The Evolution of Modern Texture Processing

Joseph P. Havlicek

School of Electrical and Computer Engineering

The University of Oklahoma

202 West Boyd, Room 219

Norman, OK 73019-0631 USA

Tel: (405) 325-4279 Fax: (405) 325-7066 E-mail: joebob@tobasco.ecn.ou.edu

Abstract – This paper studies the evolution of image texture processing techniques over the last 20 years. Although texture is a fundamental attribute of images that has been shown to play an important role in human visual perception, the quantification and characterization of texture is difficult. Early texture processing techniques described texture deterministically or statistically in terms of repeated gray level patterns and the structure of the spatial placement of these patterns. Gray level cooccurrence matrices were among the most successful such methods. Modern texture processing techniques tend to characterize texture in terms of spatio-spectrally localized coherent amplitude, frequency, and phase modulations. This paper argues that evolution of the modern methods from the early methods can be directly linked to advances in the understanding of mammalian biological visual function that occurred in the fields of psychophysics and physiology, and furthermore that the most successful modern methods have evolved to emulate biological vision systems. Evolution of modern texture processing methods is examined, and several of the most successful new techniques such as the multidimensional Teager-Kaiser operator and AM-FM modeling techniques are described in some detail. The use of computed dominant modulations to perform effective texture segmentation is demonstrated for the first time.

This research was supported in part by the Army Research Office under contract DAAH 049510494 and by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under grant number F49620-93-1-0307.

Author preprint. Submitted to Elektrik,
Turkish Journal of Electrical Engineering and Computer Sciences.

1 Introduction

The characterization, the analysis, and the representation of projected surface textures in images have long been recognized as fundamental problems in image processing and machine vision [1–5]. Texture processing may be used to segment images into objects or regions of homogeneous texture, to classify or recognize surface materials from their projections in images, and to infer three-dimensional surface shape from images. Yet, despite the facts that texture is clearly an intrinsic property of images, has been clearly shown to play a fundamental role in human vision at the preattentive level, and is demonstrably useful in image processing and machine vision algorithms, the quantification and characterization of texture has proven elusive. Indeed, no well posed definition of image texture even exists.

The earliest successful texture processing techniques sought to characterize texture (whether deterministically or statistically) in terms of repeated *primitives*, or structured groupings of gray levels, as well as in terms of the spatial arrangements of those primitives. In a comprehensive 1979 survey, Haralick reviewed what I shall refer to as *early* texture processing techniques [6]; a later survey by Van Gool, Dewaele, and Oosterlinck reviewed essentially the same methods [7].

In contrast, many modern texture processing techniques characterize and quantify texture in terms of localized, coherent groupings of instantaneous spatial frequencies or, more generally, in terms of nonstationary amplitude, frequency, and phase modulations occurring in frequency and orientation selective channels. In this paper, I will review, compare, and contrast the early methods and the modern methods of texture processing. I will discuss the evolution of texture processing with relation to profound discoveries about mammalian biological visual function that occurred in the fields of psychophysics and physiology, and I will argue that modern texture processing techniques have evolved to emulate biological vision systems to the extent possible. Some of the most recent innovations in texture processing such as the multidimensional Teager-Kaiser operator and AM-FM modeling techniques will be examined in some detail, and

effective texture segmentation using computed dominant modulations will be demonstrated for the first time.

2 Early Approaches to Texture

Statistical characterization of the spatial arrangement of texture primitives dates back at least to Julesz, who in his later work used the term *textons* to describe the primitives [8–11]. Typical textons included line segments, oriented blobs, line crossings, and the terminations of blobs and line segments. In a series of psychophysical experiments spanning more than 20 years, Julesz discovered that the human visual system is capable of *preattentively* discriminating certain textures. The culmination of this work was the assertion that the eye cannot preattentively discriminate between textures that have the same second-order statistics on a spatially local basis (but that it can preattentively discriminate between textures with the same first-order and global second-order statistics, provided that the local second-order statistics differ).

Motivated in part by what Julesz was discovering about biological vision systems, particularly that the human visual system is perceptually sensitive to second-order statistics, numerous investigators set out to develop machine vision texture processing algorithms based on the statistical approach. Perhaps the most successful of these were based on *gray level cooccurrence matrices*, which were introduced and systematically developed by Haralick [6]. Each cooccurrence matrix was associated with a distance d and an orientation θ. The matrix entries $P_{i,j}$ corresponded to the number of times that a pixel with gray level i appeared in the image displaced by a distance d and orientation θ from a pixel with gray level j. *Features*, such as angular second moment, contrast, correlation, and entropy, were then defined in terms of algebraic operations on the cooccurrence matrix entries and used to characterize, classify, and discriminate textures, typically using a linear classifier. Haralick and others obtained good classification and discrimination results using this approach [6]. Zucker and Terzopoulos

developed a statistical method for determining the best values of d and θ, as well as the best features to use for classification of textures belonging to fixed training classes [12].

A generalization of the gray level cooccurrence matrix, motivated at least partially by Hubel and Wiesel's discovery that certain cortical cells responded strongly and selectively to oriented, elongated bars and edges [13–15], was the generalized cooccurrence matrix. Generalized cooccurrence matrices were analogous to gray level cooccurrence matrices, but described the spatial distributions of textons that were more sophisticated than simple pixel gray levels [6, 7]. Hence, they were correspondingly more difficult to define, more computationally intensive, and in the absence of a complete texton theory for human visual perception, more *ad hoc*. Other methods related to the gray level cooccurrence matrix and generally deemed inferior included the *gray level difference matrix* and the *gray level run length matrix* [6, 7].

An approach using two-dimensional autoregressive models to describe texture originated in the computer graphics and texture synthesis fields, where it was found that synthetic textures of reasonably high quality could be created by estimating the model parameters for a real texture, and then iterating the model from a random initial state [6, 7]. Attempts to use the autoregressive model parameters as texture features for classification and discrimination achieved only limited success, particularly for images containing *macrotextures* characterized by structural placement rules consistent over large numbers of pixels.

Several early texture analysis techniques employed spatial frequency-based texture features computed from the autocorrelation function of the image, the Fourier power spectrum of the image, or Fourier spectra computed on blocks within the image [6, 7]. The shape of the autocorrelation function indicates the *coarseness* of the texture, whereas the *orientation* of the texture is manifest in directional dependencies in the autocorrelation function. Typical features used in the Fourier spectrum techniques included the amount of power present in annular or ring-, wedge-, and slit-shaped regions of the frequency plane. Bajcsy and Lieberman measured texture gradients from

Fourier spectra computed over square regions and used them to infer three-dimensional shape from texture [16]. At the time of the survey papers [6, 7], *viz.* 1983, spatial frequency-based techniques were generally considered to have inherent problems and to be inferior to the cooccurrence matrix methods for classifying and discriminating texture. Closely related to the spatial frequency-based methods were the so called *edginess* methods, which attempted to characterize textures by estimating the number of edges present per unit area at various scales using gradient, derivative, or Laplacian convolution kernels [6, 7].

Purely structural methods sought to describe texture in terms of deterministic rules for the spatial placement of texture primitives. The placement rules were described formally by shape, array, and tree grammars [17, 18]. The choices of what primitives to use and what their scales of definition should be were image dependent and *ad hoc*, and hence these methods were rarely used for texture analysis, classification, and discrimination in practice [6]. Morphological techniques for the structural characterization of texture were also investigated [6, 19]

The salient characteristic of the *early* methods described in this section is that they were all *ad hoc*. Some of them were effective for characterizing and discriminating textures in certain classes of images, but none of them worked well for general images. The characterization and discrimination of textures within a given limited class of images required the design of features and a corresponding classifier that were specialized to that class.

3 Psychophysical and Physiological Advances

Between 1968 and 1985, a series of advances in the understanding and modeling of biological visual function occurred in psychophysics and physiology; these advances had a profound impact on subsequent texture research in image processing and machine vision. Campbell and Robson determined that certain aspects of human visual function

could be explained by the existence of independent linear channels sensitive to narrow ranges of spatial frequencies [20]. This multiple-channels model was corroborated by Graham and Nachmias [21]. For a striking perceptual demonstration of the channels in operation, see [22, p. 144]. Blakemore and Campbell established that the channels were orientation selective as well as frequency selective [23]. Based on psychophysical experiments involving one-dimensional textures, Richards and Polit argued that the texture discrimination capability of the human visual system could be explained by the existence of only four distinct spatial frequency channels (here, *spatial frequency* refers to the *magnitude* of the multidimensional frequency vector) [24]. From similar experiments, Caelli and Bevan suggested that the maximum number of orientation channels is 18 [25].

Numerous ensuing studies characterized cortical cell responses as spatial frequency filters and advocated the modeling of cortical visual function as a Fourier frequency analysis [26–34]. Robson suggested that the Fourier decomposition occurs independently on equi-sized spatial patches, rather than globally on the entire retinal image [35]. Typically, the experiments in these studies employed sinusoidal gratings of one form or another as stimuli, and hence the results characterized cortical receptive fields as functions of spatial frequency. The corresponding spatial sensitivity profiles were obtained via inverse Fourier transforms, subject to the assumption that the spatial filtering being performed was *linear*. While there are certainly nonlinear aspects to the function of complex cells, and even simple cells, the validity of linear models for the spatial filtering stages of both simple and complex cells has been reasonably established [36–41]. Various functional models, including the *difference of Gaussians* (DoG) and *Laplacian of Gaussian* (LoG, or $\nabla^2 g$) [1, 42], were popular from early on for describing the spatial receptive fields of retinal, lateral geniculate nucleus (LGN), and cortical cells.

Both even- and odd-symmetric spatial receptive fields were observed experimentally, and in 1980 Marčelja [43] made the connection between these receptive fields

and the one-dimensional elementary functions described by Gabor in 1946 [44]. The so called *Gabor functions* take the form $f(x) = G(x)e^{j\omega x}$ in one dimension and $f(x, y) = G(x, y)e^{j(ux+vy)}$ in two dimensions, where $G(\cdot)$ is Gaussian. Hence, they are Gaussians in the frequency domain and Gaussian envelopes frequency-modulated by complex sinusoids in the time and space domains. The real and imaginary parts of $f(\cdot)$ differ in phase by $\pi/2$. For appropriate choices of $G(\cdot)$ they are even- and odd-symmetric, respectively. In 1981 Pollen and Ronner recorded responses from pairs of adjacent simple visual cortical cells in cat, and found that often the *phases* of the responses within each pair did indeed differ by $\pi/2$ [45]. Furthermore they found that, within pairs, it was often the case that the receptive field of one cell exhibited even symmetry while the other exhibited odd symmetry.

The Gabor functions admit many attractive properties. Gabor proved an uncertainty relation which limits the degree to which any one-dimensional function can be localized simultaneously in both time and frequency [44]. He showed that his functions achieved the lower bound on conjoint resolution and were unique in this respect; hence he proposed analyzing arbitrary functions in terms of linear combinations of Gabor functions to implement an optimally tempero-spectrally localized nonstationary analysis. An analytical technique for determining the coefficients in such a *Gabor expansion* was developed by Basstiaans [46–48].

MacKay and Daugman both advocated the Gabor receptive field model for visual cortical cells, and developed the two-dimensional version of Gabor's uncertainty relation [49, 50]. The lower bound on conjoint spatio-spectral localization is uniquely realized by the two-dimensional Gabor functions. Research by numerous investigators supported the Gabor function receptive field model, and it became widely accepted [41, 50–57].

Objections to the Gabor function receptive field model still remain [58]. Some investigators have argued that other functional forms such as the *difference of three Gaussians* may fit certain experimental data better than Gabor functions [42]. However, an

extensive and rigorous study by Jones, Stepnoski, and Palmer measuring the receptive fields of simple cortical cells in the striate cortex of cat, both in the two-dimensional spatial and spatial frequency domains, found the deviations from the Gabor model to be "...devoid of spatial structure and statistically indistinguishable from random noise" [41, 56, 57].

4 Modern Approaches to Texture

In the search for machine vision algorithms to characterize image textures and perform segmentation and shape analysis based on texture, a recurrent theme has been the emulation, to varying degrees, of biological vision systems. The latter constitute the only known vision systems that work well *in general*. Of the techniques described in Section 2, the gray level cooccurrence matrix methods were perhaps most successful. These were motivated by the knowledge that biological vision systems somehow utilize local second-order spatial statistics to discriminate textures. However, since nothing specific is known about *how* this statistical information is used in biological vision systems, the analogy between cooccurrence matrices and biological visual function is only a weak one.

Consider using gray level cooccurrence matrices to design an algorithm for discriminating textures in 256×256 gray scale images having eight-bit pixels. If we consider only the eight orientations $n\pi/16$, $n \in \{0, 1, \ldots, 7\}$, and compute second-order statistics over distances ranging from one to 16 pixels, then the region about every pixel in an image will be described by 128 gray level cooccurrence matrices each having $32,896$ unique entries. With this approach, texture discrimination features at each pixel will be defined on a set of $4,210,688$ numbers. Since we do not *know* how biological vision systems use the information represented by this large set of numbers, the design of the features employed by our discriminator will *not* be biologically motivated.

In view of the research described in Section 3, let us alternatively consider the

design of a texture discriminator based on decomposing images with a multiband bank of linear filters. Since real-valued images have conjugate symmetric spectra, only half of the possible orientations need be considered explicitly. Thus, the results of Richards and Polit [24] together with those of Caelli and Bevan [25] suggest that a bank of *four* filters at each of *nine* orientations, or 36 filters in total, should suffice. In general, the filters may be complex-valued. So, with this approach we seek to discriminate textures using 36 complex-valued response images. Since we do not *know* how the information represented by these response images is used in biological vision systems, we still cannot design a biologically motivated discriminator. However, in this case each pixel in the image is described by only 36 complex-valued numbers. This represents a *fantastic* simplification compared with the cooccurrence matrix approach just discussed: the size of the set of numbers to be considered in designing features has been reduced by a factor of approximately *58,481*. In a like spirit, the texture processing and analysis techniques I will review in this section made use of the advances described in Section 3. I will refer to them as *modern* approaches to differentiate them from the approaches described in Section 2.

In 1972, Stockham advocated the design of image processors based on the human visual system and recommended the use of Gaussian frequency characteristics [59]. Primarily motivated by texton theories of Julesz and Bergen [11], Ikonomopoulos and Unser used a complete, orthogonal bank of directional filters with radial slit-shaped spectra to discriminate textures [60]. They obtained reasonable rates of correct classification for images made up of juxtapositions of Brodatz textures [61]. Using features defined on the responses of isotropic, annularly-shaped frequency filters and Gaussian orientation filters, Coggins and Jain obtained a correct classification rate of 98% for a set of 200 Brodatz textures [62]. Rao and Schunck [63] and Rao and Jain [64] used the responses of $\nabla^2 g$ filters to estimate the local coherency and dominant local orientation at each pixel in an image and used these quantities to perform texture segmentation.

In 1986, Turner analyzed textured images with a bank of 16 Gabor filters arranged

in a polar tesselation with four circularly symmetric filters on each of four oriented rays [65]. He argued that effective texture segmentation could be performed using the responses of these filters. Malik and Perona formulated an approximate computational model for biological early visual function, and used it for texture segmentation [66]. They modeled area V1 simple cells as DoG filters followed by half-wave rectifiers, and also incorporated a model of nonlinear intracortical inhibition into their paradigm. In the final stage of their approach, they computed texture gradients from the channel filter responses and identified texture boundaries as local peaks in the magnitude of the texture gradient.

Extending the work of Basstiaans [46–48], Zeevi and Porat developed a formulation for the coefficients in a *two-dimensional* Gabor expansion applicable to images [67]. The Gabor expansion coefficients bear strong analogy to spatially subsampled Gabor filter response images, in that they reflect the degree to which an image correlates with various Gabor functions on a spatio-spectrally local basis. Porat and Zeevi demonstrated high quality reconstructions of textured images using Gabor expansions truncated to a few thousand terms, and used simple statistics of the Gabor coefficients as features to effectively segment juxtaposed Brodatz textures [68, 69]. In a similar scheme, Wilson and Spann performed texture segmentation using feature sets defined on the responses of a multiband bank of finite prolate spheroidal filters, which have properties similar to those of the Gabor filters [70].

In 1986, Bovik, Clark, and Geisler proposed an interpretation of image texture as a *carrier* of region information, which could be *demodulated* [71–73]. They modeled a given texture according to

$$t(x, y) = a(x, y) \cos\left[2\pi(U_0 x + V_0 y) + p(x, y)\right], \tag{1}$$

where $a(x, y)$ was a slowly varying amplitude modulation and $p(x, y)$ was a slowly varying phase term. Their approach was to filter the image with a multiband bank of Gabor filters and then assign a given pixel to a texture-segmented region according to

which channel filter produced the greatest magnitude response at the pixel. With this approach, they obtained a number of excellent segmentations using channel filters with parameters that were specifically designed for each image analyzed.

In addition, they observed that visually discriminable textures arise when two regions contain textures with identical spatial frequency and amplitude modulation characteristics, but have a spatial phase displacement relative to one another. Using frequency-modulated *derivative of Gaussian* and $\nabla^2 g$ filters in concert with the Gabor channel filters, they obtained the Laplacians of the channel filter response phases. Significant zero crossings of these Laplacians correspond to significant phase discontinuities in the texture under analysis, and they used this fact to refine the texture segmentations obtained from the channel amplitudes by also segmenting textures with identical amplitude and frequency characteristics but differing phase characteristics.

A detailed analysis of the Gabor filters used in [73] was subsequently carried out by Bovik, who established that postfiltering can be used to obtain texture segmentations and three-dimensional shape information even in the presence of significant deviations from the smooth model (1) [74]. He observed that a host of factors including occlusions, surface discontinuities, deformations and defects in surface topology, surface reflectance, shadows, specularities, and noise can all give rise to perturbations in (1) that correspond to amplitude and phase modulations that are not locally smooth. In [74], Bovik also introduced the more general texture model

$$t(x, y) = A(x, y) \cos\left[\Omega(x, y)\right], \tag{2}$$

where $A(x, y)$ is a smoothly varying amplitude modulation term and $\Omega(x, y)$ is a smoothly varying phase modulation term. The components of the phase gradient vector $\nabla\Omega(x, y)$ are the instantaneous image frequencies. Bovik briefly addressed estimation of $\nabla\Omega(x, y)$ from the Gabor filter channel responses and proposed two approximate approaches. The first was simply to estimate $\nabla\Omega(x, y)$ as the center frequency of the channel filter with maximum response amplitude at pixel (x, y). The second was to

interpolate between the center frequencies of several channel filters, weighting each by the amplitude of its response at pixel (x, y).

4.1 The Importance of Modulation and Frequency Estimation

The usefulness of term $A(x, y)$ in (2), which is interpreted as *amplitude modulation* in a bandpass filtered texture, has been established in the foregoing discussion. For example, Malik and Perona performed bandpass filtering of images, and then applied half-wave rectification and low-pass filtering to the bandpass responses [66]. This procedure yields an approximation to $A(x, y)$. Bovik, Clark, and Geisler [73] and Bovik [74] demonstrated that excellent texture segmentations could be obtained based solely on the amplitude envelopes of appropriately designed bandpass filters. Of the studies reviewed in Section 3 that advocated Gabor function receptive field models, most did so based on evidence that biological vision systems somehow detect the presence of and make use of narrowband concentrations of *spatially local* frequencies; spatial variations in the energies of such concentrated frequency bundles correspond to amplitude modulations $A(x, y)$.

Phase and frequency modulation information is also generally useful. Witkin [75] and Kass and Witkin [76] used demodulated filterbank channel responses to compute shape and orientation from texture. Schachter used smoothly varying amplitude *and* frequency modulations to synthesize realistic looking image textures [77]. Recently, Friedlander and Francos modeled images as the real part of a complex exponential with polynomial phase and demonstrated an algorithm for estimating the order and coefficients of the polynomial [78, 79].

In this section, I will discuss *why* estimation of the instantaneous image frequencies $\nabla\Omega(x, y)$, interpreted as *frequency modulations*, is an important problem. Most, if not all of the deficiencies associated with the *early* spatial frequency-based texture analysis methods described in Section 2 can be attributed to the fact that these methods utilized *global* Fourier frequencies, or Fourier frequencies computed over fixed-size subimages.

This sheds light the apparent discrepancy between the general conclusion that spatial frequency-based methods were inferior to cooccurrence matrix-based methods [6,7] and the overwhelming psychophysical and physiological evidence that biological vision systems do indeed analyze the retinal image in terms of spatial frequency, as described in Section 3. The explanation is that the latter (biological systems) perform the frequency analysis on a conjointly *localized* basis in the joint space-frequency hyperspace. Therefore, in designing texture processing algorithms to emulate biological vision systems to some degree, it makes sense to consider the *spatially localized* instantaneous frequencies $\nabla\Omega(x,y)$ rather than the Fourier frequencies.

The extent to which biological vision systems *explicitly* use spatial frequency information in texture perception remains unknown. What is clear, however, is that visual cortical cells *do* transmit contrast and spatial phase information [26]. *Contrast* information is manifest in the amplitude modulation term $A(x,y)$ in (2), while *spatial phase* information is manifest in $\Omega(x,y)$. The design of machine vision algorithms to estimate $\Omega(x,y)$ is difficult and ill-posed as a consequence of the *phase wrapping* problem: the value of the phase at any point in an image is always ambiguous by an additive factor of $2n\pi$, $n \in \mathbb{Z}$. In a machine vision system, *representing* spatial phase in terms of estimates of the spatial frequencies $\nabla\Omega(x,y)$ circumvents this problem.

Furthermore, since the linear operations of convolution and differentiation commute, several of the texture analysis techniques discussed so far may be interpreted as making *explicit* use of demodulated instantaneous spatial frequency information. For example, the methods of Bajcsy and Lieberman [16], Rao and Schunck [63], Rao and Jain [64], and Malik and Perona [66], as well as the phase-based texture segmentation technique of Bovik, Clark, and Geisler [73] all employed gradients or filters with differentiated envelopes. The responses of the latter are equivalent to the *derivatives* of the responses of the *undifferentiated* filters. So derivatives of images or derivatives of bandpass filtered images were explicitly utilized in all of these methods. To see that these differentiated quantities approximate the instantaneous spatial frequencies, consider a

possibly bandpass filtered sinusoidal grating image $t(x, y)$ modeled according to

$$t(x, y) = A \cos [\Omega(x, y)] = A \cos [u_0 x + v_0 y] . \tag{3}$$

The partial derivatives of $t(x, y)$ are

$$t^{(x)}(x, y) = \frac{\partial}{\partial x} t(x, y) = -A u_0 \sin [u_0 x + v_0 y] \tag{4}$$

and

$$t^{(y)}(x, y) = \frac{\partial}{\partial y} t(x, y) = -A v_0 \sin [u_0 x + v_0 y] . \tag{5}$$

Now

$$\arctan \left[\frac{t^{(y)}(x, y)}{t^{(x)}(x, y)} \right] = \arctan \left[\frac{v_0}{u_0} \right] = \arg \nabla \Omega(x, y), \tag{6}$$

so the *orientation* field of the differentiated image is equal to the orientation of the instantaneous frequency vector. For the special case of an image filtered by a complex-valued Gabor filter, the real and imaginary components of the response are in quadrature phase, so the response image $t_G(x, y)$ may be modeled as

$$t_G(x, y) = A \exp [j \Omega(x, y)] = A \exp [j \, (u_0 x + v_0 y)] . \tag{7}$$

The partials of $t_G(x, y)$ are

$$t_G^{(x)}(x, y) = \frac{\partial}{\partial x} t_G(x, y) = j A u_0 \exp [j \, (u_0 x + v_0 y)] \tag{8}$$

and

$$t_G^{(y)}(x, y) = \frac{\partial}{\partial y} t_G(x, y) = j A v_0 \exp [j \, (u_0 x + v_0 y)] . \tag{9}$$

As before,

$$\arctan \left[\frac{t_G^{(y)}(x, y)}{t_G^{(x)}(x, y)} \right] = \arctan \left[\frac{v_0}{u_0} \right] = \arg \nabla \Omega(x, y). \tag{10}$$

In this case we also have

$$\left|\frac{\partial}{\partial x} t_G(x,y)\right| = |A|\ |u_0| \tag{11}$$

and

$$\left|\frac{\partial}{\partial y} t_G(x,y)\right| = |A|\ |v_0|\,, \tag{12}$$

and hence the *magnitudes* of the partials of $t_G(x,y)$ are *proportional* to the instantaneous frequencies u_0 and v_0. For slowly varying modulations $A(x,y)$ and $\nabla\Omega(x,y)$, (6) and (10)–(12) still hold approximately. Furthermore, the coefficients computed in the two-dimensional Gabor expansions of Porat and Zeevi [68] indicate the degree to which certain frequencies are present in an image on a spatially local basis, and hence are *implicitly* related to the quantity $\nabla\Omega(x,y)$. More recently, techniques for three-dimensional shape recovery from texture and for computed phase-based stereopsis making *explicit* use of spatially localized instantaneous frequency information have been developed [80–84].

4.2 Computed Modulation Models

Several techniques from the fields of time-frequency analysis and multiresolution signal processing, including the windowed Fourier transform [85, 86], the complex spectrogram (Rihaczek distribution) [87, 88], the Wigner distribution [89, 90], the ambiguity function [91, 92], Cohen class and affine time-frequency distributions [93–96], the Choi-Williams distribution [97], and the wavelet transform [98–104] could be used as bases for an instantaneous frequency estimation scheme. However, *only* an approach based on Gabor filters estimates the *spatially localized* frequencies from responses that are optimally *conjointly* localized in the joint space-frequency hyperspace. Furthermore, a Gabor filtering approach is motivated by processing that we *know* occurs in simple visual cortical cells of biological vision systems. With regards to the recently popularized wavelet transform, it is worth noting that the responses of an appropriately chosen set of Gabor filters are, in many respects, strongly analogous both to a biorthogonal wavelet

transform and to the wavelet transform of Morlet (cycle-octave transform) [105].

In 1992, Bovik [106] and Bovik, *et. al.* [107] addressed in detail the problem of estimating the modulations $A(x,y)$ and $\nabla\Omega(x,y)$ in (2) from Gabor filter channel responses. For an image modeled as a sum of components of the form (2), they estimated the quantities $A(x,y)$ and $\nabla\Omega(x,y)$ corresponding to the component that dominated the local frequency spectrum on a spatially local basis. They referred to these dominant component frequencies as *emergent*.

They considered filtering the image with a complex-valued bandpass filter of the form

$$h(\mathbf{x}) = w(\mathbf{x}) \exp\left(j\mathbf{\Omega}_0^T\mathbf{x}\right), \tag{13}$$

where $\mathbf{x} = [x\ y]^T$ is an image coordinate, $w(\mathbf{x})$ is a low-pass filter, and $\mathbf{\Omega}_0 = [u_0\ v_0]^T$ is the center frequency of the bandpass filter $h(\mathbf{x})$. The frequency response of $h(\mathbf{x})$ is

$$H(\mathbf{\Omega}) = W(\mathbf{\Omega} - \mathbf{\Omega}_0), \tag{14}$$

where $W(\mathbf{\Omega})$ is the Fourier transform of the low-pass equivalent filter $w(\mathbf{x})$.

The magnitude response of the filter (13) to an image component $t(\mathbf{x})$ of the form (2) is given *exactly* by

$$m(\mathbf{x};\mathbf{\Omega}_0) = \left|t(\mathbf{x}) * h(\mathbf{x})\right|. \tag{15}$$

In general, it is impossible to obtain a closed form expression for $m(\mathbf{x};\mathbf{\Omega}_0)$ in (15). This fact complicates the problem of estimating $A(\mathbf{x})$ and $\nabla\Omega(\mathbf{x})$ — a problem which is already ill-posed since, corresponding to any real image component, there exist uncountably infinitely many pairs of modulating functions $A(x,y)$ and $\nabla\Omega(x,y)$ for which the model (2) is exactly equal to the component.

In [106] and [107], Bovik, *et. al.*, *approximated* the magnitude response (15) using the quantity

$$\widehat{m}(\mathbf{x};\mathbf{\Omega}_0) = A(\mathbf{x})\Big|W\left[\mathbf{\Omega}_0 - \nabla\Omega(\mathbf{x})\right]\Big|. \tag{16}$$

Approximations of this type are generally useful in the analysis of linear system responses to nonstationary inputs, and are now known as *quasi-eigenfunction approximations* [107–112]. In particular, note that if $t(\mathbf{x})$ is monochromatic and takes the form (3), then $m(\mathbf{x};\boldsymbol{\Omega}_0) = \widehat{m}(\mathbf{x};\boldsymbol{\Omega}_0)$ identically. Thus, Bovik, *et. al.*, introduced the *emergent frequency constraint equation*

$$m(\mathbf{x};\boldsymbol{\Omega}_0) \approx \widehat{m}(\mathbf{x};\boldsymbol{\Omega}_0). \tag{17}$$

In [106, 107], they bounded the approximation error in (17) for an image component admitting *general* modulating functions. They sought to estimate locally coherent dominant modulations $A(\mathbf{x})$ and $\nabla\Omega(\mathbf{x})$ that would minimize the approximation error in (17) using filters $h(\mathbf{x})$ also designed to minimize the error, arguing that such solutions would at once be both the most physically meaningful and the most accurately measurable from amongst the infinite locus of possible solutions.

Minimization of the error bound required simultaneously minimizing both the generalized p-energy variances (spatial spreads) and the Sobolev norms (smoothness functionals) of the equivalent low-pass filter $w(\mathbf{x})$. In [106, 107] they gave an uncertainty relation generalizing those of Gabor [44], MacKay [49], and Daugman [50], which placed a lower bound on the degree to which the p-energy variances and Sobolev norms of $w(\mathbf{x})$ could simultaneously be minimized. They found that this lower bound was uniquely realized by the Gabor functions. Intuitively this should come as no surprise, since minimization of the error bound on (17) is equivalent to demanding filters $h(\mathbf{x})$ which are highly spatially localized to capture nonstationarities, but also highly spectrally localized to effectively separate the dominant image component from subemergent components. Consequently, for filters $h(\mathbf{x})$ they employed a bank of forty unity aspect ratio Gabor filters arranged in a polar frequency tesselation.

To compute the estimated emergent frequencies $\nabla\widehat{\Omega}(\mathbf{x})$, they substituted the Gaus-

sian

$$W(\mathbf{\Omega}) = 2\sigma\sqrt{2\pi}\exp\left[-\sigma^2\mathbf{\Omega}^T\mathbf{\Omega}\right] \tag{18}$$

into (17), and upon taking the logarithm of both sides obtained

$$\begin{aligned} M_\sigma(\mathbf{x};\mathbf{\Omega}_0) &\approx \widehat{M}_\sigma(\mathbf{x};\mathbf{\Omega}_0) \\ &= \frac{1}{\sigma^2}\log A(\mathbf{x}) + \frac{1}{\sigma^2}\log\left[\frac{2\sigma\sqrt{2\pi}}{m_\sigma(\mathbf{x};\mathbf{\Omega}_0)}\right] \\ &= [\mathbf{\Omega}_0 - \nabla\Omega(\mathbf{x})]^T[\mathbf{\Omega}_0 - \nabla\Omega(\mathbf{x})]. \end{aligned} \tag{19}$$

Letting $\mathbf{\Omega}^*(\mathbf{x})$ denote the center frequency of the channel filter with maximum response magnitude at domain coordinate $\mathbf{x}$ and $\widehat{M}^*(\mathbf{x})$ denote the value of $\widehat{M}_\sigma(\mathbf{x};\mathbf{\Omega}_0)$ for this filter, they formulated the penalty functional

$$\mathcal{P}(\nabla\widehat{\Omega}) = \int_{\mathbb{R}^2}\left\{\left[\mathbf{\Omega}^*(\mathbf{x}) - \nabla\widehat{\Omega}(\mathbf{x})\right]^T\left[\mathbf{\Omega}^*(\mathbf{x}) - \nabla\widehat{\Omega}(\mathbf{x})\right] - \widehat{M}^*(\mathbf{x})\right\}^2 d\mathbf{x}. \tag{20}$$

They then minimized the energy functional

$$\mathcal{E}(\nabla\widehat{\Omega}) = \lambda\mathcal{P}(\nabla\widehat{\Omega}) + \mathcal{S}(\nabla\widehat{\Omega}), \tag{21}$$

where $\mathcal{S}(\nabla\widehat{\Omega})$ was a smoothness constraint term involving the Sobolev norms of $\nabla\widehat{\Omega}$ and where $\lambda > 0$ was a Lagrange multiplier balancing the smoothness constraint against agreement of the estimated emergent frequency solution with (17). The Euler-Lagrange equations for (21) lead to a coupled system of partial differential equations, which they solved numerically using an iterative relaxation. The estimated dominant amplitude modulation function $\widehat{A}(\mathbf{x})$ was then obtained by backsubstitution of the estimated emergent frequencies $\nabla\widehat{\Omega}(\mathbf{x})$ into (17). Using this technique, they obtained several excellent texture segmentations by finding the significant zero crossings in the response of a $\nabla^2 g$ filter applied to the estimated emergent frequency magnitudes$|\nabla\widehat{\Omega}(\mathbf{x})|$.

4.3 Multidimensional Teager-Kaiser Operator

In the course of his notable work on nonlinear speech modeling, Teager developed a computationally efficient discrete operator for estimating the *energy* required to generate a narrowband speech signal with instantaneous frequencies concentrated about a formant [113–117]. The operator was systematically developed by Kaiser [118, 119]. Ensuing research on the operator, now known as the *Teager-Kaiser operator*, established that it may be used to efficiently estimate amplitude and frequency modulations in a signal. Hence, the operator is significant to texture processing in that it represents a direct, computationally efficient alternative to the iterative constrained optimization procedure used to solve (21) for the amplitude and frequency modulations in [107].

The two-dimensional, discrete-domain Teager-Kaiser operator was first introduced by Yu, Mitra, and Kaiser in 1991 for performing edge detection [120]. For a two-dimensional, real-valued image $f(m,n) : \mathbb{Z}^2 \to \mathbb{R}$, the operator is defined by

$$\begin{aligned} \Psi_d\left[f(m,n)\right] &= 2f^2(m,n) - f(m-1,n)f(m+1,n) \\ &\quad - f(m,n-1)f(m,n+1). \end{aligned} \tag{22}$$

This operator was also used by Mitra and others for image enhancement [121–125] and noise removal [126]. The corresponding discrete-domain energy separation algorithm was developed by Maragos, Bovik, and Quatieri [127]. For an image modeled by the discrete equivalent of (2), *viz.*,

$$f(m,n) = a(m,n)\cos\left[\varphi(m,n)\right], \tag{23}$$

define the components of the phase gradient vector field $\nabla\varphi(m,n)$ according to

$$U(m,n) = \left.\frac{\partial}{\partial x}\varphi(x,y)\right|_{x=m,y=n}, \tag{24}$$

$$V(m,n) \;=\; \frac{\partial}{\partial y}\varphi(x,y)\bigg|_{x=m,y=n}. \tag{25}$$

Then the discrete two-dimensional energy separation algorithm of [127] may be expressed as

$$|U(m,n)| \approx \arcsin\left\{\left[\frac{\Phi_d\left[f(m+1,n)-f(m-1,n)\right]}{4\Phi_d\left[f(m,n)\right]}\right]^{\frac{1}{2}}\right\}, \tag{26}$$

$$|V(m,n)| \approx \arcsin\left\{\left[\frac{\Phi_d\left[f(m,n+1)-f(m,n-1)\right]}{4\Phi_d\left[f(m,n)\right]}\right]^{\frac{1}{2}}\right\}, \tag{27}$$

$$|a(m,n)| \approx \left[\frac{\Phi_d\left[f(m,n)\right]}{\sin^2\left[U(m,n)\right]+\sin^2\left[V(m,n)\right]}\right]^{\frac{1}{2}}. \tag{28}$$

In [127], Maragos, Bovik, and Quatieri also introduced a multidimensional continuous-domain Teager-Kaiser operator for a signal $f(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$:

$$\Psi_c\left[f(\mathbf{x})\right] = \left|\nabla f(\mathbf{x})\right|^2 - f(\mathbf{x})\nabla^2 f(\mathbf{x}). \tag{29}$$

With $\mathbf{x} = [x_1\ x_2\ \ldots\ x_n]^T$ and $f(\mathbf{x}) = a(\mathbf{x})\cos\left[\varphi(\mathbf{x})\right]$, the energy separation algorithm associated with (29) is

$$\left|\frac{\partial}{\partial x_i}\varphi(\mathbf{x})\right| \approx \left\{\frac{\Psi_c\left[\frac{\partial}{\partial x_i}f(\mathbf{x})\right]}{\Psi_c\left[f(\mathbf{x})\right]}\right\}^{\frac{1}{2}}, \tag{30}$$

$$|a(\mathbf{x})| \approx \frac{\Psi_c\left[f(\mathbf{x})\right]}{\left\{\sum_{i=1}^{n}\Psi_c\left[\frac{\partial}{\partial x_i}f(\mathbf{x})\right]\right\}^{\frac{1}{2}}}. \tag{31}$$

Conditions under which the approximate equalities in (26) – (28), (30), and (31) hold accurately were investigated by Maragos and Bovik [128, 129], who also gave examples of applying the discrete algorithms (26) – (28) to several natural images. Extensions of the multidimensional Teager-Kaiser operator applicable to vector-valued functions were discussed in [127, 129].

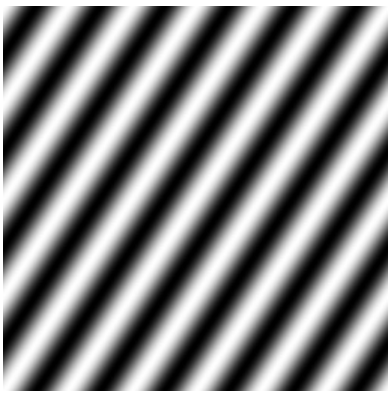
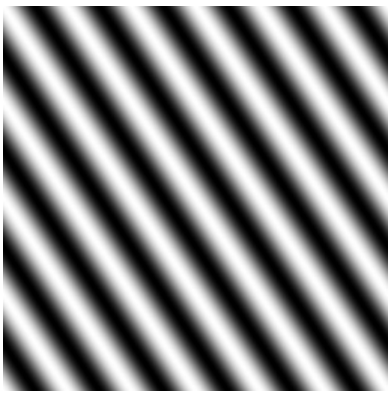

(a)

(b)

Figure 1: Two sinusoidal gratings which demonstrate that negative frequency is required to unambiguously represent signals of dimension greater than one. (a) $t(x, y) = \cos(U_0 x + V_0 y)$. (b) $t(x, y) = \cos(U_0 x - V_0 y)$.

4.4 Demodulation with Signed Frequency

For *images* and other signals with dimension greater than one, the *signs* of the components of the frequency vector are generally significant. Even for the case of a pure sinusoidal grating, the *relative* signs of the spatial frequencies embody *orientation*. Figure 1(a) shows the image $t(x, y) = \cos(U_0 x + V_0 y)$. The image $t(x, y) = \cos(U_0 x - V_0 y)$ is shown in Figure 1(b). These two images differ only in the *relative* signs of their spatial frequencies, yet they are obviously distinct.

However, from (26) and (27), it is clear that the multidimensional Teager-Kaiser operator is inherently incapable of estimating the *signs* of the components of the multidimensional instantaneous frequency vector.

However, it is straightforward to demodulate an arbitrary, *complex-valued* n-dimensional signal

$$t(\mathbf{x}) = a(\mathbf{x}) \exp\left[j\varphi\left(\mathbf{x}\right)\right], \tag{32}$$

where $\mathbf{x} = [x_1 \; x_2 \; \ldots \; x_n]^T$, $a(\mathbf{x}) : \mathbb{R}^n \to [0, \infty)$, $\varphi(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$, and $t(\mathbf{x}) : \mathbb{R}^n \to \mathbb{C}$ with correctly signed instantaneous frequencies. The amplitude modulation function

$a(\mathbf{x})$ is obtained through

$$a(\mathbf{x}) = |t(\mathbf{x})| = |a(\mathbf{x}) \exp\left[j\varphi\left(\mathbf{x}\right)\right]|. \tag{33}$$

Taking the gradient of $t(\mathbf{x})$ yields

$$\nabla t(\mathbf{x}) = jt(\mathbf{x})\nabla\varphi(\mathbf{x}) + \exp\left[j\varphi\left(\mathbf{x}\right)\right]\nabla a(\mathbf{x}), \tag{34}$$

from which it follows that $\forall\ \mathbf{x}$ where $a(\mathbf{x}) \neq 0$,

$$\frac{\nabla t(\mathbf{x})}{jt(\mathbf{x})} = \nabla\varphi(\mathbf{x}) - j\frac{\nabla a(\mathbf{x})}{a(\mathbf{x})}. \tag{35}$$

Equation (35) leads immediately to the spatially local nonlinear frequency algorithm

$$\nabla\varphi(\mathbf{x}) = \mathrm{Re}\left[\frac{\nabla t(\mathbf{x})}{jt(\mathbf{x})}\right]. \tag{36}$$

The demodulation algorithm (33), (36) is *exact* for *any* arbitrary complex-valued, n-dimensional AM-FM signal interpreted as a single component, and gives correct signs for the components of the instantaneous frequency vector $\nabla\varphi(\mathbf{x})$ [112].

As a demodulation example, Figure 2 depicts the single-component complex-valued image *Bent Chirp*. The real part of the image is shown in Figure 2(a), while the imaginary part is appears in Figure 2(b). Along curves which are everywhere normal to the wavefronts, the phase of the image is a slowly varying quadratic. The estimated amplitude modulation function for this image obtained by applying (33) is unity to within floating point precision. The estimated frequency modulation function is shown in Figure 2(c), where needle lengths are directly proportional to the magnitude frequency $|\nabla\varphi(\mathbf{x})|$. The needles point in the direction $\arg\nabla\varphi(\mathbf{x})$. These frequency estimates are in near perfect agreement with the model values.

Now consider an n-dimensional real-valued image $s(\mathbf{x}) : \mathbb{R}^n \rightarrow \mathbb{R}$, where $\mathbf{x} =$

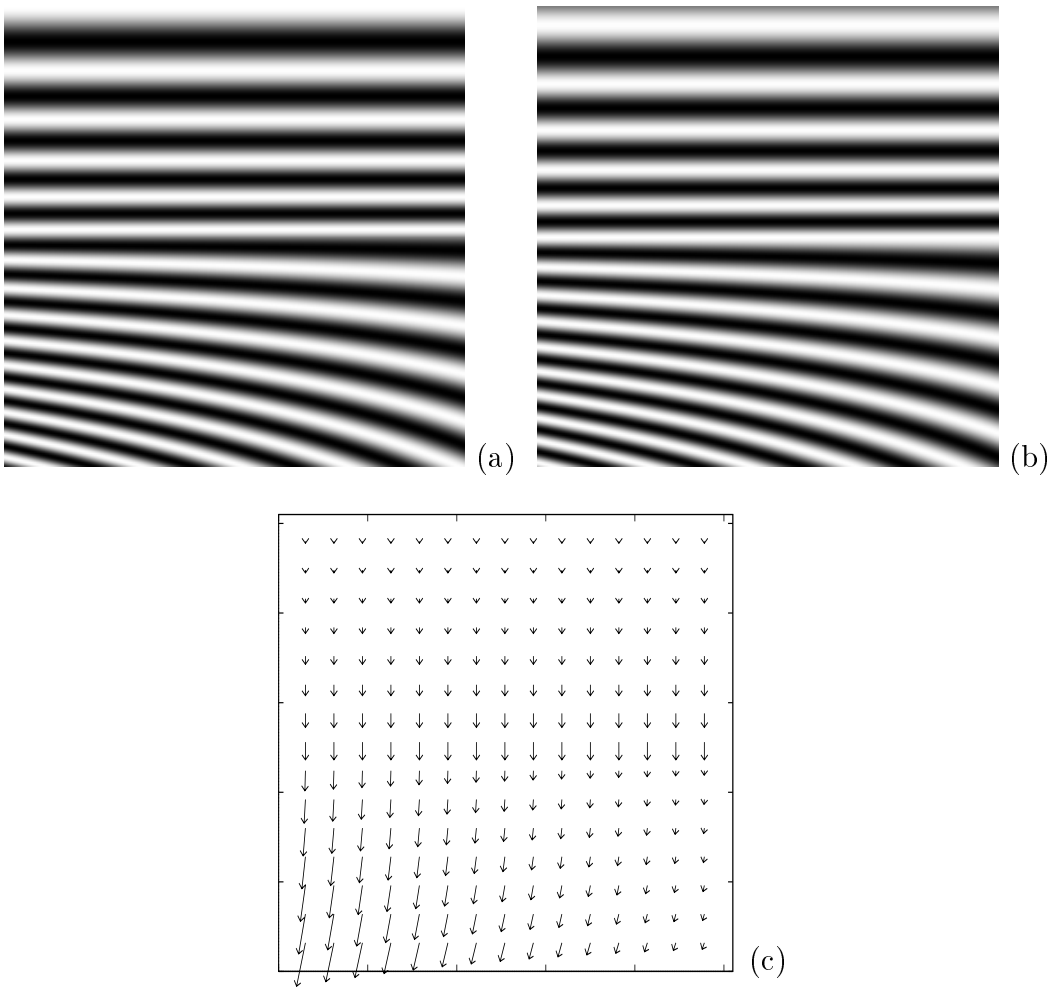


Figure 2: AM-FM analysis of *Bent Chirp* image. (a) *Bent Chirp* image. (b) Imaginary part. (c) Estimated frequency modulation function. Needle length is proportional to the *frequency*.

$[x_1\ x_2\ \ldots\ x_n]^T$. Denote the unit vector in the x_i direction by

$$\mathbf{e}_i = [0\ 0\ \ldots\ 0\ 1\ 0\ \ldots\ 0]^T, \tag{37}$$

where the solitary 1 appears in position i. The image $s(\mathbf{x})$ may be demodulated by applying the algorithm (33), (36) to the unique complex extension

$$t(\mathbf{x}) = s(\mathbf{x}) + j\mathcal{H}\left[s(\mathbf{x})\right], \tag{38}$$

where

$$\mathcal{H}\left[s(\mathbf{x})\right] = \frac{1}{\pi}\int_{\mathbb{R}} s(\mathbf{x} - \xi\mathbf{e}_i)\frac{d\xi}{\xi} = s(\mathbf{x}) * \frac{1}{\pi\mathbf{x}^T\mathbf{e}_i}\prod_{k\neq i}\delta(\mathbf{x}^T\mathbf{e}_k), \tag{39}$$

is the multidimensional Hilbert transform of $s(\mathbf{x})$. The integral in (39) takes its Cauchy principal value and the image $t(\mathbf{x})$ in (38) is refered to as the *analytic image* associated with the real-valued image $s(\mathbf{x})$ [130].

4.5 Dominant Component Texture Processing

Many signals of practical interest are inherently multipartite in character. For example, an image may contain a single textured region that the human observer perceives as having multiple distinct characteristics. An image patch textured by a crosshatched pattern is one instance of such a region. Generally, no single-component interpretation of such signals will admit locally coherent modulating functions. Thus, single-component demodulation of a multi-component image does not in general produce a physically meaningful characterization or interpretation of the image texture, even if the modulating functions of each individual image component are exceedingly smooth.

A complex-valued multipartite signal $t(\mathbf{x})$ may be analyzed against the K-component AM-FM model

$$t(\mathbf{x}) = \sum_{i=1}^{K} a_i(\mathbf{x})\exp[j\varphi_i(\mathbf{x})], \tag{40}$$

where $a_i(\mathbf{x}) : \mathbb{R}^n \rightarrow [0, \infty)$ and $\varphi_i(\mathbf{x}) : \mathbb{R}^n \rightarrow \mathbb{R}$. Uncountably infinitely many such decompositions exist for any signal $t(\mathbf{x})$. Of greatest interest are those decompositions which admit only a few locally coherent components. Since the multidimensional Hilbert transform (39) is a continuous linear operator on $L^2(\mathbb{R}^n)$, the multi-component complex-valued image $t(\mathbf{x})$ in (40) may be generated from a real-valued image $s(\mathbf{x})$ by setting $t(\mathbf{x}) = s(\mathbf{x}) + j\mathcal{H}[s(\mathbf{x})]$.

In this section, I will introduce a texture analysis paradigm called *dominant component analysis* [131]. At each point in the image domain, the dominant component paradigm estimates the modulating functions of the component that locally dominates the image spectrum. Thus, dominant component analysis associates a single pair of modulating functions with a multi-component image, just as did the variational approach of Bovik, *et al.* [107]. Since different components may be dominant at different points in the domain, the modulating functions obtained via dominant component analysis may not be everywhere locally smooth.

The dominant component frequency estimates are precisely the *emergent* frequencies which characterize the local texture structure of the image [107]. For example, many images arising from natural physical, chemical, biological, and erosive processes contain textured regions or quasi-repetitive structures. Examples noted by Bovik, *et al.* [107] include crystals, rock strata, a zebra's stripes, wind patterns in sand, and wood grains. The emergent frequency estimates characterize the orientation, roughness, or flow of such patterns. Perspective distortion also gives rise to nonstationary, quasi-regular patterns because the process of image formation involves projecting three-dimensional surfaces onto the two-dimensional focal plane. Besides nonstationary signal analysis, estimates of the dominant component modulating functions are also useful in a variety of computer vision applications. They may be used for texture segmentation and classification [65, 67–69, 71–73], phase-based computational stereopsis [83, 84], three-dimensional shape from texture algorithms [80–82], and in the computation of image flow lines [131, 132].

However, the nonlinear demodulation algorithm (33), (36) is applicable only to an image modeled as a single AM-FM function. With modification, it may be used to estimate the modulating functions of the various components of a multi-component image, provided that the components are isolated from one another prior to the estimation. This separation need not be effected on a global scale. Rather, what is required is that the components be isolated from one another on a pointwise basis. This may be accomplished using a multiband linear filterbank, as described in [107, 133].

A suitable bank of conjointly spatio-spectrally localized Gabor filters is depicted in the frequency domain in Figure 3. There are 40 filters arranged along eight rays spaced equally at angles of approximately 20.6418°, with five filters per ray. Hence the filterbank comprises channels at eight orientations and at five magnitude spatial frequencies. Recall from Section 3 that the research of Richards and Polit and of Caelli and Bevan suggested that human visual function could be explained by the existence of channels at 18 orientations and at four magnitude spatial frequencies. However, in analyzing complex-valued analytic images, only *half* of the 18 orientations need be explicitly implemented, which suggests that human visual function could be emulated by a filterbank comprising channels at *nine* orientations and at *four* magnitude spatial frequencies. Thus, the channel structure of the filterbank depicted in Figure 3 agrees *roughly* with that proposed for the human visual system by the work of Richards and Polit and of Caelli and Bevan. For display, each filter in the figure has been *independently* scaled for maximum dynamic range in the available 256 gray scales and additional scaling has been applied to accentuate the intersections between individual filters.

Let $t(\mathbf{x})$ be a complex-valued multi-component image and let $g(\mathbf{x})$ be the complex-valued impulse response of a particular one of the Gabor functions in an n-dimensional multiband linear bank similar to the one depicted in Figure 3. Suppose that the filter frequency response is $G(\mathbf{\Omega}) = \mathcal{F}[g(\mathbf{x})]$. Let $y(\mathbf{x}) = g(\mathbf{x}) * t(\mathbf{x})$ be the filter response to input $t(\mathbf{x})$, and suppose that, for some particular point $\mathbf{x}$ in the image domain, this

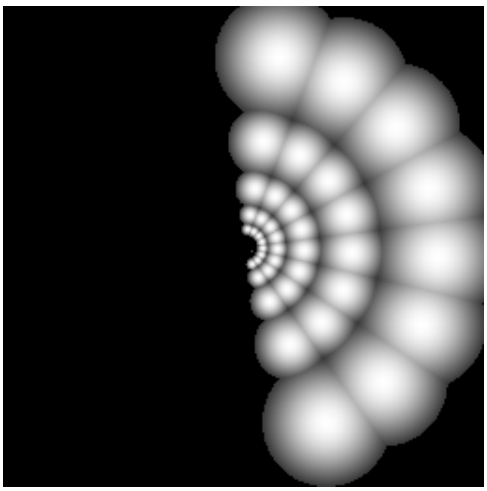


Figure 3: Frequency domain representation of multiband Gabor filterbank suitable for isolating image components from one another on a spatio-spectrally localized basis. There are 40 filters arranged in a polar wavelet-like tesselation along eight rays with five filters per ray. Each of the 40 filters in the figure has been independently scaled for maximum dynamic range in the available gray levels.

response is *dominated* by the image component $t_i(\mathbf{x}) = a_i(\mathbf{x})e^{j\varphi_i(\mathbf{x})}$. Then the modulating functions $a_i(\mathbf{x})$ and $\nabla\varphi_i(\mathbf{x})$ may be estimated by the *approximate* demodulation algorithm

$$\nabla\varphi(\mathbf{x}) \approx \mathrm{Re}\left[\frac{\nabla y(\mathbf{x})}{jy(\mathbf{x})}\right], \tag{41}$$

$$a(\mathbf{x}) \approx \left|\frac{y(\mathbf{x})}{G[\nabla\widehat{\varphi}(\mathbf{x})]}\right|. \tag{42}$$

The approximate algorithm (41), (42) is based on a *quasi-eigenfunction theorem* [107–112] which guarantees that the approximation errors may be expected to be negligible provided that $t(\mathbf{x})$ is locally coherent and that $g(\mathbf{x})$ is reasonably spatially localized.

A block diagram of the dominant component paradigm is shown in Figure 4, where the multicomponent image $t(\mathbf{x})$ is analyzed with an M-channel multiband filterbank to isolate the components from one another on a spatio-spectrally localized basis. At each point in the domain, the response of each filterbank channel is demodulated in the blocks marked *ESA* in Figure 4. Amplitude and frequency estimates are computed from

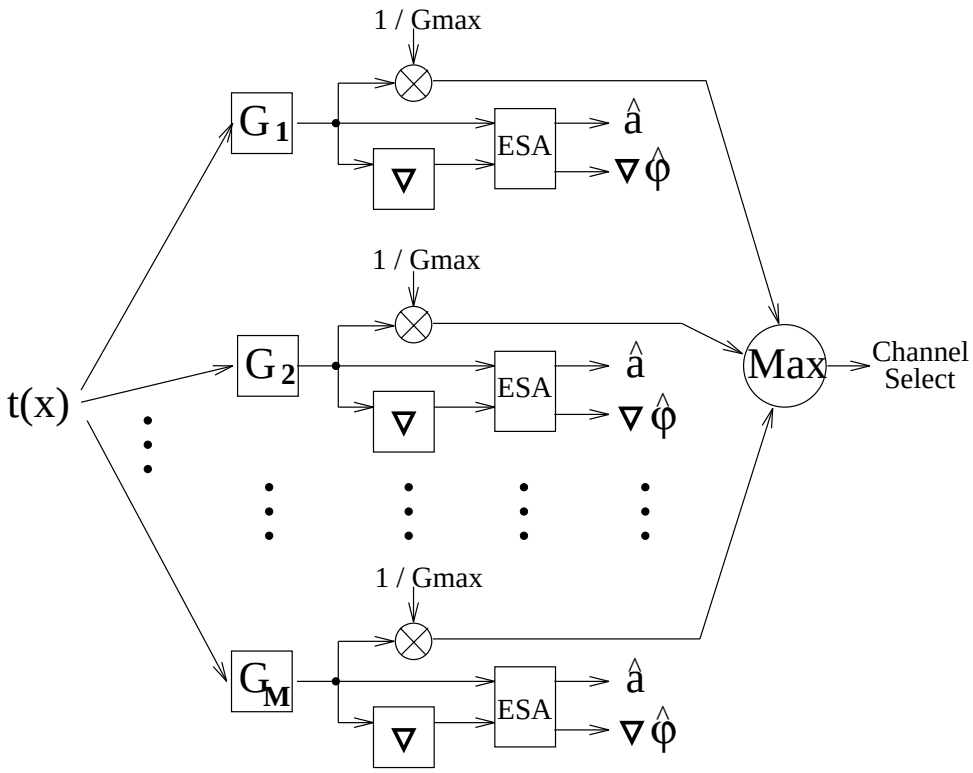


Figure 4: Block diagram of the dominant component paradigm. The demodulation algorithm is performed in the blocks marked ESA. The magnitude of each channel response is divided by the peak value of the channel filter frequency response magnitude to obtain the filter selection criterion. On a pixel by pixel basis, the estimation is performed using the demodulation results from the channel which maximizes this quantity.

each filterbank channel at each point in the image domain. The dominant component demodulation problem is then that of determining which channel should be used at each point. This determination cannot be based solely on the magnitudes of the channel responses, since all of the filters have identical octave bandwidths and unity L^2-norms. Hence, the peak magnitudes of the frequency responses of low frequency channels are much greater than those of high frequency channels. For each channel, a metric called the *filter selection criterion* is computed at each point in the domain, and estimates of the dominant component modulating functions at a given point are taken from the filterbank channel which maximizes the filter selection criterion at that point. The filter selection criterion $\Psi_m(\mathbf{x})$ for channel m is defined by

$$\Psi_m(\mathbf{x}) = \frac{|y_m(\mathbf{x})|}{\max_{\mathbf{\Omega}} |G_m(\mathbf{\Omega})|}, \tag{43}$$

where $y_m(\mathbf{x})$ is the response of filterbank channel m and where $G_m(\mathbf{\Omega})$ is the frequency response of the m'th channel filter. In view of (42), the quantity $\Psi_m(\mathbf{x})$ may be regarded as a crude estimate of the true amplitude modulation function of the image component which dominates the response of channel m at the point $\mathbf{x}$.

While reasonable estimates of the dominant component modulating functions could in many cases be obtained from any one of several channel filters having center frequencies near the dominant instantaneous frequency vector, the technique based on (43) tends to select the channel with the best signal-to-noise ratio. Hence, it affords maximal rejection of cross-component interference and other out-of-band information such as random noise.

As an example, dominant component analysis of the Brodatz texture image *Tree* is shown in Figure 5. The original image appears in Figure 5(a). A needle diagram depicting the the emergent frequency estimates appears in Figure 5(b), where one needle is shown for each block of 10×10 pixels. The needle lengths in Figure 5(b) are inversely proportional to the magnitude of the instantaneous frequency vector. Thus they are proportional to the dominant *instantaneous period*. As before, the needles

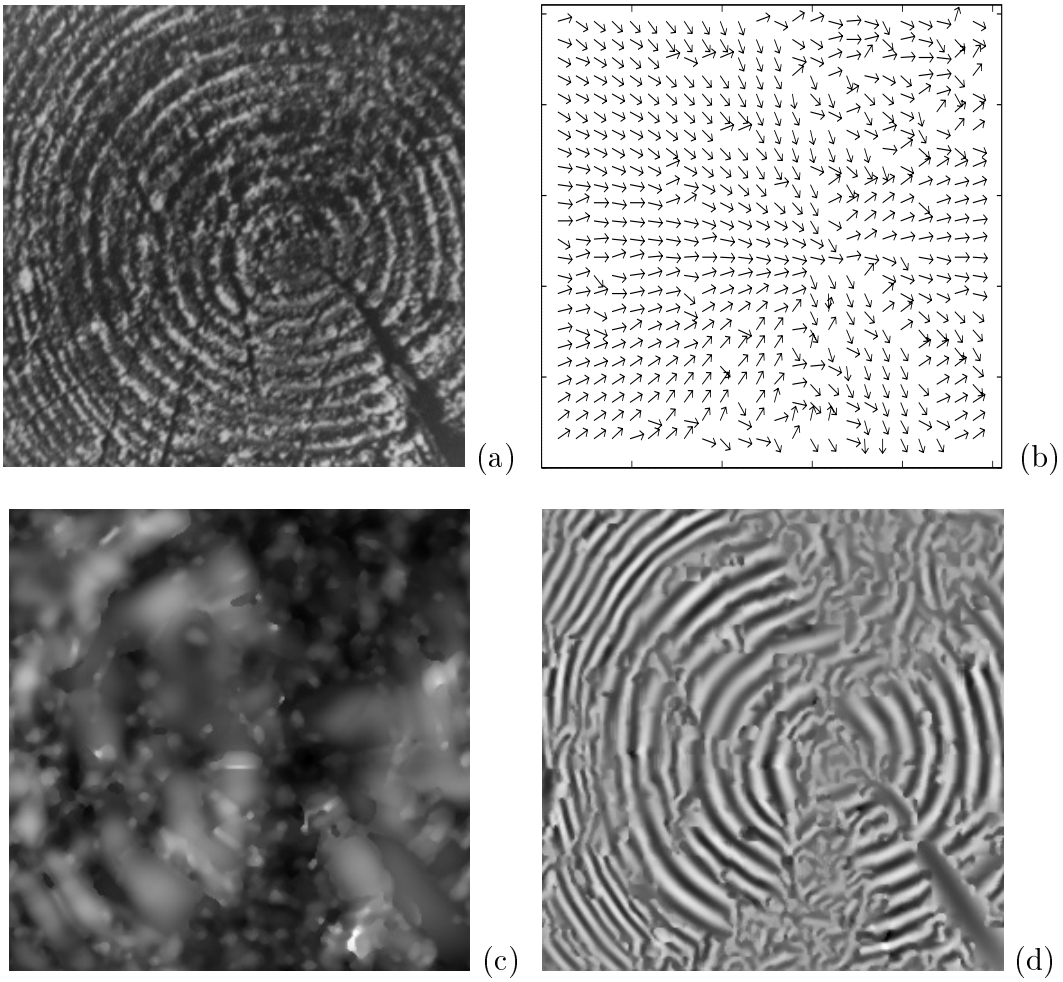


Figure 5: Dominant component analysis of *Tree* image. (a) *Tree* image. (b) Needle diagram depicting estimated emergent frequencies. (c) Estimated dominant component amplitude modulation function. (d) Reconstruction of dominant component from estimated modulating functions.

are oriented with the emergent frequency vector. Note that the needles are normal to dominant texture features, as expected. The estimated dominant component amplitude modulation function is shown in Figure 5(c), and may generally be interpreted as *contrast* in the dominant image component. A reconstruction of the dominant image component from the estimated modulating functions is given in Figure 5(d). Although highly nonstationary, this image possess significant locally coherent texture structure. The similarity between the dominant component reconstruction and the original image is striking and suggests that, over much of the image domain, a substantial fraction of the total texture structure of the image has been captured in a single computed AM-FM component.

5 Example: Texture Segmentation

In this section I will illustrate how the estimated modulating functions delivered by the dominant component analysis paradigm can be used to solve the classical texture segmentation problem. The *Laplacian of Gaussian,* or *LoG filter* is an edge detection filter with a circularly symmetric impulse response [1]. Let $\mathbf{x} = [x \ y]^T$ and $r = \sqrt{x^2 + y^2}$. Then, the impulse response of the LoG filter is given by

$$\nabla^2 g(r) = \left(2 - \frac{r^2}{\sigma^2}\right) \exp\left(-\frac{r^2}{2\sigma^2}\right), \tag{44}$$

where σ, the filter space constant, controls the *scale* at which the filter detects edges. For large values of σ, the filter detects edges which are present across large regions in an image. Likewise, small values of σ yield a filter which detects local edges.

If $s(\mathbf{k})$ is a discrete real-valued image, $\nabla^2 g(\mathbf{k})$ is the unit pulse response of a sampled LoG filter, and $y(\mathbf{k}) = s(\mathbf{k}) * \nabla^2 g(\mathbf{k})$, then $y(\mathbf{k})$ will contain zero crossings where there are edges present in $s(\mathbf{k})$ at the scale σ. If one constructs an edge map $\mathcal{Y}(\mathbf{k})$ by setting $\mathcal{Y}(\mathbf{k}) = 1$ where $y(\mathbf{k})$ contains zero crossings and setting $\mathcal{Y}(\mathbf{k}) = 0$ elsewhere, then the edge contours in $\mathcal{Y}(\mathbf{k})$ will be connected and will be one pixel in width. In some

applications, it is of interest to suppress weak edges in $s(\mathbf{k})$ from appearing in the edge map $\mathcal{Y}(\mathbf{k})$. This may be accomplished by thresholding the gradient magnitude of $y(\mathbf{k})$: one sets $\mathcal{Y}(\mathbf{k}) = 1$ only if $y(\mathbf{k})$ contains a zero crossing at $\mathbf{k}$ *and* $|\nabla y(\mathbf{k})| > \tau$ for some threshold value τ.

Figure 6(a) shows the 256×256 image *Mica-Burlap*, which has a sample mean of the zero and extremes of 1 and -0.8680. A reconstruction of the dominant component obtained by applying the dominant component analysis paradigm to the image is shown in Figure 6(b). The estimated dominant component amplitude modulation function is shown in Figure 6(c), while the estimated emergent frequencies are depicted in the needle diagram of Figure 6(d). Arrow lengths in the needle diagram are proportional to the instantaneous period, and have also been squared for display to accentuate the differences between frequency estimates in different image regions. Figure 6(e) presents the magnitudes of the estimated emergent frequency vectors as a gray scale image. Thus, arrow lengths in Figure 6(d) are reciprocally related to the gray scales in Figure 6(e).

A simple iterative search technique was used to find a LoG filter space constant σ and threshold value τ that would effectively segment the two textures in the *Mica-Burlap* image. First, a sequence of LoG filters with dyadically decreasing space constants was applied to the magnitude frequency image of Figure 6(e). The first, second, and third filters in the sequence had space constants of $\sigma = 64$, 32, and 16 pixels. None of the edge maps corresponding to these first three LoG filters contained a closed contour. For $\sigma = 8$, however, the edge map contained many closed contours. The space constant search sequence was continued by decreasing σ from 16. For $\sigma = 15$ pixels the LoG edge map contained five closed contours, and the space constant search was terminated at this value.

A similar search procedure was then used to find a suitable threshold value for the magnitude LoG response image. In the coarse stage of the search, thresholds with values $\tau = 0.2n$ were applied for $n \in \mathbb{N}$. For $\tau = 1.6$, no closed contours remained

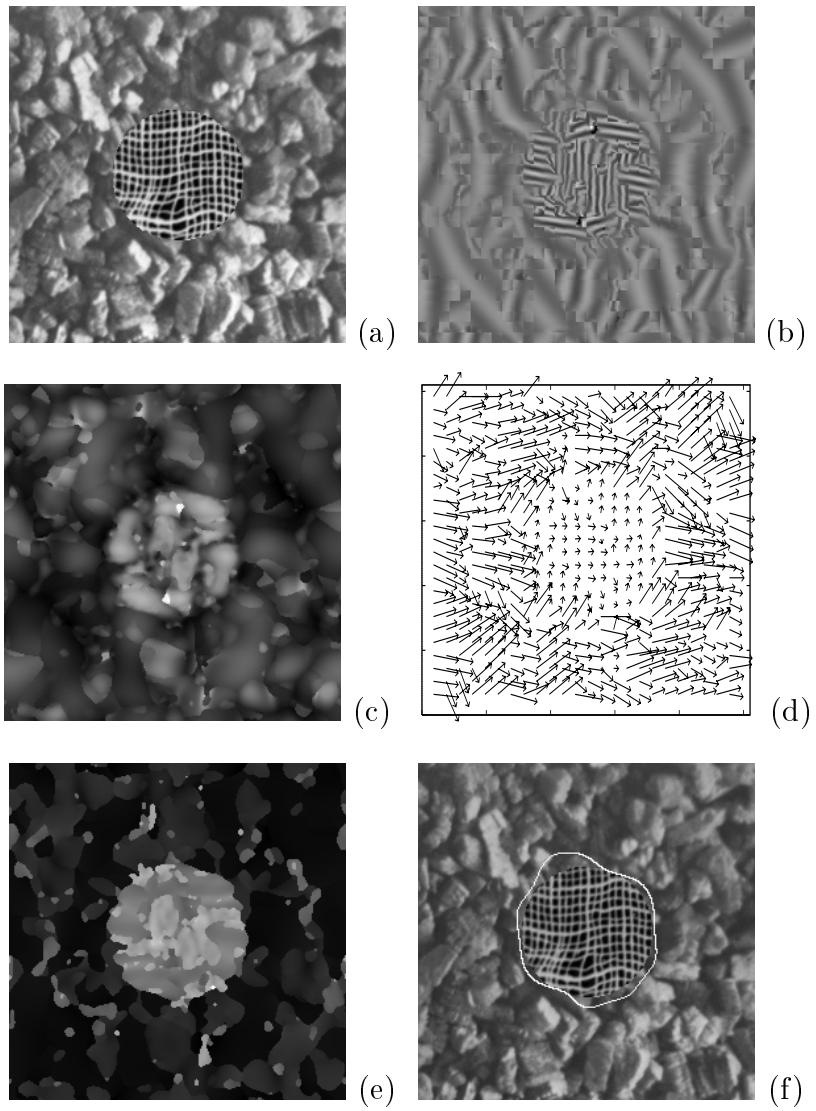


Figure 6: Dominant component texture segmentation of *Mica-Burlap* image. (a) *Mica-Burlap* image. (b) Dominant component reconstruction. (c) Estimated amplitude modulation function. (d) Estimated emergent frequencies. (e) Estimated emergent frequency magnitudes displayed as a gray scale image. (f) Texture segmentation computed by applying a $\nabla^2 g$ filter to the estimated emergent frequency magnitudes.

in the edge map. A fine search was then used to select a threshold value between 1.6 and 1.4. For $\tau = 1.5$, only one closed contour remained in the edge map. This contour is shown overlayed on the original image in Figure 6(f), where it is seen to effectively segment the two textures of the *Mica-Burlap* image.

In addition to the emergent frequency magnitudes, the other quantities delivered by the dominant component analysis paradigm may also be used to perform texture segmentation. For example, the image *Wood-Wood* is shown in Figure 7(a). This 256×256 image was generated by rotating the original image counterclockwise by 45° and subsequently replacing pixels in the central diamond-shaped region of the original image with their counterparts in the rotated image. The sample mean of the *Wood-Wood* image in Figure 7(a) is zero and the extremes are 0.5605 and -1. A reconstruction of the dominant component obtained from the dominant component paradigm is shown in Figure 7(b). The estimated dominant component amplitude modulation function is depicted in Figure 7(c) and the estimated emergent frequencies are given in Figure 7(d). As before, arrows lengths in Figure 7(d) are proportional to the instantaneous period and have been squared for display. Figure 7(e) gives the *orientations* of the estimated emergent frequency vectors as a gray scale image. As in the example of Figure 6, an iterative search was applied to select the LoG filter space constant and threshold values. For a LoG filter with space constant $\sigma = 14$ pixels and a threshold value of $\tau = 11$, the edge map computed from the orientations in Figure 7(e) contained only one closed contour. This contour is shown overlayed on the image in Figure 7(f).

The examples of this section demonstrate that the estimated quantities delivered by a dominant component image analysis are useful in solving the classical texture segmentation problem. The main advantage of the texture segmentation technique illustrated in this section is that the emergent frequency estimates are computed using the spatially localized, computationally efficient algorithm (41) as opposed to the computationally expensive constrained iterative relaxation procedure described in [107].

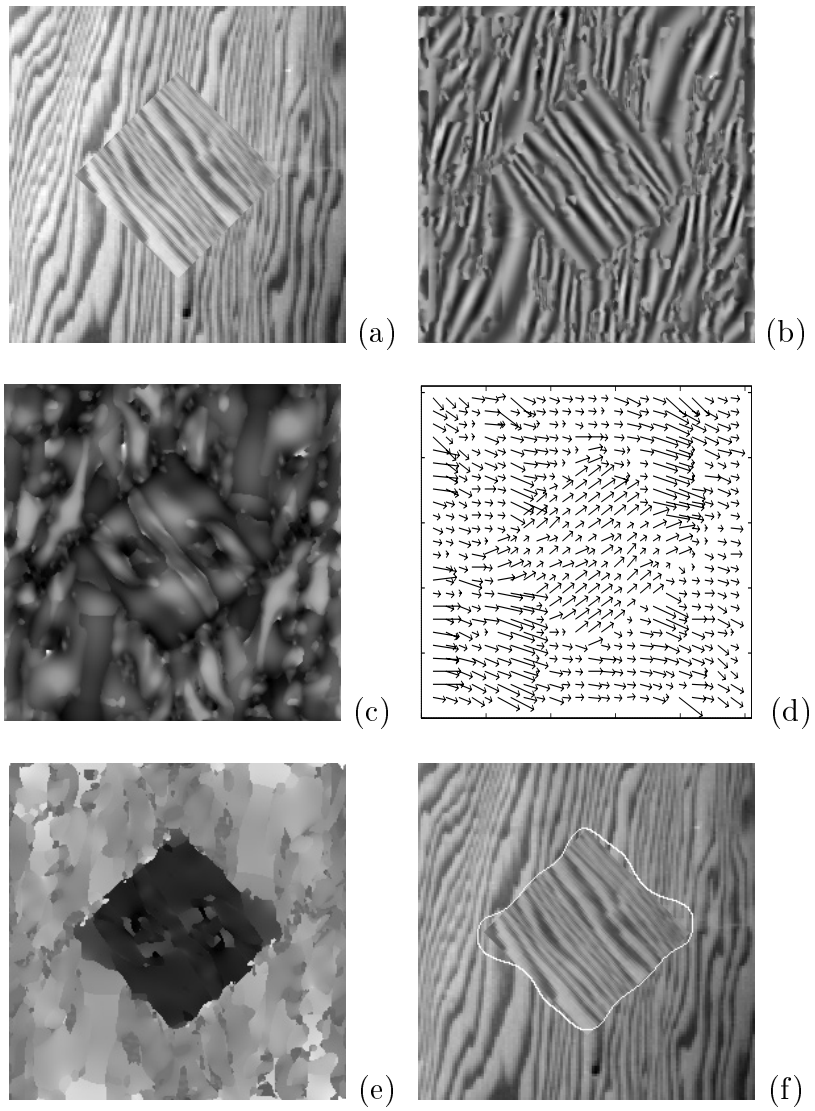


Figure 7: Dominant component texture segmentation of *Wood-Wood* image. (a) *Wood-Wood* image. (b) Dominant component reconstruction. (c) Estimated amplitude modulation function. (d) Estimated emergent frequencies. (e) Estimated emergent frequency orientations displayed as a gray scale image. (f) Texture segmentation computed by applying a $\nabla^2 g$ filter to the estimated emergent frequency orientations.

6 Conclusion

In this paper, I have contended that the philosophy and motivation behind modern approaches to texture processing are fundamentally different from from the philosophy and motivation behind the *early* techniques, which constituted the state of the art in texture processing *c.* 1979. Specifically, the early techniques were primarily *ad hoc*, and generally lacked strong physical, phenomenological, or biological motivation. Some of the early methods were *weakly* motivated by biological considerations, most notably by the Julesz conjectures and by the experimental work of Hubel and Wiesel. While some of these techniques worked well and provided reasonable solutions to classical texture problems for certain classes of images, none of them worked well in general. Methods based on the gray level cooccurrence matrices of Haralick were probably the best from among them.

The modern approaches characterize image texture in terms of nonstationary, locally coherent modulations occurring in frequency and orientation selective channels. These modulations bear strong physical relationships to certain electrochemical cell responses in mammalian biological vision systems. Thus, the modern texture processing techniques posses stronger biological motivation than the *early* methods. This is significant, since the only known vision systems that work well *in general* are biological vision systems. I argued for interpreting contrast and phase in textured images as *modulations*, and also for the representation of these quantities in image processing and machine vision algorithms by computed amplitude and frequency modulation functions. In addition to being generally useful for nonstationary spatio-spectral analysis, computed modulations have demonstrable utility in solving a broad range of texture processing problems in image processing and machine vision, including identification and classification, object recognition, phase-based computational stereopsis, three-dimensional shape from texture, and texture-based scene segmentation. Only the last of these was demonstrated in this paper.

The main thesis of this paper was that the evolution of texture processing from the

early methods to the modern methods was driven in large part by a dramatic series of developments in psychophysics and physiology that occurred between 1968 and 1985. In the context of image processing and machine vision, the most significant results from among these advances were the determination that certain elements of biological vision systems function as spatially localized, orientation selective narrowband frequency filters, that the responses of visual cortical cells explicitly convey contrast and phase information, and that the spatial receptive fields of simple visual cortical cells are well modeled by complex-valued Gabor functions. The fact that we still know relatively little about the higher-level cognitive processing that occurs in biological vision systems suggests that texture processing techniques will continue to evolve and improve with future psychophysical and physiological discoveries, as well as that the future of texture processing research is indeed exciting.

References

[1] D. Marr, *Vision*, W. H. Freeman, New York, 1982.

[2] W. K. Pratt, *Digital Image Processing*, John Wiley & Sons, New York, 1991.

[3] A. K. Jain, *Fundamentals of Digital Image Processing*, Prentice Hall, Englewood Cliffs, NJ, 1989.

[4] R. C. Gonzalesz and R. E. Woods, *Digital Image Processing*, Addison-Wesley, Reading, MA, 1992.

[5] D. H. Ballard and C. M. Brown, *Computer Vision*, Prentice-Hall, Englewood Cliffs, NJ, 1982.

[6] R. M. Haralick, "Statistical and structural approaches to texture", *Proc. IEEE*, vol. 67, no. 5, pp. 786–804, May 1979.

[7] L. Van Gool, P. Dewaele, and A. Oosterlinck, "Texture analysis anno 1983", *Comput. Vision, Graphics, Image Proc.*, vol. 29, no. 3, pp. 336–357, March 1985.

[8] B. Julesz, "Visual pattern discrimination", *IRE Trans. Info. Theory*, vol. 8, pp. 84–92, 1962.

[9] B. Julesz, "Experiments in the visual perception of texture", *Sci. Amer.*, vol. 232, pp. 84–92, 1975.

[10] B. Julesz, "Textons, the elements of texture perception, and their interactions", *Nature*, vol. 290, pp. 91–97, 1981.

[11] B. Julesz and J. R. Bergen, "Textons, the fundamental elements in preattentive vision and perception of textures", *Bell Syst. Tech. J.*, vol. 62, pp. 1619–1645, 1983.

[12] S. W. Zucker and D. Terzopoulos, "Finding structure in co-occurrence matrices for texture analysis", *Comput. Vision, Graphics, Image Proc.*, vol. 12, no. 3, pp. 286–308, March 1980.

[13] D. H. Hubel and T. N. Wiesel, "Receptive fields of single neurons in the cat's striate cortex", *J. Physiol. London*, vol. 148, pp. 574–591, 1959.

[14] D. H. Hubel and T. N. Wiesel, "Receptive fields, binolcular interaction, and functional architecture in the cat's visual cortex", *J. Physiol. London*, vol. 160, pp. 106–154, 1962.

[15] D. H. Hubel and T. N. Wiesel, "Receptive fields and functional architecture of monkey striate cortex", *J. Physiol. London*, vol. 195, pp. 215–243, 1968.

[16] R. Bajcsy and L. Lieberman, "Texture gradients as a depth cue", *Comput. Graphics Image Proc.*, vol. 5, pp. 52–67, 1976.

[17] J. E. Hopcroft and J. D. Ullman, *Introduction to Automata Theory, Languages, and Computation*, Addison-Wesely, Reading, MA, 1979.

[18] A. Rosenfeld and A. C. Kak, *Digital Picture Processing*, Academic Press, San Diego, CA, 2nd edition, 1982.

[19] P. Maragos and R. W. Schafer, "Morphological systems for multidimensional signal processing", *Proc. IEEE*, vol. 78, no. 4, pp. 690–710, April 1990.

[20] F. W. Campbell and J. G. Robson, "Applications of Fourier analysis of the visibility of gratings", *J. Physiol. London*, vol. 197, pp. 551–556, 1968.

[21] N. Graham and J. Nachmias, "Detection of grating patterns containing two spatial frequencies: A comparison of single-channel and multiple-channels models", *Vision Res.*, vol. 11, pp. 251–259, March 1971.

[22] J. M. Woodhouse and H. B. Barlow, "Vision: Spatial and temporal resolution and analysis", in *The Senses*, H. B. Barlow and J. D. Mollon, Eds., pp. 133–164. Cambridge University Press, Cambridge, 1982.

[23] C. Blakemore and F. W. Campbell, "On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images", *J. Physiol. London*, vol. 203, pp. 237–260, 1969.

[24] W. Richards and A. Polit, "Texture matching", *Kybernetic*, vol. 16, no. 3, pp. 155–162, 1974.

[25] T. Caelli and P. Bevan, "Probing the spatial frequency spectrum for orientation sensitivity in stochastic textures", *Vision Res.*, vol. 23, pp. 39–45, 1983.

[26] L. Maffei and A. Fiorentini, "The visual cortex as a spatial frequency analyser", *Vision Res.*, vol. 13, pp. 1255–1267, July 1973.

[27] V. D. Glezer, V. A. Ivanoff, and T. A. Tscherbach, "Investigation of complex and hypercomplex receptive fields of visual cortex of the cat as spatial frequency filters", *Vision Res.*, vol. 13, pp. 1875–1904, 1973.

[28] N. Graham, "Visual detection of aperiodic spatial stimuli by probability summation among narrowband channels", *Vision Res.*, vol. 17, pp. 637–652, 1977.

[29] V. D. Glezer and A. M. Cooperman, "Local spectral analysis in the visual cortex", *Biol. Cybern.*, vol. 28, pp. 101–108, 1977.

[30] J. A. Movshon, I. D. Thompson, and D. J. Tolhurst, "Spatial and temporal contrast sensitivity of neurones in areas 17 and 18 of the cat's visual cortex", *J. Physiol. London*, vol. 283, pp. 101–120, 1978.

[31] L. Maffei, M. C. Morrone, M. Pirchio, and G. Sandini, "Responses of visual cortical cells to periodic and non-periodic stimuli", *J. Physiol. London*, vol. 296, pp. 27–47, 1979.

[32] D. J. Tolhurst and I. D. Thompson, "On the variety of spatial frequency selectivities shown by neurons in area 17 of the cat", *Proc. R. Soc. London Ser. B*, vol. 213, pp. 183–199, 1981.

[33] J. J. Kulikowski and P. O. Bishop, "Linear analysis of the responses of simple cells in the cat visual cortex", *Exp. Brain Res.*, vol. 44, pp. 386–400, 1981.

[34] R. L. De Valois, D. G. Albrecht, and L. G. Thorell, "Spatial frequency selectivity of cells in macaque visual cortex", *Vision Res.*, vol. 22, pp. 545–559, 1982.

[35] J. G. Robson, "Receptive fields: Spatial and intensive representation of the visual image", in *Handbook of Perception, Vol. 5: Vision*, E. C. Carterette and M. P. Friedman, Eds., pp. 81–112. Academic Press, New York, 1975.

[36] J. A. Movshon and D. J. Tolhurst, "On the response linearity of neurons in cat visual cortex", *J. Physiol. London*, vol. 249, pp. 56–57, 1975.

[37] J. A. Movshon, I. D. Thompson, and D. J. Tolhurst, "Spatial summation in the receptive fields of simple cells in the cat's striate cortex", *J. Physiol. London*, vol. 283, pp. 53–77, 1978.

[38] K. K. De Valois, R. L. De Valois, and E. W. Yund, "Responses of striate cortical cells to grating and checkerboard patterns", *J. Physiol. London*, vol. 291, pp. 483–505, 1979.

[39] B. W. Andrews and D. A. Pollen, "Relationship between spatial frequency selectivity and receptive field profile of simple cells", *J. Physiol. London*, vol. 287, pp. 163–176, 1979.

[40] J. J. Kulikowski and P. O. Bishop, "Fourier analysis and spatial representation in the visual cortex", *Experientia*, vol. 37, pp. 160–163, 1981.

[41] J. P. Jones and L. A. Palmer, "An evaluation of the two-dimensional Gabor model of simple receptive fields in cat striate cortex", *J. Neurophysiol.*, vol. 58, no. 6, pp. 1233–1258, 1987.

[42] H. R. Wilson, D. Levi, L. Maffei, J. Rovamo, and R. De Valois, "The perception of form: Retina to striate cortex", in *Visual Perception: the Neurophysiological Foundations*, L. Spillmann and J. S. Werner, Eds., pp. 232–271. Academic Press, San Diego, CA, 1990.

[43] S. Marčelja, "Mathematical description of the responses of simple cortical cells", *J. Opt. Soc. Am.*, vol. 70, no. 11, pp. 1297–1300, November 1982.

[44] D. Gabor, "Theory of communication", *J. Inst. Elect. Eng. London*, vol. 93, no. III, pp. 429–457, 1946.

[45] D. A. Pollen and S. F. Ronner, "Phase relationship between adjacent simple cells in the visual cortex", *Science*, vol. 212, pp. 1409–1411, 1981.

[46] M. J. Basstiaans, "Gabor's expansion of a signal into Gaussian elementary signals", *Proc. IEEE*, vol. 68, pp. 538–539, 1980.

[47] M. J. Basstiaans, "A sampling theorem for the complex spectrogram and Gabor expansion of a signal into Gaussian elementary signals", *Opt. Eng.*, vol. 20, no. 4, pp. 594–598, 1981.

[48] M. J. Basstiaans, "Gabor's signal expansion and degrees of freedom of a signal", *Opt. Acta*, vol. 29, no. 9, pp. 1223–1229, 1982.

[49] D. M. MacKay, "Strife over visual cortical function", *Nature*, vol. 289, pp. 176–218, 1981.

[50] 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*, vol. 2, no. 7, pp. 1160–1169, July 1985.

[51] J. J. Kulikowski, S. Marčelja, and P. O. Bishop, "Theory of spatial position and spatial frequency relations in the receptive fields of simple cells in the visual cortex", *Biol. Cybern.*, vol. 43, no. 3, pp. 187–198, 1982.

[52] B. Sakitt and H. B. Barlow, "A model for the economical encoding of the visual image in cerebral cortex", *Biol. Cybern.*, vol. 43, pp. 97–108, 1982.

[53] D. A. Pollen and S. F. Ronner, "Visual cortical neurons as localized spatial frequency filters", *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-13, no. 5, pp. 907–916, September 1983.

[54] T. Caelli and G. Moraglia, "On the detection of Gabor signals and discrimination of Gabor textures", *Vision Res.*, vol. 25, no. 5, pp. 671–684, 1985.

[55] M. A. Webster and R. L. De Valois, "Relationship between spatial-frequency and orientation tuning of striate-cortex cells", *J. Opt. Soc. Am. A*, vol. 2, no. 7, pp. 1124–1132, July 1985.

[56] J. P. Jones and L. A. Palmer, "The two-dimensional structure of simple receptive fields in cat striate cortex", *J. Neurophysiol.*, vol. 58, no. 6, pp. 1187–1211, 1987.

[57] J. P. Jones, A. Stepnoski, and L. A. Palmer, "The two-dimensional spectral structure of simple receptive fields in cat striate cortex", *J. Neurophysiol.*, vol. 58, no. 6, pp. 1212–1232, 1987.

[58] D. G. Stork and H. R. Wilson, "Analysis of Gabor function descriptions of visual receptive fields", in *Proc. ARVO Ann. Meeting Investig. Opth., Visual Sci.*, Sarasota, FL, May 1-6, 1988, vol. 29, p. 398.

[59] T. G. Stockham, Jr., "Image processing in the context of a visual model", *Proc. IEEE*, vol. 60, no. 7, pp. 828–842, July 1972.

[60] A. Ikonomopoulos and M. Unser, "A directional filtering approach to texture discrimination", in *Proc. 7th Int'l. Conf. Pattern Recog.*, Montreal, Canada, July 30 - August 2, 1984, pp. 87–89.

[61] P. Brodatz, *Textures: a Photographic Album for Artists and Designers*, Dover, New York, 1966.

[62] J. M. Coggins and A. K. Jain, "A spatial filtering approach to texture analysis", *Pattern Recog. Lett.*, vol. 3, no. 3, pp. 195–203, 1985.

[63] A. R. Rao and B. G. Schunck, "Computing oriented texture fields", in *Proc. IEEE Comput. Soc. Conf. Comput. Vision Pattern Recog.*, San Diego, CA, June 4-8, 1989, pp. 61–68.

[64] A. R. Rao and R. C. Jain, "Computerized flow field analysis: Oriented texture fields", *IEEE. Trans. Pattern Anal. Machine Intell.*, vol. 14, no. 7, pp. 693–709, July 1992.

[65] M. R. Turner, "Texture discrimination by Gabor functions", *Biol. Cybern.*, vol. 55, no. 2, pp. 71–82, 1986.

[66] J. Malik and P. Perona, "Preattentive texture discrimination with early vision mechanisms", *J. Opt. Soc. Am. A*, vol. 7, no. 5, pp. 923–932, May 1990.

[67] Y. Y. Zeevi and M. Porat, "Combined frequency-position scheme of image representation in vision", *J. Opt. Soc. Am. A*, vol. 1, no. 12, p. 1248, December 1984.

[68] M. Porat and Y. Y. Zeevi, "The generalized Gabor scheme of image representation in biological and machine vision", *IEEE. Trans. Pattern Anal. Machine Intell.*, vol. 10, no. 4, pp. 452–468, July 1988.

[69] M. Porat and Y. Y. Zeevi, "Localized texture processing in vision: Analysis and synthesis in the Gaborian space", *IEEE. Trans. Biomed. Eng.*, vol. 36, no. 1, pp. 115–129, January 1989.

[70] R. Wilson and M. Spann, "Finite prolate spheroidal sequences and their application II: Image feature description and segmentation", *IEEE. Trans. Pattern Anal. Machine Intell.*, vol. 10, no. 2, pp. 193–203, March 1988.

[71] M. Clark and A. C. Bovik, "Texture discrimination using a model of visual cortex", in *Proc. IEEE Int'l. Conf. Syst., Man, Cybern.*, Atlanta, GA, 1986.

[72] A. C. Bovik, M. Clark, and W. S. Geisler, "Computational texture analysis using localized spatial filtering", in *Proc. IEEE Comput. Soc. Workshop Comput. Vision*, Miami Beach, FL, December 1987.

[73] A. C. Bovik, M. Clark, and W. S. Geisler, "Multichannel texture analysis using localized spatial filters", *IEEE. Trans. Pattern Anal. Machine Intell.*, vol. 12, no. 1, pp. 55–73, January 1990.

[74] A. C. Bovik, "Analysis of multichannel narrow-band filters for image texture segmentation", *IEEE. Trans. Signal Proc.*, vol. 39, no. 9, pp. 2025–2043, September 1991.

[75] A. Witkin, "Recovering surface shape and orientation from texture", *Artificial Intell.*, vol. 17, pp. 17–45, 1981.

[76] M. Kass and A. Witkin, "Analyzing oriented patterns", *Comput. Vision, Graphics, Image Proc.*, vol. 37, pp. 362–385, 1987.

[77] B. Schachter, "Long crested wave models", *Comput. Vision, Graphics, Image Proc.*, vol. 12, no. 2, pp. 187–201, 1980.

[78] J. M. Francos and B. Friedlander, "The polynomial phase differencing operator for modeling of nonhomogeneous images", in *Proc. IEEE Int'l. Conf. Image Proc.*, Washington, DC, October 22-25, 1995, vol. II, pp. 276 – 279.

[79] B. Friedlander and J. M. Francos, "An estimation algorithm for 2-D polynomial phase signals", *IEEE Trans. Image Proc.*, vol. 5, no. 6, pp. 1084–1087, June 1996.

[80] B. J. Super and A. C. Bovik, "Solution to shape-from-texture by wavelet-based measurement of local spectral moments", Tech. Rept. TR-92-5-80, Computer and Vision Research Center, The University of Texas at Austin, November 1991.

[81] B. J. Super and A. C. Bovik, "Shape from texture using local spectral moments", *IEEE. Trans. Pattern Anal. Machine Intell.*, vol. 17, no. 4, pp. 333–343, 1995.

[82] B. J. Super and A. C. Bovik, "Planar surface orientation from texture spatial frequencies", *Pattern Recog.*, vol. 28, no. 5, pp. 728–743, 1995.

[83] T. -Y. Chen, A. C. Bovik, and B. J. Super, "Multiscale stereopsis via Gabor filter phase response", in *Proc. IEEE Int'l. Conf. Syst., Man, and Cyber.*, SanAntonio, TX, October 2-5, 1994, pp. 55–60.

[84] T. -Y. Chen and A. C. Bovik, "Stereo disparity from multiscale processing of local image phase", in *Proc. IEEE Int'l. Symp. Comput. Vision*, Coral Gables, FL, November 20-22, 1995.

[85] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*, Prentice Hall, Englewood Cliffs, NJ, 1989.

[86] M. R. Portnoff, "Time-frequency representation of digital signals and systems based on short-time Fourier analysis", *IEEE. Trans. Acoust., Speech, Signal Proc.*, vol. 28, no. 1, pp. 55–69, January 1980.

[87] J. G. Kirkwood, "Quantum statistics of almost classical ensembles", *Phys. Rev.*, vol. 44, pp. 31–37, 1933.

[88] A. W. Rihaczek, "Signal energy distribution in time and frequency", *IEEE Trans. Info. Theory*, vol. 14, pp. 369–374, 1968.

[89] E. Wigner, "On the quantum correction for thermodynamic equilibrium", *Phys. Rev.*, vol. 40, pp. 749–759, 1932.

[90] G. F. Boudreaux-Bartels and T. W. Parks, "Time-varying filtering and signal estimation using Wigner distribution synthesis techniques", *IEEE. Trans. Acoust., Speech, Signal Proc.*, vol. 36, pp. 442–451, June 1986.

[91] P. M. Woodward, *Probability and Information Theory, with Applications to Radar*, McGraw-Hill, New York, 1953.

[92] A. Papoulis, *Signal Analysis*, McGraw-Hill, New York, 1977.

[93] L. Cohen, "Time-frequency distributions – a review", *Proc. IEEE*, vol. 77, no. 7, pp. 941–981, July 1989.

[94] L. Cohen, *Time-Frequency Analysis*, Prentice Hall, Englewood Cliffs, NJ, 1995.

[95] P. Flandrin and O. Rioul, "Affine smoothing of the Wigner-Ville distribution", in *Proc. IEEE Int'l. Conf. Acoust., Speech, Signal Proc.*, Albuquerque, NM, April 3-6, 1990.

[96] J. Bertrand and P. Bertrand, "Affine time frequency distributions", in *Time-Frequency Signal Analysis — Methods and Applications*, B. Boashash, Ed. Longman-Chesire, Melbourne, Australia, 1991.

[97] H. I. Choi and W. J. Williams, "Improved time-frequency representation of multicomponent signals using exponential kernels", *IEEE. Trans. Acoust., Speech, Signal Proc.*, vol. 37, no. 6, pp. 862–871, 1989.

[98] A. Grossmann and J. Morlet, "Decomposition of Hardy functions into square integrable wavelets of constant shape", *SIAM J. Math. Anal.*, vol. 15, no. 4, pp. 723–736, July 1984.

[99] I. Daubechies, "Orthonormal bases of compactly supported wavelets", *Commun. Pure Appl. Math.*, vol. 51, pp. 909–996, 1988.

[100] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation", *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 31, pp. 674–693, July 1989.

[101] S. G. Mallat, "Multiresolution approximations and wavelet orthonormal bases of $L^2(R)$", *Trans. Amer. Math. Soc.*, vol. 315, no. 1, pp. 69–87, September 1989.

[102] S. G. Mallat, "Multifrequency channel decompositions of images and wavelet models", *IEEE Trans. Acoust., Speech, Signal Proc.*, vol. 37, no. 12, pp. 2091–2110, December 1989.

[103] G. Strang, "Wavelets and dilation equations: A brief introduction", *SIAM Rev.*, vol. 31, no. 4, pp. 614–627, December 1989.

[104] I. Daubechies, "The wavelet transform, time-frequency localization, and signal analysis", *IEEE Trans. Info. Theory*, vol. 36, no. 5, pp. 961–1005, September 1990.

[105] P. Goupillaud, A. Grossmann, and J. Morlet, "Cycle-octave and related transforms in seismic signal analysis", *Geoexploration*, vol. 23, pp. 85–102, April 1984.

[106] A. C. Bovik, "Variational pattern analysis using Gabor wavelets", in *Proc. IEEE Int'l. Conf. Acoust., Speech, Signal Proc.*, San Francisco, CA, March 1992, vol. IV, pp. 669–672.

[107] A. C. Bovik, N. Gopal, T. Emmoth, and A. Restrepo, "Localized measurement of emergent image frequencies by Gabor wavelets", *IEEE. Trans. Info. Theory*, vol. 38, no. 2, pp. 691–712, March 1992.

[108] J. P. Havlicek, D. S. Harding, and A. C. Bovik, "Discrete quasi-eigenfunction approximation for AM-FM image analysis", in *Proc. IEEE Int'l. Conf. Image Proc.*, Lausanne, Switzerland, September 16-19, 1996, pp. 633–636.

[109] A. C. Bovik, J. P. Havlicek, D. S. Harding, and M. D. Desai, "Limits on discrete modulated signals", *IEEE Trans. Signal Proc.*, vol. 45, no. 4, pp. 867–879, April 1997.

[110] J. P. Havlicek and A. C. Bovik, "AM-FM models, the analytic image, and nonlinear demodulation techniques", Tech. Rept. TR-95-001, Center for Vision and Image Sciences, The University of Texas at Austin, March 1995.

[111] J. P. Havlicek and A. C. Bovik, "Multi-component AM-FM image models and wavelet-based demodulation with component tracking", in *Proc. IEEE Int'l. Conf. Image Proc.*, Austin, TX, November 13-16, 1994, vol. I, pp. 41–45.

[112] J. P. Havlicek, A. C. Bovik, and P. Maragos, "Modulation models for image processing and wavelet-based image demodulation", in *Proc. 26th IEEE Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, October 26-28, 1992, pp. 805–810.

[113] H. M. Teager, "Some observations on oral air flow during phonation", *IEEE. Trans. Acoust., Speech, Signal Proc.*, vol. 28, no. 5, pp. 599–601, October 1980.

[114] H. M. Teager and S. M. Teager, "The effects of separated air flow on vocalization", in *Vocal Fold Physiology: Contemporary Research and Clinical Issues*, D. M. Bless and J. H. Abbs, Eds., pp. 124–143. College-Hill Press, San Diego, CA, 1983.

[115] H. M. Teager and S. M. Teager, "Active fluid dynamic voice production models or 'there is a unicorn in the garden' ", in *Vocal Fold Physiology: Biomechanics, Acoustics, and Phonatory Control*, I. Titze and R. Scherer, Eds., pp. 387–401. Denver Center for the Performing Arts, Denver, CO, 1983.

[116] H. M. Teager and S. M. Teager, "A phenomenological model for vowel production in the vocal tract", in *Speech Sciences: Recent Advances*, R. G. Daniloff, Ed., pp. 73–109. College-Hill Press, San Diego, CA, 1983.

[117] H. M. Teager and S. M. Teager, "Evidence for nonlinear sound production mechanisms in the vocal tract", in *Speech Production and Speech Modelling*, W. J. Hardcastle and A. Marchal, Eds., vol. 55 of *NATO Advanced Study Institute Series D*, pp. 241–261. Kluwer Academic Publishers, Boston, MA, 1990.

[118] J. F. Kaiser, "On a simple algorithm to calculate the 'energy' of a signal", in *Proc. IEEE Int'l. Conf. Acoust., Speech, Signal Proc.*, Albuquerque, NM, April 1990, pp. 381–384.

[119] J. F. Kaiser, "On Teager's energy algorithm and its generalization to continuous signals", in *Proc. 4th IEEE DSP Workshop*, New Paltz, NY, September 1990.

[120] T. -H. Yu, S. K. Mitra, and J. F. Kaiser, "Novel algorithm for image enhancement", in *Proc. SPIE/SPSE Conf. Image Proc. Alg. and Tech. II*, San Jose, CA, February 1991.

[121] S. K. Mitra, H. Li, I. S. Lin, and T. -H. Yu, "A new class of nonlinear filters for image enhancement", in *Proc. IEEE Int'l. Conf. Acoust., Speech, Signal Proc.*, Toronto, Ont., Canada, May 1991, pp. 2525–2528.

[122] S. Thurnhofer and S. K. Mitra, "Nonlinear detail enhancement of error-diffused images", in *Proc. IS&T/SPIE Symp. Elect. Imag.: Science & Tech.*, San Jose, CA, February 6-10, 1994, SPIE vol. 2179, pp. 170–181.

[123] T. -H. Yu and S. K. Mitra, "Unsharp masking with nonlinear filters", in *Signal Processing VII: Theories and Applications*, M. Holt, C. Cowan, P. Grant, and W. Sandham, Eds. European Association for Signal Processing, 1994.

[124] N. Strobel and S. K. Mitra, "Quadratic filters for image contrast enhancement", in *Proc. 28th Annual Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, November 1994, pp. 208–212.

[125] 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, June 20-22, 1995, pp. 959–962.

[126] S. K. Mitra and T. -H. Yu, "A new nonlinear algorithm for the removal of impulse noise from highly corrupted images", in *Proc. 1994 IEEE Int'l Symp. Circ. Syst.*, London, U. K., May 1994, pp. 17–20.

[127] P. Maragos, A. C. Bovik, and T. F. Quatieri, "A multidimensional energy operator for image processing", in *Proc. SPIE Symp. Visual Commun. Image Proc.*, Boston, MA, November 16-18, 1992, pp. 177–186.

[128] P. Maragos and A. C. Bovik, "Demodulation of images modeled by amplitude-frequency modulation using multidimensional energy separation", in *Proc. IEEE Int'l. Conf. Image Proc.*, Austin, TX, 1994, vol. III, pp. 421–425.

[129] P. Maragos and A. C. Bovik, "Image demodulation using multidimensional energy separation", *J. Opt. Soc. Amer. A*, vol. 12, no. 9, pp. 1867–1876, September 1995.

[130] J. P. Havlicek, J. W. Havlicek, and A. C. Bovik, "The analytic image", in *Proc. IEEE Int'l. Conf. Image Proc.*, Santa Barbara, CA, October 26-29 1997, *to appear.*

[131] J. P. Havlicek, M. S. Pattichis, D. S. Harding, A. C. Christofides, and A. C. Bovik, "AM-FM image analysis techniques", in *Proc. IEEE Southwest Symp. Image Anal., Interp.*, San Antonio, TX, April 8-9, 1996, pp. 195–199.

[132] M. S. Pattichis and A. C. Bovik, "Multi-dimensional frequency modulation in texture images", in *Proc. Int'l. Conf. on Digital Signal Proc.*, Limassol, Cyprus, June 26-28, 1995, pp. 753–758.

[133] J. P. Havlicek, D. S. Harding, and A. C. Bovik, "The multi-component AM-FM image representation", *IEEE Trans. Image Proc.*, vol. 5, no. 6, pp. 1094–1100, June 1996.