(n_1, n_2)$ in (21). This last assumption is satisfied automatically in practice, since the decomposition of the image into

components as indicated in (21) is not unique and is, in fact, determined by the structure of the filterbank. The demodulation problem then becomes one of estimating the AM and FM functions in (21) from the filterbank channel responses.

Assume first that the filters are real-valued and let $g_i(n_1, n_2)$ and $G_i(U, V)$ be the impulse response and frequency response of one of the filterbank channels. At a particular pixel, suppose that the channel response $y_i(n_1, n_2)$ is dominated by the image component $f_k(n_1, n_2)$, so that, in a neighborhood about the pixel, $y_i(n_1, n_2) = f(n_1, n_2) * g_i(n_1, n_2) \approx f_k(n_1, n_2) * g_i(n_1, n_2)$. Then, under mild and realistic assumptions one may show that, in the neighborhood,

$$\begin{aligned} \Phi_d[y_i(n_1, n_2)] &\approx a_k^2(n_1, n_2) |\nabla \varphi_k(n_1, n_2)|^2 |G_i[\nabla \varphi_k(n_1, n_2)]|^2 \\ &\approx \Phi_d[f_k(n_1, n_2)] |G_i[\nabla \varphi_k(n_1, n_2)]|^2. \end{aligned} \quad (23)$$

Note that the energy operator appears in both the numerators and denominators of (9) and (10). If these frequency demodulation algorithms are applied to $y_i(n_1, n_2)$, then the scaling by $|G_i[\nabla \varphi_k(n_1, n_2)]|^2$ indicated in (23) is approximately canceled by division. Thus, (9) and (10) can be applied directly to a channel response to estimate the FM function of the component that dominates that channel at any given pixel.

Moreover, multiband filtering provides a convenient means of approximating the relative signs of the frequency estimates (9) and (10). Let $|U_k(n_1, n_2)|$ and $|V_k(n_1, n_2)|$ be the magnitude frequency estimates obtained by demodulating $y_i(n_1, n_2)$ at a particular pixel. Since $g_i(n_1, n_2)$ is by assumption real-valued in this case, the frequency response $G_i(U, V)$ will be conjugate symmetric. Hence, the bandpass characteristic $|G_i(U, V)|$ will have two mainlobes with center frequencies located either in quadrants one and three or in quadrants two and four of the 2D frequency plane. If the signs of the horizontal and vertical components of these center frequencies agree, we take $\text{sgn } U_k(n_1, n_2) = \text{sgn } V_k(n_1, n_2) = +1$. Otherwise, we set $\text{sgn } U_k(n_1, n_2) = +1$ and $\text{sgn } V_k(n_1, n_2) = -1$.

Unlike the frequency algorithms just discussed, the amplitude demodulation algorithm (11) requires explicit modification before it can be applied directly to a filterbank channel response. This is because the image components $f_k(n_1, n_2)$ in (21) are individually scaled as they pass through the filterbank. In particular, the amplitude estimates obtained by applying (11) to $y_i(n_1, n_2)$ must be divided by $|G_i[\nabla \varphi_k(n_1, n_2)]|$. Thus, once $\nabla \varphi_k(n_1, n_2)$ has been estimated by applying (9) and (10) to $y_i(n_1, n_2)$, the modified amplitude estimation algorithm of the ESA is given by

$$a_k(n_1, n_2) \approx \frac{\sqrt{\Phi_d[y_i(n_1, n_2)]}}{\left\{ \begin{array}{l} |G_i[U_k(n_1, n_2), V_k(n_1, n_2)]| \\ \times \sqrt{\sin^2(|U_k(n_1, n_2)|) + \sin^2(|V_k(n_1, n_2)|)} \end{array} \right\}}. \quad (24)$$

We next consider filtered versions of the demodulation algorithms (16)–(20). As before, the real-valued discrete image $f(n_1, n_2)$ is modeled according to (21). Since the 2D directional Hilbert transform (15) is linear, it commutes with sums. By applying the transform to both sides of (21) we obtain the complex-valued multicomponent analytic image

$$\begin{aligned} z(n_1, n_2) &= f(n_1, n_2) + j\mathcal{H}_2[f(n_1, n_2)] \\ &= \sum_{k=1}^K a_k(n_1, n_2) \exp[j\varphi_k(n_1, n_2)]. \end{aligned} \quad (25)$$

When $z(n_1, n_2)$ is input to the filterbank, the channel responses are given by $y_i(n_1, n_2) = z(n_1, n_2) * g_i(n_1, n_2) \approx z_k(n_1, n_2) * g_i(n_1, n_2)$, where filterbank channel i is dominated by component $z_k(n_1, n_2)$ in a neighborhood about the pixel (n_1, n_2) .

Because the Hilbert transform is implemented using the spectral multiplier $\tilde{\mathcal{H}}_2(u, v)$ given in (15) and since 2D multiband linear filtering is almost always implemented using pointwise multiplication of DFTs, great computational savings can be obtained by combining the linear filtering and analytic image generation into a single operation. If $\tilde{G}_i(u, v)$ is the DFT of $g_i(n_1, n_2)$, then the channel response $y_i(n_1, n_2)$ can be obtained by taking the inverse DFT of

$$\tilde{Y}_i(u, v) = \tilde{G}_i(u, v) \left[1 + j\tilde{\mathcal{H}}_2(u, v) \right] \tilde{F}(u, v). \quad (26)$$

Half of the frequency samples in $\tilde{Z}(u, v)$ are identically zero, so (26) actually saves half of the complex multiplies required to implement the convolution performed by each filterbank channel. It should be noted that the combined LTI operation indicated in (26) implicitly involves channel filters having complex-valued impulse responses.

Similar to the case of the filtered TKEO and ESA, careful application of the quasi-eigenfunction approximation as described in [13] verifies that the frequency demodulation algorithms (17)–(20) may be applied directly to the channel response $y_i(n_1, n_2)$ to estimate the FM function of the component that locally dominates the response on a pixel-by-pixel basis. Once this has been done, the amplitude demodulation algorithm (16) must be explicitly modified to account for the scaling effects of the filter. The resulting filtered amplitude algorithm is given by

$$a_k(n_1, n_2) \approx \frac{|y_i(n_1, n_2)|}{|G_i[U_k(n_1, n_2), V_k(n_1, n_2)]|}. \quad (27)$$

Thus, with an appropriate bank of LTI bandpass filters $g_i(n_1, n_2)$, the AM and FM functions of all components in the multicomponent image model (21) may be estimated using (9), (10), and (24) or using (17)–(20) and (27).

3.2 Filterbank Design Considerations

The following list summarizes the six main filterbank design requirements that have been implied thus far in Section 3:

1. The channel filters should be linear and translation invariant (LTI).
2. Each filter should be spatially localized to avoid too much smoothing, which could average out important nonstationary structure in the image. In addition, filters that are well localized spatially will tend to minimize approximation errors in the filtered demodulation algorithms [13, 14].
3. The filters should be spectrally localized to resolve input components that are closely spaced in frequency and to avoid cross-component interference.
4. The response of each filter should be dominated by at most one AM-FM image component at each pixel.
5. The filterbank channels should sample the 2D frequency plane densely so that all of the visually important structure of the input image is manifest in the channel responses.
6. The filterbank should yield a decomposition of the image into a sum of nonstationary components having smoothly varying modulations. This is critical for obtaining a representation in terms of visually significant choices for the AM and FM functions, as well as for minimizing approximation errors in the filtered demodulation algorithms. What this implies is that the channels must have frequency responses with smooth transition bands and must be selective in terms of both the magnitude and the orientation of the 2D frequency vector.

Unfortunately, there are no known filters that satisfy all six requirements. The second and third requirements are conflicting as a consequence of the reciprocal spreading property of the Fourier transform: as one attempts to make a filter more localized in one domain, it becomes less localized in the other. This fact strongly suggests the use of 2D Gabor filters as was argued in [13, 14]. The continuous-space Gabor functions uniquely realize the uncertainty principle lower bound on joint spatirospectral localization [34, 36, 46–48] and discrete Gabor functions are approximately optimal in this regard. Furthermore, requirements four and six tend to be best satisfied by filters that have good joint localization properties, again suggesting the use of Gabor filters. In fact, Gabor filters are suboptimal for AM-FM image modeling in only one respect: taken together, requirements four through six imply that the filters should be mutually orthogonal and, ideally, should constitute a perfect reconstruction system. Since Gabor filters fail to admit finite support in either domain, they can provide neither orthogonality nor perfect reconstruction. Nevertheless, as described in

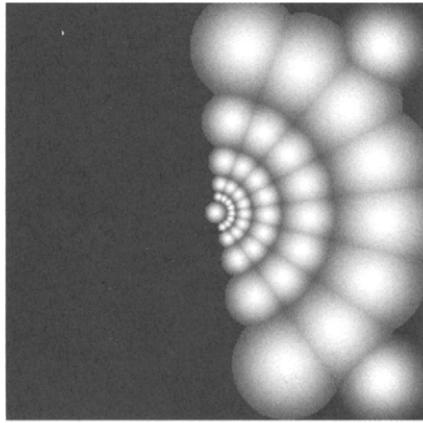


FIGURE 3 Frequency response of a two-dimensional Gabor filterbank suitable for AM-FM image modeling. Nonlinear scaling of the gray levels has been applied to accentuate the intersections between filters.

Chapter 4.9, Gabor filters are popular choices in a variety of image processing and computer vision applications due to their other good properties. Moreover, as described in Chapter 4.1, Gabor filtering is a plausible explanation for several of the processes known to occur in the early stages of human vision.

For the remainder of Section 3, we assume the use of a densely spaced bank of complex-valued Gabor filters like the one described in [10, 13–15]. The frequency response of a filterbank of this type is shown in Fig. 3, where two high-frequency filters have been added to cover the corners of the right frequency half-plane. The design of this filterbank is described in detail in [49]. Each filter in the regular polar tessellation has a half-peak bandwidth of one octave, and any four adjacent filters intersect at a single frequency where each one is at half of its peak response. The baseband (DC) filter in the center of the figure is also designed to intersect the innermost ring of AC filters at half-peak.

3.3 Dominant Component Analysis

In this section we describe a multicomponent computational technique called dominant component analysis, or DCA, which at every pixel delivers estimated modulating functions $a_D(n_1, n_2)$ and $\nabla\varphi_D(n_1, n_2)$ corresponding to the AM-FM component that is locally dominant at that pixel [13, 15, 49]. The dominant frequency vectors $\nabla\varphi_D(n_1, n_2)$ are often referred to as the *emergent frequencies* of the image [14]. Generally, different components in the sums (21) and (25) are expected to be dominant in different image regions. A block diagram of DCA is shown in Fig. 4. The real image $f(n_1, n_2)$ or the analytic image $z(n_1, n_2)$ is passed through a multiband linear filterbank such as the one shown in Fig. 3. The demodulation algorithms (9), (10), and (24) or (17)–(20) and (27) are then applied to the response of every filterbank channel in the blocks labeled DEMOD in Fig. 4.

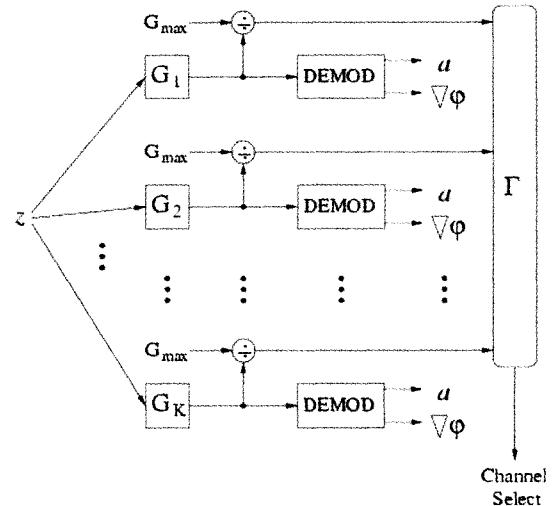


FIGURE 4 Block diagram of dominant component analysis (DCA).

The dominant component at each pixel is defined as the one that dominates the response of the channel that maximizes a channel selection criterion $\Gamma_i(n_1, n_2)$. For the ESA-based and analytic image-based demodulation approaches, $\Gamma_i(n_1, n_2)$ is given respectively by

$$\Gamma_i(n_1, n_2) = \frac{\Phi_d[y_i(n_1, n_2)]}{|\nabla\varphi_k(n_1, n_2)|^2 \max_{U,V} |G_i(U, V)|^2}, \quad (28)$$

and

$$\Gamma_i(n_1, n_2) = \frac{|y_i(n_1, n_2)|}{\max_{U,V} |G_i(U, V)|}, \quad (29)$$

where $\nabla\varphi_k(n_1, n_2)$ in (28) is the vector-valued frequency estimate obtained by demodulating $y_i(n_1, n_2)$. As explained in [13], these criteria tend to select the channel having its center frequency closest to the instantaneous frequency vector of the component with the largest amplitude. They also provide superior rejection of “false” or bogus components arising from cross-channel interference or random noise that lies far outside the filter passband. Dominant modulations $a_D(n_1, n_2)$ and $\nabla\varphi_D(n_1, n_2)$ are extracted from the channel that maximizes $\Gamma_i(n_1, n_2)$ on a pointwise basis.

The dominant modulations provide a rich description of the local texture structure of the image. As we pointed out in Section 1, they have been used with great success in a number of important image processing and computer vision applications. The dominant AM and FM functions shown in Fig. 2C, D, G, and H were computed by DCA using the filtered demodulation algorithms (17)–(20) and (27) with the filterbank of Fig. 3. Since the Fourier spectra of most images are dominated by extremely low frequency information representing the mean pixel value and by large-scale shading and contrast variations due to lighting conditions, the mean

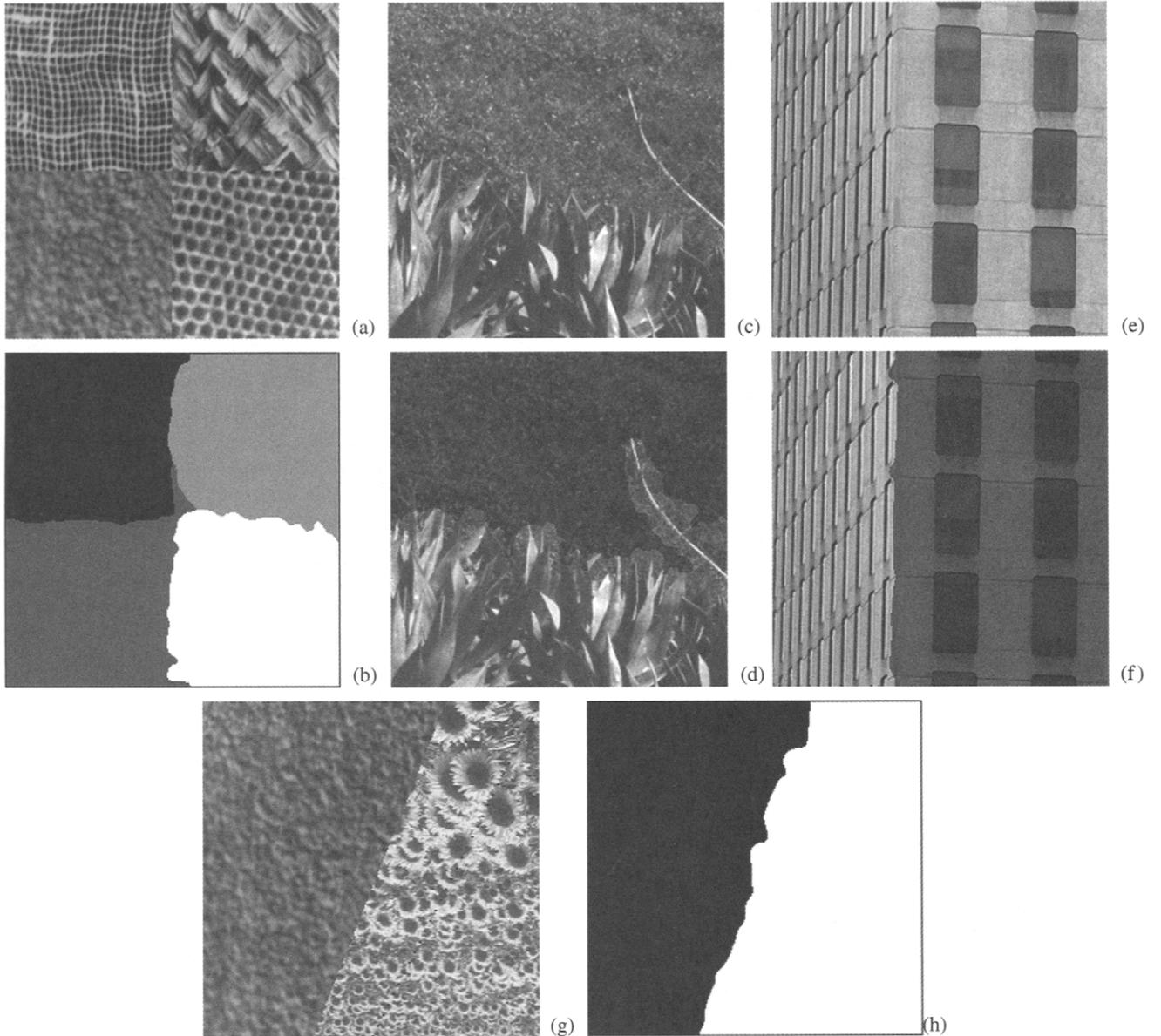


FIGURE 5 Dominant component analysis (DCA) segmentation examples. A: Four-texture image. B: DCA segmentation; each segment is shown with a distinct gray tone. (From [52], with permission). C: Natural two-texture image. D: DCA segmentation; pixel values in the upper segment are divided by two. (From [52], with permission). E: Natural two-texture image. F: DCA segmentation; pixel values in the right segment are divided by two. (From [18], with permission). G: Two-texture image. H: DCA segmentation; each segment is shown with a distinct gray tone. (From [16], with permission). (© 2005 IEEE.)

value of the image was subtracted out prior to filtering and the DC filter shown in Fig. 3 was removed for DCA. Postfiltering was also applied as described in Section 3.6 and in Fig. 8.

To illustrate a DCA-based application, several texture segmentation examples are shown in Fig. 5. In each case, the dominant modulations $a_D(n_1, n_2)$ and $\nabla\varphi_D(n_1, n_2)$ were used to construct feature vectors for performing segmentation in the modulation domain. The image in Fig. 5A is a juxtaposition of four Brodatz textures [50], while Fig. 5C is a natural image from the VisTex database [51]. These two examples are partially supervised in the sense that the number of texture

segments — four in the case of Fig. 5A and two in the case of Fig. 5C — was used as *a priori* information. However, no information about the types of textures or their characteristics was supplied to the algorithm. After subtracting out the mean of each image, DCA was applied using the filterbank of Fig. 3 (with the DC channel removed) and using postfiltering as described in Section 3.6. The segmentations shown in Fig. 5B and 5D were computed by applying k -means clustering to the modulation domain feature vectors as described in [52]. In the result of Fig. 5B, 97.3% of the pixels are correctly classified. While ground truth is not available for the image of Fig. 5C,

the result in Fig. 5D is remarkably consistent with our intuition about where the boundary between the two regions should be located.

The image of Fig. 5E is also from the VisTex database. The segmentation in Fig. 5F was obtained by applying k -means clustering with cross-validation to the dominant modulations computed by DCA. As described in [18], this example is *fully unsupervised*: absolutely no *a priori* information whatsoever about the number of textures or their characteristics was used. Likewise, the example of Fig. 5G and 5H is fully unsupervised. In this case, the segmentation was computed by applying the maximum likelihood approach described in [18]. The percentage of pixels correctly classified in Fig. 5H is 97.0%.

3.4 Channelized Components Analysis

A general block diagram for performing image processing in the modulation domain is shown in Fig. 6. On the left half, Fig. 6A shows the analysis portion, which involves using (9), (10), and (24) to compute a real-valued multicomponent AM-FM model according to (21) or using (17)–(20) and (27) to compute a complex-valued model according to (25). On the right half, Fig. 6B shows the signal processing operation, the definition of which is currently an open research topic, and the synthesis portion (reconstruction) as described in Section 3.5. In this section, we consider the simplest case wherein the signal processing block of Fig. 6B is the “null” or “do nothing” operation. The block diagram then represents an *AM-FM transform*.

The question of how best to decompose a complicated image into a sum of components for AM-FM modeling is a difficult one and sophisticated techniques for following

the frequency tracks of the various components have been investigated [5, 7, 53]. Here we consider a particularly simple approach that often works well in practice: we consider that the filterbank channels partition the image into components on a spatially *global* basis. Each resulting AM-FM component is restricted to lie in a single channel over the entire image domain. Therefore, we refer to this approach as channelized components analysis (CCA). With CCA, the number of components in the computed image model is necessarily equal to the number of channels in the filterbank. As depicted in Fig. 6A, estimates of the modulating functions of each individual component are obtained by simply demodulating the corresponding filter response on a global basis. The CCA representation provides a dense description characterizing not only the dominant image structures, but subtle subemergent texture features as well. It should be noted that exact reconstruction from the CCA model is possible in general only if the analysis and synthesis processes constitute a perfect reconstruction system, which they usually do not. Some of the reasons for this were discussed in Section 3.2, and we will look into the issue in more detail in Section 4.

The multicomponent woodgrain image reconstruction shown in Fig. 2B was computed using CCA with the reconstruction algorithm described in Section 3.5. The mean of the image was subtracted out and the complex analytic image was processed with the filterbank of Fig. 3. The channel responses were demodulated using (17)–(20) and (27). Postfiltering was applied as described in Section 3.6. While it is not a perfect reconstruction, Fig. 2B demonstrates that virtually all of the visually important information in the image has been captured by the CCA multicomponent AM-FM model. Another example is given in Fig. 7. The DCA reconstruction of the dominant AM-FM component is shown in

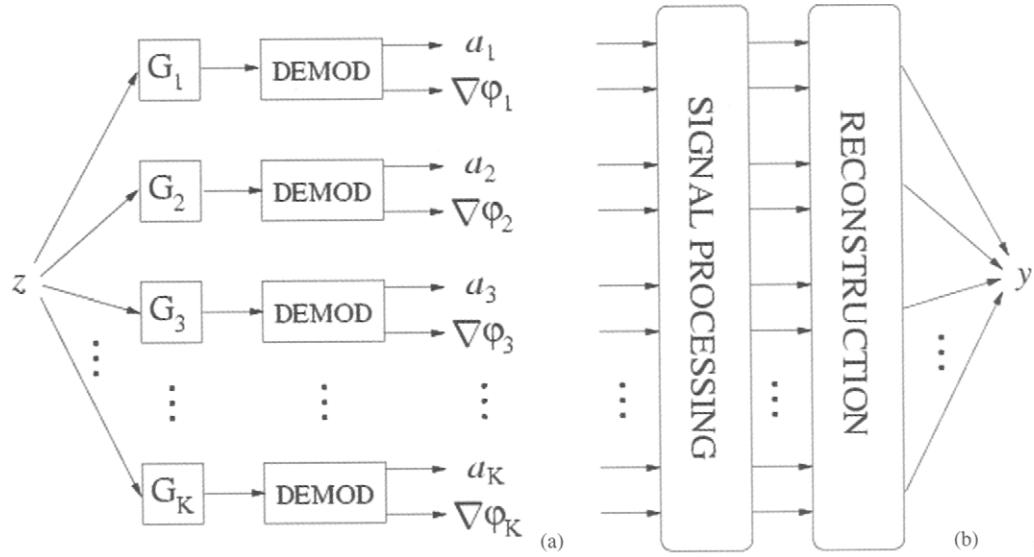


FIGURE 6 Modulation domain signal processing. **A:** Block diagram of channelized components analysis (CCA). **B:** Block diagram of signal processing and reconstruction.

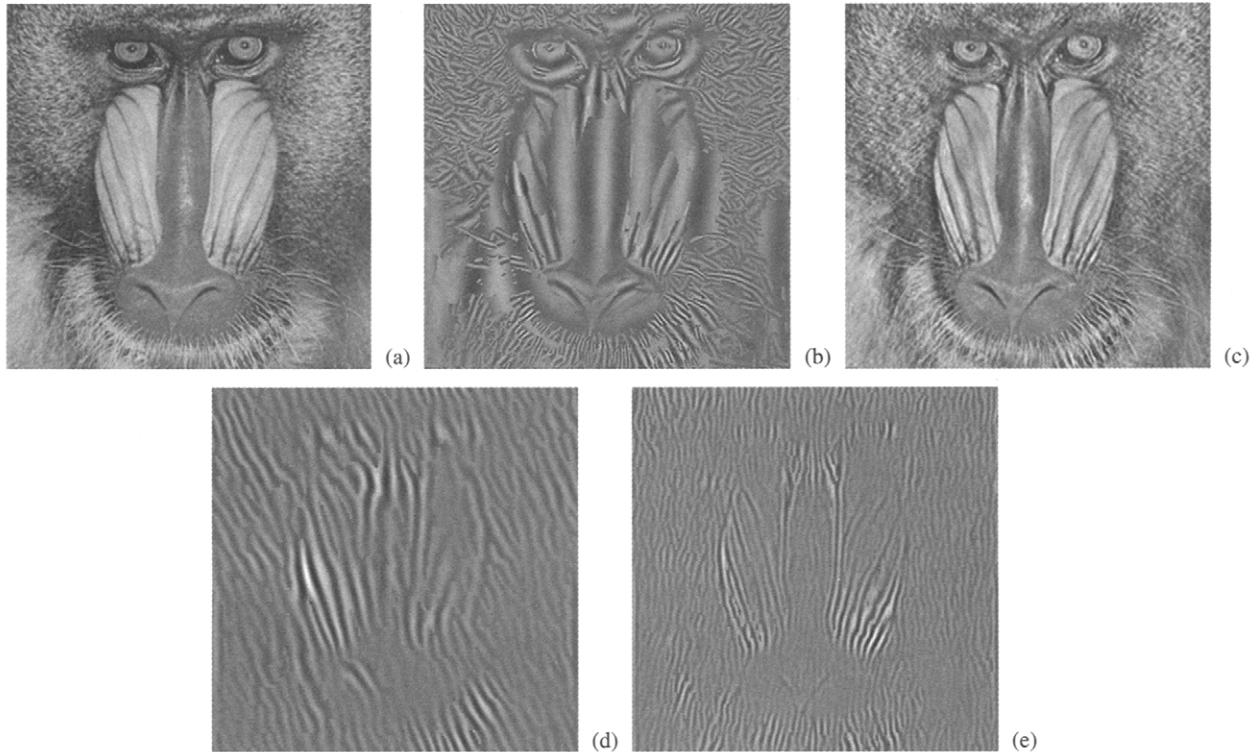


FIGURE 7 Multicomponent example. A: Original *mandril* image. B: Reconstruction of dominant component computed by dominant component analysis (DCA). C: 43-component channelized components analysis (CCA) reconstruction. D, E: Two of the individual AM-FM [amplitude modulation (AM); frequency modulation (FM)] components from the CCA image reconstruction in (C).

Fig. 7B, while Fig. 7C gives the 43-component CCA reconstruction. Two of the individual channelized AM-FM image components are shown in Fig. 7D and 7E.

3.5 Reconstruction

Reconstructing AM-FM image components from their estimated amplitude and frequency modulations is an important part of the overall modulation domain signal processing paradigm shown in Fig. 6. The application of DCA or CCA to an image involves estimating the AM and FM functions of all image components at every pixel. Although it is a straightforward matter to reconstruct any given component (22) in (21) once the AM function $a_k(n_b, n_2)$ and instantaneous phase function $\varphi_k(n_b, n_2)$ have been estimated, reconstruction of the phase from the FM function $\nabla\varphi_k(n_b, n_2)$ is in general a difficult and ill-posed inverse problem. In this section we present a simple algorithm for performing phase reconstruction [53]. While it is certainly suboptimal, this algorithm is often effective in practice, and it was used to compute the AM-FM reconstructions shown in Fig. 2B, 2E, and 2I, as well as those shown in Figs. 7B–E.

Suppose that we want to reconstruct the component $f_k(n_b, n_2)$ given modulating function estimates $a_k(n_b, n_2)$ and $\nabla\varphi_k(n_b, n_2)$. For any pixel that does not lie in the uppermost row or leftmost column of the image, we can easily obtain

independent horizontal and vertical estimates of the phase using a first-order Taylor series approximation for the derivative and truncating the higher order terms. Averaging these two estimates, we have

$$\begin{aligned} \varphi_k(n_1, n_2) = & \frac{1}{2} [\varphi_k(n_1 - 1, n_2) + U_k(n_1 - 1, n_2) \\ & + \varphi_k(n_1, n_2 - 1) + V_k(n_1, n_2 - 1)], \end{aligned} \quad (30)$$

where $\nabla\varphi_k(n_b, n_2) = [U_k(n_b, n_2) \ V_k(n_b, n_2)]^T$ as defined in (2) and (3). Equation (30) can be used to reconstruct the phase on the entire image domain, provided that we save phase samples $\varphi_k(1, n_2)$ and $\varphi_k(n_1, 1)$ as initial conditions during the DCA or CCA analysis phase. There is no difficulty in doing this, since Gabor filters have real-valued spectra.³ Thus, for any given channel, the QEA [13] implies that the instantaneous phase of the channel response is approximately equal to that of the input component that dominates the channel.

³Assuming that the impulse response has not been spatially shifted (i.e., that it is centered at the origin of the spatial coordinate system), any complex-valued Gabor filter has a frequency response that is a real-valued Gaussian. Similarly, the frequency response of a real-valued, origin-centered Gabor cosine filter is the sum of two real-valued Gaussians. In either of these cases, the frequency response has a phase that is identically zero.

The main problem with this approach is that the frequency estimates $\nabla\varphi_k(n_1, n_2)$ generally contain approximation errors that propagate cumulatively unless the phase reconstruction process is periodically reset. Therefore, it is preferable to partition the image into disjoint blocks and perform phase reconstruction independently on each block. A block size of 4×4 pixels was used for all of the examples in this chapter. The value of the phase at the upper left corner of each block was saved as an initial condition during analysis. Phase reconstruction was then performed along the first row of the block using the horizontal frequency estimates alone according to $\varphi_k(n_1, n_2) = \varphi_k(n_1 - 1, n_2) + U_k(n_1 - 1, n_2)$. Similarly, phase reconstruction was performed on the first column of the block according to $\varphi_k(n_1, n_2) = \varphi_k(n_1, n_2 - 1) + V_k(n_1, n_2 - 1)$. Equation (30) was then used to reconstruct the phase at all interior pixels of the block.

Once the phase reconstruction is completed on all blocks, the component $f_k(n_1, n_2)$ is reconstructed using (22), as illustrated in Fig. 7D and 7E. The CCA multicomponent AM-FM image reconstructions in Fig. 2B and Fig. 7C were obtained by summing the individual reconstructed components according to (21). A slightly different approach was required for the DCA reconstructions shown in Fig. 2E, 2I, and Fig. 7B. All of the individual components $f_k(n_1, n_2)$ were reconstructed exactly as described above. At each pixel, the dominant components were then set equal to the value of the reconstructed component that maximized the channel selection criterion (28) or (29) at the pixel.

3.6 Postfiltering and Postprocessing

The DCA and CCA techniques presented in Sections 3.3 and 3.4 generally yield high-quality AM-FM models for images that are locally coherent in the sense of having smooth modulating functions. Synthetic images generated from analytic equations are often of this type, as are some naturally occurring images. For many other images, however, additional processing may be required to compensate for certain perturbations that can occur in the Gabor filter channel responses and in the computed AM and FM functions themselves. The postfiltering operations described in Section 3.6.1 are useful for smoothing the computed modulations and were applied in all of the examples shown in Fig. 2, Fig. 5, and Fig. 7. Less frequently, the postprocessing described in Section 3.6.2 is needed to remove outliers and compress the dynamic range when displaying AM-FM image reconstructions or when displaying the computed modulating functions as images. This kind of post processing was used only for the AM image of Fig. 2G, where a 3×3 median filter was applied to reduce the dynamic range of the floating point data for 8-bit display.

3.6.1 Postfiltering

It is well known that the textures occurring in images of general practical interest can contain shifts, breaks, or

discontinuities in their local surface patterns that may lead to Gabor filter responses that fail to be everywhere smooth. Such perturbations from the smooth, locally coherent image model considered in this chapter are most often manifest as nonsmooth variations or localized excursions in the amplitude envelopes and phases of the filter responses. Occlusions, shadows, specularities, and surface defects are but a few among the many factors that can give rise to these perturbations, which are almost always problematic. For example, Gabor filter magnitude and phase responses were used as features for segmenting textured images in [12] and [54], where it was observed that nonsmooth perturbations in these features could result in the failure to obtain a consistent segmentation. Low-pass postfilters were applied to the Gabor filter responses prior segmentation to effectively compensate for this problem. Each postfilter was a Gaussian with a spatial standard deviation (or *space constant*) $\frac{3}{2}$ as large as that of the corresponding Gabor channel filter. Thus, the postfilters had envelopes of the same shape as the channel filters, but were somewhat larger in spatial extent and correspondingly more narrowband.

When, as is the case with the DCA and CCA techniques presented in Sections 3.3 and 3.4, the desired features are explicitly computed AM and FM functions as opposed to channel filter magnitude and phase responses, it is appropriate to apply postfiltering directly to the computed modulations. This approach was adopted in [2] where median filtering was applied to the computed FM function and in [13, 55] where Gaussian postfilters were used. The importance of postfiltering cannot be overemphasized: In the vicinity of nonsmooth amplitude and phase variations, the demodulation algorithms (9)–(11) and (16)–(20) are subject to significant approximation errors. The postfiltered channel model used for all of the examples in Fig. 2, Fig. 5, and Fig. 7 is shown in Fig. 8, where the postfilters P_i are low-pass Gaussians with space constants equal to those of the corresponding Gabor channel filters G_i . Frequency demodulation is applied to the channel response and the components of the obtained FM vectors are independently smoothed by postfiltering. The smoothed frequencies are used in the amplitude demodulation algorithm, and the resulting AM function is then postfiltered.

3.6.2 Postprocessing

Further postprocessing in addition to the postfiltered channel model of Fig. 8 is sometimes required when the input image contains phase discontinuities. At the site of a phase

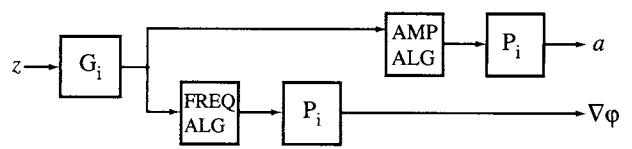


FIGURE 8 Postfiltered channel model.

discontinuity, the FM function $\nabla\varphi(n_1, n_2)$ can exhibit essentially unbounded excursions that may remain large in magnitude even after postfiltering [55]. Since the magnitudes of the channel filter frequency responses necessarily fall off rapidly outside the passband, the term $|G_i[U_k(n_1, n_2), V_k(n_1, n_2)]|$ in the denominators of (24) and (27) may become vanishingly small. As a result, these filtered amplitude algorithms, which were designed under the assumption of locally smooth input modulations, may yield absurdly large amplitude spikes in a small neighborhood about the phase discontinuity. When these spikes occur, they are sometimes sufficiently large so as to render it impossible to meaningfully display the estimated AM function and reconstructed image using 8-bit gray scales. This condition should be suspected any time the reconstructed image or computed AM function exhibit values that are many times larger than the largest gray scales occurring in the original image. Depending on the severity and spatial extent of the amplitude spikes, one or a few applications of a 3×3 median filter, as was used in Fig. 2G, often yields a satisfactory result. If the dynamic range of the AM function is still too large after median filtering, then additional compression techniques should be considered. One approach that we have found useful is to empirically determine a threshold and logarithmically compress all pixels that exceed it in magnitude.

4 Emerging Trends in AM-FM Image Modeling

The DCA and CCA techniques presented in Section 3 deliver computed AM-FM models that are of sufficient quality to be demonstrably useful and effective for solving a variety of important image processing and computer vision analysis problems; we enumerated several of these in the concluding paragraph of Section 1. However, these techniques fail to provide perfect reconstruction within the framework of Fig. 6, even when the signal processing block is set to the “do nothing” operation. Thus, they do not provide an invertible AM-FM transform and this fact has slowed the emergence of important applications including AM-FM image coding and the development of general theories for modulation domain image processing. We have already mentioned the reasons why CCA models as described in Section 3 cannot provide perfect reconstruction and we reiterate these below with some additional discussion:

1. The amplitude and frequency demodulation algorithms contain inherent approximation errors. The unfiltered algorithms (5) and (6) are exact when applied to a real-valued monochromatic signal, while (13) and (14) are exact when applied to a complex signal. However, the discrete unfiltered versions (9)–(11) and (16)–(20) involve approximations with nonzero errors. Moreover,

all of the discrete filtered algorithms given in Section 3.1 involve approximation errors that are often small but nevertheless nontrivial in the context of Fig. 6 unless the input is purely monochromatic.

2. Gabor filterbanks cannot provide perfect reconstruction, as was discussed in some detail in Section 3.2.⁴ While the filtered demodulation algorithms given in Section 3.1 are essentially independent of the particular filterbank choice and many mutually orthogonal filterbanks are available, these generally involve channel frequency responses with crisp edges that imply poor joint localization, whereas good joint localization is critical for minimizing approximation errors in the demodulation algorithms. This problem is compounded by the fact that there is no straightforward way to extend the uncertainty principle, which quantifies joint localization, from the continuous to the discrete case.
3. The reconstruction algorithm given in Section 3.2 is admittedly simplistic. The presence of frequency estimation errors means that the computed phase gradient field is nonconservative and thus the phase reconstruction problem is ill-posed. Even if the frequency estimates were perfect, nonzero errors arise in (30) due to modeling the partials of the phase with first-order differences. It is unclear how any discrete algorithm can resolve this issue in a way that is completely free of approximation errors. The block-based approach described in Section 3.2 also tends to produce visible blocking artifacts when the block size is large.

In this section we briefly sketch some exciting techniques that have begun to emerge very recently and that seem to offer solutions to all of these problems.

4.1 Perfect Reconstruction Filterbanks

The theory of 1D continuous-time orthogonal wavelet transforms has been recognized as an attractive framework for developing discrete subband signal decompositions that are capable of providing perfect reconstruction [56–58]. These are referred to as discrete wavelet transforms (DWTs). As described in [59] and in Chapter 4.2, the application of this theory for constructing separable 2D filterbanks is well understood. So it is natural to ask why wavelets were not selected for building the filterbank in Section 3.2.

One reason is that only Gabor filters can achieve the uncertainty principle lower bound on joint localization in the continuous case. With reference to Fig. 6A in Chapter 4.2, another reason is that the computation of wavelet subbands

⁴It should be noted that perfect reconstruction can be achieved with the *Gabor transform*. But that is a biorthogonal transform involving inner products (or convolutions) between the input signal and a family of non-Gabor auxiliary functions as opposed to with the Gabor functions themselves. The auxiliary functions are not smooth and do not possess optimal joint localization [49].

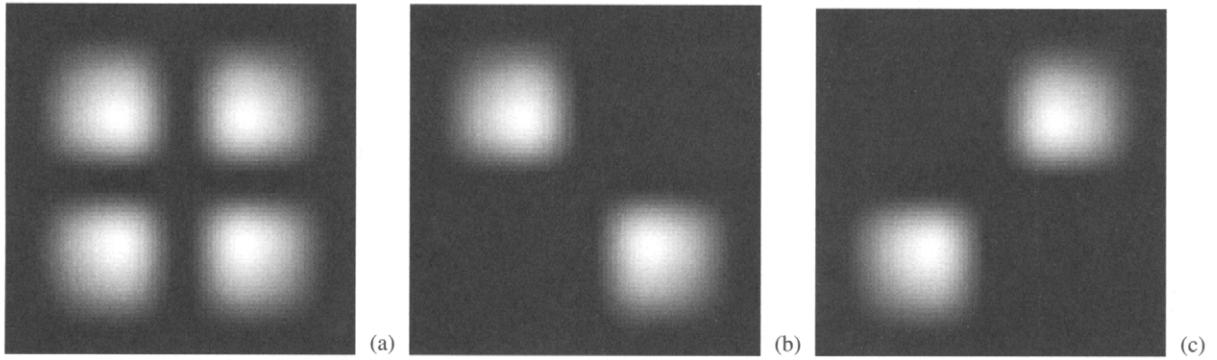


FIGURE 9 Decomposition of a non-orientation selective, separable parallel filterbank channel (A) into two channels that are orientation selective but nonseparable (B, C). The log-magnitude spectra are shown. (From [60], with permission.) (© 2005 IEEE.)

involves downsampling between levels. The various subbands in a multilevel DWT are of different sizes and this complicates interpretation of the results when these subbands are demodulated. The cascaded structure shown in Fig. 7A and 7B of Chapter 4.2 also involves iterated application of the analysis and synthesis filters, which is inefficient from a computational standpoint. More importantly, as is readily evident in Fig. 1C and 6B of Chapter 4.2, the subbands in a separable DWT are neither orientation selective nor smooth in general. Thus, one can expect nontrivial approximation errors when applying the discrete demodulation algorithms of Section 3.1 to these subbands.

The key that has emerged for resolving these problems is to implement the DWT in a tricky way as described in [60, 61]. First, the Noble identities are applied to transform the 1D M -level DWT into an $M + 1$ channel parallel filterbank [62]. Each channel of this filterbank produces one of the 1D wavelet subbands in a single convolution operation followed by appropriate decimation. For example, a four-level DWT with low- and high-pass analysis filters $f_a[n]$ and $g_a[n]$ can be implemented as a parallel filterbank with five channels having transfer functions given by

$$H_0(z) = F_a(z)F_a(z^2)F_a(z^4)F_a(z^8), \quad (31)$$

$$H_1(z) = F_a(z)F_a(z^2)F_a(z^4)G_a(z^8), \quad (32)$$

$$H_2(z) = F_a(z)F_a(z^2)G_a(z^4), \quad (33)$$

$$H_3(z) = F_a(z)G_a(z^2), \quad (34)$$

$$H_4(z) = G_a(z). \quad (35)$$

These transfer functions are converted to DFTs by taking equally spaced samples around the unit circle of the z -plane. Taking all of the pairwise products $H_i(u)H_k(v)$ between (31)–(35) then yields a separable 2D DWT implemented as a 25-channel parallel filterbank that is magnitude frequency selective but not orientation selective. However, it is shown in [61] that if the 1D analysis and synthesis filters satisfy the requirements for a perfect reconstruction orthogonal or biorthogonal DWT (see, e.g., [63]), then the DFT array

$H_i(u)H_k(v)$ will admit zeros along the line $u=0$ whenever H_i is different from the low-pass filter (31) and along the line $v=0$ whenever H_k is different from (31). These lines of zeros lie between the mainlobes of the frequency response and make it possible to decompose each nonorientation selective channel into a pair of channels that are orientation selective. An example based on the length-six orthogonal Daubechies wavelet is shown in Fig. 9. The log-magnitude DFT spectrum of the original 2D separable channel $H_3(u)H_3(v)$ is shown in Fig. 9A, while the decomposition into two orientation selective channels is given by the log-magnitude spectra of Fig. 9B and 9C. It should be noted that these orientation selective channels are distinctly nonseparable and cannot be implemented through the usual iteration of 1D convolutions along the rows and columns of the image.

For our example 25-channel separable DWT perfect reconstruction parallel filterbank based on (31)–(35), this procedure can be applied to each of the 16 channels $H_i(u)H_k(v)$ for $i, j \in [1, 4]$ to obtain 32 nonseparable orientation selective channels. While the procedure could also be applied to the eight channels $H_0(u)H_k(v)$ and $H_k(u)H_0(v)$ for $k \in [1, 4]$, it is unnecessary to do so because these channels each have a pair of mainlobes aligned either along the u -axis or along the v -axis and are hence already orientation selective. Finally, there is the nondirectional low-pass channel $H_0(u)H_0(v)$ to which the procedure cannot be applied. This results overall in a nonseparable orientation selective filterbank with $32 + 8 + 1 = 41$ channels. If the 41 wavelet subbands produced by this analysis filterbank are input to the corresponding synthesis filterbank, then the original image is recovered. However, what is most important for our purposes is the fact that *the individual image components corresponding to each subband* can be reconstructed at full resolution by submitting only that subband to the synthesis filterbank with all other inputs set to zero. The individual components obtained in this way can be summed to obtain a perfect reconstruction of the original image. It is most significant that this approach provides us with a means of gaining direct access to the individual components $f_k(n_1, n_2)$

in (21) and (22), which can be demodulated using unfiltered techniques.

What has been lost is the optimal joint localization provided by the continuous Gabor filters. But that is not too great a concern because the orientation selective parallel 2D DWT filterbanks generated by several of the most popular wavelets actually have quite good joint localization properties. What is needed is a means of quantifying the joint spatirospectral localization of a pair of discrete 2D analysis and synthesis perfect reconstruction parallel DWT filterbanks. This is provided by the novel discrete-discrete uncertainty measures given in [48, 61]. Over a reasonably large class of well-known length-six orthogonal and biorthogonal wavelets, these measures are optimized by the Coiflet, use of which will be assumed in Section 4.2.

4.2 Perfect Reconstruction FM Algorithm

The perfect reconstruction filterbank described in Section 4.1 provides us with direct access to the individual image components $f_k(n_1, n_2)$ in (21), which can be demodulated using either the TKEO-based techniques given in Section 2.1

or the analytic image-based techniques given in Section 2.2. However, when these techniques are applied, one finds that the approximation errors inherent in the demodulation algorithms themselves preclude the possibility of perfect reconstruction. This implies that more accurate demodulation algorithms are needed to realize the modulation domain signal processing paradigm depicted in Fig. 6.

In view of the fact that the approximation errors in the ESA associated with the continuous-time TKEO are generally smaller than those of the discrete ESA, it was proposed in [64] to fit the discrete signal with the cubic splines described in [65–67]. This provides a continuous-time piecewise polynomial interpolation of the signal that ensures continuous second-order derivatives. In [64], the continuous-time ESA was then applied analytically to this interpolating function. The approximation errors in the demodulation algorithms were appreciably reduced, but not eliminated, and the possibility of perfect reconstruction from the computed modulations was therefore still precluded.

Here, an advantage can be gained by using the analytic image-based approach described in Section 2.2. For each of the individual image components delivered by the filterbank



FIGURE 10 Perfect reconstruction example using wavelet-based filterbank and spline-based demodulation. **A:** Original *Lena* image. **B, C:** Two of the individual AM-FM, [amplitude modulation (AM) frequency modulation (FM)] components making up the model. **D:** Reconstruction.

of Section 4.1, the discrete 2D Hilbert transform (15) can be applied to generate a complex extension. For each component, (16) can then be applied to obtain error-free estimates of the AM function. For the component $f_k(n_1, n_2)$ with complex extension $z_k(n_1, n_2)$, error-free samples of the phase are also available via $\varphi_k(n_1, n_2) = \arctan\{\text{Im}[z_k(n_1, n_2)]/\text{Re}[z_k(n_1, n_2)]\}$. The discrete phase can then be fit with third-order 2D tensor product splines to yield a continuously differentiable polynomial interpolation [65–67]. The discrete FM function is obtained by differentiating this model analytically and sampling the continuous-space derivatives. This provides an error-free means of calculating the discrete FM function, from which the discrete phase can be perfectly reconstructed. This is true regardless of whether phase unwrapping is performed prior to the spline fit. However, if phase unwrapping is not performed, then the computed FM function will contain high frequencies that arise purely from the branch cuts in the arctan function and have no visual significance whatsoever. Therefore, phase unwrapping should always be performed prior to fitting the tensor product splines.

The filterbank of Section 4.1 provides a decomposition of the image into components with perfect reconstruction. Once the Hilbert transform (15) is applied to generate complex extensions of these components, the algorithm (16) delivers error-free estimates of the AM functions. By applying cubic tensor splines, we obtain error-free estimates of the FM functions with perfect reconstruction. Together, these techniques constitute a perfect reconstruction AM-FM transform. Figure 10A shows the original *Lena* image. A 41-component AM-FM model was computed using the Coiflet-based filterbank described in Section 4.1 and the spline-based demodulation algorithm described in this section. Two of the individual AM-FM image components are shown in Fig. 10B and 10C, where the orientation selectivity of the filterbank can clearly be seen. The reconstructed image, which is identical to the original, is given in Fig. 10D.

5 Conclusion

We reviewed the fundamentals of AM-FM image modeling and discussed practical demodulation algorithms. The DCA and CCA techniques discussed in Section 3 use jointly localized Gabor filterbanks to deliver high-quality AM-FM models that have proven useful in a wide range of applications including texture segmentation and classification, edge detection and image enhancement, estimating 3D shape from texture, texture-based computational stereopsis, and content-based image retrieval. In applications where Gabor filter magnitude responses or center frequencies have traditionally been used as features, we have found that a significant performance gain can generally be achieved by replacing these features with explicitly computed amplitude and frequency modulations. We also introduced new and exciting

techniques for computing perfect reconstruction AM-FM image transforms that set the stage for the development of AM-FM image coding schemes and open the door to the development of general theories of modulation domain image processing.

Acknowledgment

The authors wish to acknowledge the contributions of Roy Sivley of the University of Oklahoma, who designed and implemented the splines-based demodulation and reconstruction algorithms described in Section 4.2.

References

- [1] A. C. Bovik, J. P. Havlicek, D. S. Harding, and M. D. Desai, “Limits on discrete modulated signals,” *IEEE Trans. Signal Process.* **45**, 867–879 (1997).
- [2] P. Maragos, J. F. Kaiser, and T. F. Quatieri, “Energy separation in signal modulations with applications to speech analysis,” *IEEE Trans. Signal Process.* **41**, 3024–3051 (1993).
- [3] 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).
- [4] 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).
- [5] S. Lu and P. C. Doerschuk, “Nonlinear modeling and processing of speech based on sums of AM-FM formant models,” *IEEE Trans. Signal Proc.* **44**, 773–782 (1996).
- [6] J. Foote, D. Mashao, and H. Silverman, “Stop classification using DESA-1 high-resolution formant tracking,” in *Proc. IEEE Intl. Conf. Acoust. Speech, Signal Proc.* 720–723 (1993).
- [7] H. M. Hanson, P. Maragos, and A. Potamianos, “A system for finding speech formants and modulations via energy separation,” *IEEE Trans. Speech Audio Process.* **2**, 436–443 (1994).
- [8] D. A. Cairns and J. H. L. Hansen, “Nonlinear analysis and classification of speech under stressed conditions,” *J. Acoust. Soc. Amer.* **96**, 3392–3400 (1994).
- [9] G. S. Ying, C. D. Mitchell, and L. H. Jamieson, “Endpoint detection of isolated utterances based on a modified Teager energy measurement,” in *Proc. IEEE Intl. Conf. Acoust. Speech, Signal Proc.* 732–735 (1993).
- [10] P. Maragos and A. C. Bovik, “Image demodulation using multidimensional energy separation,” *J. Opt. Soc. Amer. A* **12**, 1867–1876 (1995).
- [11] B. Friedlander and J. M. Francos, “An estimation algorithm for 2-D polynomial phase signals,” *IEEE Trans. Image Process.* **5**, 1084–1087 (1996).
- [12] A. C. Bovik, M. Clark, and W. S. Geisler, “Multichannel texture analysis using localized spatial filters,” *IEEE Trans. Pattern Anal. Machine Intell.* **12**, 55–73 (1990).
- [13] J. P. Havlicek, D. S. Harding, and A. C. Bovik, “Multi-dimensional quasi-eigenfunction approximations and multi-component AM-FM models,” *IEEE Trans. Image Process.* **9**, 227–242 (2000).

- [14] A. C. Bovik, N. Gopal, T. Emmoth, and A. Restrepo, "Localized measurement of emergent image frequencies by Gabor wavelets," *IEEE Trans. Info. Theory* **38**, 691–712 (1992).
- [15] J. P. Havlicek, "The evolution of modern texture processing," *Elektrik, Turkish Journal of Electrical Engineering and Computer Sciences* **5**, 1–28 (1997).
- [16] T. B. Yap, J. Havlicek, and V. DeBrunner, "Bayesian segmentation of AM-FM texture images," in *Proc. 35th IEEE Asilomar Conf. Signals Syst. Comput.* (Pacific Grove, CA, Nov. 4–7, 2001), pp. 1156–1160.
- [17] N. Zray, J. Havlicek, S. Acton, and M. Pattichis, "Active contour segmentation guided by AM-FM dominant component analysis," in *Proc. IEEE Intl. Conf. Image Proc.* (Thessaloniki, Greece, Oct. 7–10, 2001), pp. 78–81.
- [18] T. B. Yap, T. Tangsuksorn, P. C. Tay, N. D. Mamuya, and J. P. Havlicek, "Unsupervised texture segmentation using dominant image modulations," in *Proc. 34th IEEE Asilomar Conf. Signals, Syst., Comput.* (Pacific Grove, CA, Oct. 29–31, 2000), pp. 911–915.
- [19] S. K. Mitra, S. Thurnhofer, M. Lightstone, and N. Strobel, "Two-dimensional Teager operators and their image processing applications," in *Proc. 1995 IEEE Workshop Nonlin. Signal and Image Proc.* (Neos Marmaras, Halkidiki, Greece, Jun. 20–22, 1995), pp. 959–962.
- [20] B. J. Super and A. C. Bovik, "Shape from texture using local spectral moments," *IEEE Trans. Pattern Anal. Machine Intell.* **17**, 333–343 (1995).
- [21] B. J. Super and A. C. Bovik, "Planar surface orientation from texture spatial frequencies," *Pattern Recog.* **28**, 728–743 (1995).
- [22] T. -Y. Chen, A. C. Bovik, and L. K. Cormack, "Stereoscopic ranging by matching image modulations," *IEEE Trans. Image Process.* **8**, 785–797 (1999).
- [23] M. S. Pattichis, G. Panayi, A. C. Bovik, and S. -P. Hsu, "Fingerprint classification using an AM-FM model," *IEEE Trans. Image Process.* **10**, 951–954 (2001).
- [24] N. Kitiyanan and J. P. Havlicek, "Modulation domain reference point detection for fingerprint recognition," in *Proc. IEEE Southwest Symp. Image Anal., Interp.* (Lake Tahoe, NV, Mar. 28–30, 2004), pp. 147–151.
- [25] J. P. Havlicek, J. Tang, S. T. Acton, R. Antonucci, and F. N. Ouandji, "Modulation domain texture retrieval for CBIR in digital libraries," in *Proc. 35th IEEE Asilomar Conf. Signals, Syst., Comput.* (Pacific Grove, CA, Nov. 9–12, 2003), pp. 1580–1584.
- [26] S. T. Acton, D. P. Mukherjee, J. P. Havlicek, and A. C. Bovik, "Oriented texture completion by AM-FM reaction-diffusion," *IEEE Trans. Image Process.* **10**, 885–896 (2001).
- [27] J. Shekel, "Instantaneous' frequency," *Proc. Inst. Radio Eng.* **41**, 548 (1958).
- [28] D. E. Vakman, "Do we know what are the instantaneous frequency and instantaneous amplitude of a signal?," *Trans. Radio Eng., Electron. Phys.* **21**, 95–100 (1976).
- [29] D. Vakman, "On the analytic signal, the Teager-Kaiser energy algorithm, and other methods for defining amplitude and frequency," *IEEE Trans. Signal Proc.* **44**, 791–797 (1996).
- [30] L. Mandel, "Interpretation of instantaneous frequencies," *Am. J. Phys.* **42**, 840–846 (1974).
- [31] M. S. Gupta, "Definition of instantaneous frequency and frequency measurability," *Am. J. Phys.* **43**, 1087–1088 (1975).
- [32] A. Potamianos and P. Maragos, "A comparison of the energy operator and the Hilbert transform approach to signal and speech demodulation," *Signal Proc.* **37**, 95–120 (1994).
- [33] A. C. Bovik and P. Maragos, "Conditions for positivity of an energy operator," *IEEE Trans. Signal Process.* **42**, 469–471 (1994).
- [34] J. Ville, "Théorie et applications de la notation de signal Analytique," *Cables et Transmission*, vol. 2A, pp. 61–74 (1948), translated from the French in I. Selin, "Theory and applications of the notion of complex signal," Tech. Rept. T-92, The RAND Corporation, Santa Monica, CA, August 1958.
- [35] A. Papoulis, *Signal Analysis* (McGraw-Hill, New York, 1977).
- [36] L. Cohen, *Time-Frequency Analysis* (Prentice Hall, Englewood Cliffs, NJ, 1995).
- [37] E. M. Stein, *Singular Integrals and Differentiability Properties of Functions* (Princeton University Press, Princeton, NJ, 1970).
- [38] R. R. Read and S. Treitel, "The stabilization of two-dimensional recursive filters via the discrete Hilbert transform," *IEEE Trans. Geosci. Electron.* **11**, 153–160 (1973).
- [39] F. Peyrin, Y. M. Zhu, and R. Goutte, "Extension of the notion of analytic signal for multidimensional signals. Application to images," in *Signal Processing III: Theories and Applications*, I. T. Young et al., eds, (Elsevier Science Publishers, Amsterdam, B. V. (North-Holland), 1986), pp. 677–680.
- [40] Y. M. Zhu, F. Peyrin, and R. Goutte, "The use of a two-dimensional Hilbert transform for Wigner analysis of 2-dimensional real signals," *Signal Process.* **19**, 205–220 (1990).
- [41] S. L. Hahn, "Multidimensional complex signals with single-orthant spectra," *Proc. IEEE* **80**, 1287–1300 (1992).
- [42] J. P. Havlicek, J. W. Havlicek, and A. C. Bovik, "The analytic image," in *Proc. IEEE Intl. Conf. Image Proc.* (Santa Barbara, CA, October 26–29 1997), 446–449.
- [43] J. P. Havlicek, J. W. Havlicek, N. D. Mamuya, and A. C. Bovik, "Skewed 2D Hilbert transforms and computed AM-FM models," in *Proc. IEEE Intl. Conf. Image Proc.* (Chicago, IL, Oct. 4–7, 1998), pp. 602–606.
- [44] D. Wei and A. C. Bovik, "On the instantaneous frequencies of Multicomponent AM-FM Signals," *IEEE Signal Proc. Let.* **5**, 84–86 (1998).
- [45] J. P. Havlicek, D. S. Harding, and A. C. Bovik, "Multi-component multidimensional signals," *Multidimensional Syst. Signal Proces.* **9**, 391–398 (1998).
- [46] D. Gabor, "Theory of communication," *J. Inst. Elect. Eng. London* **93**, 429–457 (1946).
- [47] J. G. 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).
- [48] P. C. Tay, J. P. Havlicek, and V. DeBrunner, "A novel translation and modulation invariant discrete-discrete uncertainty measure," *Proc. IEEE Intl. Conf. Acoust. Speech Signal Proc.* (Orlando, FL, May 13–17, 2002), pp. 1461–1464.
- [49] J. P. Havlicek, A. C. Bovik, and D. Chen, "AM-FM image modeling and Gabor analysis," in *Visual Information Representation, Communication, and Image Processing*, C. W.

- Chen and Y. Zhang, eds. (Marcel Dekker, New York, 1999), pp. 343–385.
- [50] P. Brodatz, *Textures: a Photographic Album for Artists and Designers* (Dover, New York, 1966).
- [51] <http://vismod.media.mit.edu/vismod/imagery/VisionTexture/vistex.html>, visited June, 2004.
- [52] T. Tangsukson and J. P. Havlicek, “AM-FM image segmentation,” in *Proc. IEEE Intl. Conf. Image Proc.* (Vancouver, Canada, September 10–13, 2000), 104–107.
- [53] J. P. Havlicek, D. S. Harding, and A. C. Bovik, “The multi-component AM-FM image representation,” *IEEE Trans. Image Proc.* **5**, 1094–1100 (1996).
- [54] A. C. Bovik, “Analysis of multichannel narrow-band filters for image texture segmentation,” *IEEE Trans. Signal Proc.* **39**, 2025–2043 (1991).
- [55] J. P. Havlicek, D. S. Harding, and N. D. Mamuya and A. C. Bovik, “Wideband frequency excursions in computed AM-FM image models,” in *Proc. IEEE Southwest Symp. Image Anal. Interp.* (Tucson, AZ, Apr. 5–7, 1998), pp. 211–216.
- [56] I. Daubechies, “Orthonormal bases of compactly supported wavelets,” *Commun. Pure Appl. Math.* **51**, 909–996 (1988).
- [57] G. Strang and T. Nguyen, *Wavelets and Filter Banks* (Wellesley-Cambridge Press, Wellesley, MA, 1996).
- [58] M. Vetterli and J. Kovačević, *Wavelets and Subband Coding* (Prentice Hall, Englewood Cliffs, NJ, 1995).
- [59] S. G. Mallat, “Multifrequency channel decompositions of images and wavelet models,” *IEEE Trans. Acoust. Speech Signal Proc.* **37**, 2091–2110 (1989).
- [60] P. C. Tay and J. P. Havlicek, “Frequency implementation of discrete wavelet transforms,” in *Proc. IEEE Southwest Symp. Image Anal., Interp.* (Lake Tahoe, NV, Mar. 28–30, 2004), pp. 167–171.
- [61] P. C. Tay, *An Optimally Well Localized Multi-Channel Parallel Perfect Reconstruction Filter Bank*, Ph.D. thesis, The University of Oklahoma, Dec. 2003.
- [62] P. P. Vaidyanathan, “Multirate digital filters, filter banks, polyphase networks, and applications: A tutorial,” *Proc. IEEE* **78**, 56–93 (1990).
- [63] M. Vetterli and C. Herley, “Wavelets and filter banks: theory and design,” *IEEE Trans. Signal in Process.* **40**, 2207–2232 (1992).
- [64] D. Dimitriadis, and P. Maragos, “An improved energy demodulation algorithm using splines,” in *Proc. IEEE Intl. Conf. Acoust. Speech Signal Proc.* (Salt Lake City, UT, May 7–11, 2001), pp. 3481–3484.
- [65] M. Unser, A. Aldroubi, and M. Eden, “B-spline signal processing: Part I-theory,” *IEEE Trans. Signal Proc.* **41**, 821–833 (1993).
- [66] M. Unser, A. Aldroubi, and M. Eden, “B-spline signal processing: Part II-efficient design and applications,” *IEEE Trans. Signal Process.* **41**, 834–848 (1993).
- [67] M. Unser, “Splines: A perfect fit for signal and image processing,” *IEEE Signal Process. Mag.* **16**, 22–38 (1999).