AM-FM Models, the Analytic Image, and Nonlinear Demodulation Techniques.

Joseph P. Havlicek and Alan C. Bovik

CVIS-TR-95-001 March, 1995

CENTER FOR VISION AND IMAGE SCIENCES
The University of Texas at Austin, Mezes Hall 330, Austin, TX 78712

AM-FM Models, the Analytic Image, and Nonlinear Demodulation Techniques

Joseph P. Havlicek and Alan C. Bovik *

Laboratory for Vision Systems

University of Texas, Austin, TX 78712-1084

Tech. Rept. TR-95-001

Center for Vision and Image Sciences, The University of Texas at Austin

March 28, 1995

Abstract – We present very general nonlinear AM-FM models, and use them to analyze and represent complicated nonstationary images. The nonstationarities in such images, which often contribute significantly to visual perception and interpretation, are difficult to capture using stationary frequency techniques. Computationally efficient, nonlinear demodulation algorithms are developed to estimate the modulating functions associated with the image component which dominates the frequency spectrum on a spatially local basis. Theoretical bounds on the estimation errors are developed using a multidimensional quasi-eigenfunction approximation. If multiple components or noise are present, the dominant component must be isolated by multiband filtering. The filters are designed to realize the lower bound of the estimation error. Since the demodulation algorithm operates on a complex valued image, we also introduce the relatively little known multidimensional Hilbert transform and use it to generate the *analytic image* associated with a real valued image. We show that some of the properties that have made the 1D analytic signal so widely successful carry over to the multidimensional case as well. The demodulation algorithm and filter bank design are described in sufficient detail to permit a full implementation of the technique. High quality frequency estimates are obtained with a minimum of computation and can serve as low level input to to a whole range of texture based processing including analysis, segmentation, texture modeling, shape from texture, and frequency based stereo.

*This research was supported in part by a grant from the Texas Advanced Research Projects Agency and by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under grant number F49620-93-1-0307.

Contents

List of Figures

9 Relationship between the conjugate antisymmetric spectrum of $j\boldsymbol{\mathcal{H}}[s(\mathbf{x})]$ and the conjugate symmetric spectrum of $s(\mathbf{x})$. (a) For cancellation between the spectra in the left-half plane, we must have $S_{CA}(\boldsymbol{\Omega}) = -S_{CS}(\boldsymbol{\Omega})$ in quadrants II and III. This implies that $S_{CA}(\boldsymbol{\Omega}) = S^*_{CS}(-\boldsymbol{\Omega})$ in quadrants I and IV. (b) The final relation, which results from the conjugate symmetry of $S_{CS}(\boldsymbol{\Omega})$. . 32

List of Tables

1 Introduction

Many images contain regions of surface texture or patterned markings that can be effectively modeled as being locally narrowband, in the sense that on a spatially local basis the image spectrum is dominated by a single *emergent* frequency, or by a narrow distribution of frequencies. If the locally emergent frequencies do not vary too wildly or discontinuously, then such regions are described as being *locally coherent*. Certainly, uniformity and stationarity are not necessary attributes of a pattern occurring on a cohesive surface. Many natural patterns are in fact uniform, and may even have a regular structure. Others are nearly regular, but for some kind of physical deformation, *e.g.*, a zebra's camouflage, or wind patterns in sand. In these cases, the deformation contributes as much to the perception of the current state of the surface pattern as does the pattern itself. Similarly, perspective distortion in the process of image formation also gives rise to nonuniformities in the imaged patterns. Still other natural patterns are irregular but are nevertheless highly oriented on a local basis, such as clouds or a wheat field. The development of techniques for characterizing and representing such images in a way that facilitates and simplifies analysis in terms of the locally occurring nonstationarities, which manifest important structural and perception cues, remains an important open problem.

Significant recent research has demonstrated the efficacy of AM-FM modeling techniques for analyzing and characterizing locally coherent nonstationary signals and images [1], [2], [3], [4], [5], [6], [7], [8], [9], [10], [11], [12], [13]. In 1-D, AM-FM models can be efficiently computed using the Teager-Kaiser Energy Operator [2], [4], [6], [8], [9], [10]. In terms of analyzing and interpreting the structure, information content, and origin of certain important signals, the inherent ability of AM-FM models to capture local nonstationarities offers significant advantages over traditional time-frequency distributions.

A multi-dimensional version of the Teager-Kaiser operator, as well as the demodulation algorithm to be developed in this report, have been used to compute AM-FM models of images [3], [5]. In image processing and computational vision, AM-FM modeling has great utility both in estimating instantaneous frequencies on a spatio-spectrally localized basis and in formulating an efficient representation. Indeed, the characterization of images in

terms of their spatially localized instantaneous frequency content is fundamental to an increasing variety of multi-dimensional processing techniques including analysis, segmentation and modeling of texture [14], [15], [16], [17],[18],[19], shape from texture [20], and emerging techniques in stereopsis [21].

We model a locally coherent image component with the nonlinear, complex-valued AM-FM function

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

where in the general n-dimensional case $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, $t : \mathbb{R}^n \to \mathbb{C}$, $a : \mathbb{R}^n \to [0, 1]$, and $\varphi : \mathbb{R}^n \to \mathbb{R}$. For a given image, the modulating functions $a(\mathbf{x})$ and $\varphi(\mathbf{x})$ in the model (1) are not unique. We are most interested in those representations where the modulating functions are smooth in the sense of being *slowly varying*, since they tend to be efficient, and also facilitate analysis in terms of the nonstationarities.

While *any* image may be represented with the model (1), in some cases no such representation will possess smooth modulating functions. This is true for several classes of images which are not well suited to AM-FM modeling, including those that are everywhere highly discontinuous, fractal or self-similar in nature, or otherwise incoherent. It is also true for locally coherent images which exhibit a multi-partite nature. The latter contain regions in which the local frequency spectrum is dominated by multiple emergent frequencies. They can be effectively treated with the multi-component model

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

For images comprising a small number of components, each with slowly-varying modulating functions, representation in terms of (2) allows multichannel processing analogous to that which is known to occur in biological vision systems, including those of cat and macaque. The model is also useful in analyzing and representing the locally coherent components of an image that contains both coherent and incoherent information, as well as locally coherent images that have been corrupted by additive noise, *e.g.* while being transmitted through a noisy channel.

The fact that the models (1), (2) are complex valued simplifies the demodulation problem,

which is concerned with estimating the modulating functions. In this report we describe the procedure for formulating a particular complex extension of a real image known as the associated *analytic image*. The analytic image can be analyzed against the models. We also develop algorithms for simultaneously demodulating all components of a multi-component image on a pointwise basis, and for computing the *dominant component AM-FM representation*, which consists of a single component that contains the dominant emergent frequencies at each point in the image.

2 Complex Representation of a Real Image

Consider a real valued image $s(\mathbf{x})$ that we wish to analyze against the complex valued model. We must generate an extension by adding an imaginary component $jw(\mathbf{x})$ prior to demodulation. This raises the question of how $w(\mathbf{x})$ should be chosen. This is ill-posed, because for arbitrary $s(\mathbf{x})$ there are uncountably infinitely many sets of modulating functions $\{a_k(\mathbf{x}), \nabla\varphi_k(\mathbf{x})\}_{k\in[1,K]}$ for which $s(\mathbf{x})$ is exactly equal to the real part of the model (2).

Assume for the moment that $s(\mathbf{x})$ is *known* to be a single component image. Then $s(\mathbf{x}) = \mathrm{Re}\{a(\mathbf{x})\exp[j\varphi(\mathbf{x})]\} = a(\mathbf{x})\cos[\varphi(\mathbf{x})]$, and so clearly $w(\mathbf{x}) = \mathrm{Im}\{a(\mathbf{x})\exp[j\varphi(\mathbf{x})]\}$. But without *a priori* knowledge of $a(\mathbf{x})$ and $\varphi(\mathbf{x})$, this fact does not help us generate $w(\mathbf{x})$. There has been considerable dialogue on this in 1-D [22], [23], [24], [25], [26], [27], where one normally starts out with a specific pair of modulating functions $\{\widehat{a}(\mathbf{x}), \nabla\widehat{\varphi}(\mathbf{x})\}$, called the *imposed modulations*, in mind. For various choices of $w(\mathbf{x})$, with $a(\mathbf{x})\exp[j\varphi(\mathbf{x})] = s(\mathbf{x}) + jw(\mathbf{x})$, one can then proceed to formulate how closely $a(\mathbf{x})$ and $\nabla\varphi(\mathbf{x})$ agree with the imposed modulations. This approach makes sense in a communications engineering context, where the signal under analysis was in fact created by imposing information bearing modulations onto a carrier signal.

In computational vision and image processing, however, $s(\mathbf{x})$ is more likely to be a natural image that we wish to represent or interpret. Choices for $w(\mathbf{x})$ that lead to modulating functions which are smooth yield representations that are efficient and facilitate processing analogous to that which occurs in biological vision systems. We argue that a natural and sensible choice for $w(\mathbf{x})$ is one that cancels the redundancy in the conjugate symmetric

spectrum $S(\Omega)$. This results in a complex extension whose spectrum has support only on a frequency half plane. Following the 1-D convention [28], we refer to the extension generated by cancelling the redundancy in $S(\Omega)$, such that the extended spectrum is supported only in the right-half plane, as the *analytic image* associated with $s(\mathbf{x})$. The analytic image can be conveniently calculated using a multidimensional Hilbert transform.

2.1 2-Dimensional Hilbert Transform

The Fourier transform of a real valued one-dimensional signal $s(x)$ is conjugate symmetric, and hence the signal is completely determined by its Fourier transform on a half-line. Therefore, $s(x)$ can be uniquely associated with a complex analytic signal $t(x)$ whose Fourier transform is supported only on the nonnegative half-line, where it is equal to twice $\mathcal{F}[s(x)]$. Clearly then, $t(x)$ is obtained by cancelling all spectral redundancy in $s(x)$. The two signals are related by $s(x) = \mathrm{Re}[t(x)]$ and $t(x) = s(x) + j\boldsymbol{\mathcal{H}}[s(x)]$, where $\boldsymbol{\mathcal{H}}[\cdot]$ denotes Hilbert transform.

Likewise, the two-dimensional Fourier transform of a real image is conjugate symmetric about the origin. So a real valued image $s(\mathbf{x})$ is uniquely determined by its Fourier spectrum $S(\boldsymbol{\Omega}) = \mathcal{F}[s(\mathbf{x})]$ on a half-plane. If arbitrary real images are to be represented, this implies that any stationary frequency representation must incorporate $S(\boldsymbol{\Omega})$ for *negative* values of one of the frequency variables. As an illustration of this fact, consider the two pure sinusoidal grating images shown in Figure 1, which differ only in the signs of their vertical frequencies. These images could not both be represented in a spectral representation scheme based solely on nonnegative frequencies. From this important observation it is clear that, in spaces of dimension greater than one, a general algorithm for instantaneous frequency demodulation and estimation must be capable of estimating negative frequencies. This is the primary shortcoming of higher-dimensional energy separation algorithms based on the Teager-Kaiser operator, which does not preserve sign information.

In computing the analytic image associated with $s(\mathbf{x})$, we make use of the relatively little known multidimensional Hilbert transform acting in the direction of the unit vector $\mathbf{e}_i$,

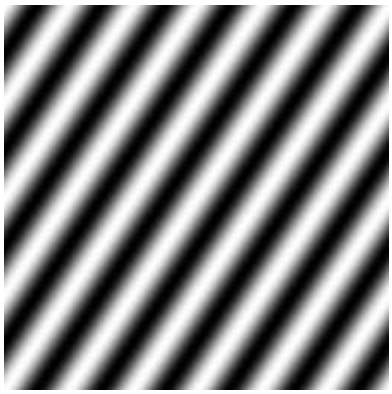
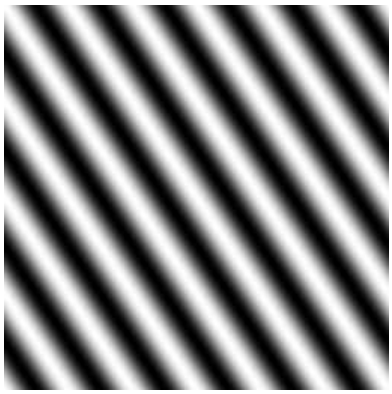

(a)

(b)

Figure 1: Two sinusoidal gratings which demonstrate that negative stationary frequency is required to uniquely represent signals of dimension greater than one. (a) $\cos(u_0 x + v_0 y)$. (b) $\cos(u_0 x - v_0 y)$.

defined by

$$\begin{aligned} \boldsymbol{\mathcal{H}}[s(\mathbf{x})] &= \frac{1}{\pi} \int_{\mathbb{R}} s(\mathbf{x} - y\mathbf{e}_i) \frac{dy}{y} && (3) \\ &= s(\mathbf{x}) * \frac{\delta(\mathbf{x}^T \mathbf{e}_j)}{\pi \mathbf{x}^T \mathbf{e}_i}, && (4) \end{aligned}$$

where the integral is interpreted as a Cauchy principle value, $\mathbf{e}_j$ is the unit vector orthogonal to $\mathbf{e}_i$, and δ is the Dirac delta. If $w(\mathbf{x}) = \boldsymbol{\mathcal{H}}[s(\mathbf{x})]$, then the Fourier transforms of $w(\mathbf{x})$ and $s(\mathbf{x})$ are related by

$$W(\boldsymbol{\Omega}) = \mathcal{F}[w(\mathbf{x})] = -j \mathrm{sgn}\left[\boldsymbol{\Omega}^T \mathbf{e}_i\right] S(\boldsymbol{\Omega}). \quad (5)$$

Henceforth, we will always take the transform in the $\mathbf{e}_x = [1, 0]^T$ direction, which means that the spectrum of the complex signal $s(\mathbf{x}) + j\boldsymbol{\mathcal{H}}[s(\mathbf{x})]$ is supported only in quadrants I and IV of the $\boldsymbol{\Omega} = [u, v]^T$ frequency plane. Appendix A is a derivation of (3), (5) proceeding from engineering motivations. The result is alternatively obtained by deriving a directional bounded operator on $L^2(\mathbb{R}^n)$ which commutes with translations and dilations, and which anticommutes with reflections in its direction of action [29].

For general $w(\mathbf{x})$, the spectrum of $s(\mathbf{x}) + jw(\mathbf{x})$ will contain stationary (Fourier) frequencies that are not present in $S(\Omega)$. We feel that the choice $w(\mathbf{x}) = \boldsymbol{\mathcal{H}}[s(\mathbf{x})]$ is a natural one because it gives $t(\mathbf{x}) = s(\mathbf{x}) + j\boldsymbol{\mathcal{H}}[s(\mathbf{x})]$, the analytic image associated with $s(\mathbf{x})$, whose

Fourier transform differs from $S(\mathbf{\Omega})$ only by the cancellation of the redundant information in the left frequency half-plane. The analytic image also associates a *unique* pair of modulating functions $a(\mathbf{x})$, $\nabla\varphi(\mathbf{x})$ with $s(\mathbf{x})$, thereby giving us in $\nabla\varphi(\mathbf{x})$ an unambiguous defintion of the instantaneous frequencies of $s(\mathbf{x})$. In Section 2.2 we argue that the modulating functions associated with the analytic image also tend to be slowly varying, in a certain sense.

2.2 An Amplitude Sifting Property

We introduce a theorem which indicates that image features which are slowly varying in the sense that they contain only low Fourier frequency information in the horizontal direction will, under reasonable assumptions, be interpreted as amplitude modulations in the analytic image $t(\mathbf{x})$.

Theorem 1 *Let* $a(\mathbf{x}), \cos[\varphi(\mathbf{x})] : \mathbb{R}^2 \to \mathbb{R}$ *be in* $L^2(\mathbb{R}^2)$ *and have Fourier integrals* $\mathcal{F}[a(\mathbf{x})] = A(\mathbf{\Omega})$ *and* $\mathcal{F}[\cos[\varphi(\mathbf{x})]] = \Phi(\mathbf{\Omega})$. *With* $\mathbf{\Omega} = [u, v]^T$, *suppose* $\exists\ u_0 \in \mathbb{R}$ *such that* $|A(u,v)| = 0$ $\forall\ |u| > u_0$ *and* $|\Phi(u,v)| = 0\ \forall\ |u| < u_0$. *Then*

$$\boldsymbol{\mathcal{H}}\Big[a(\mathbf{x})\cos[\varphi(\mathbf{x})]\Big] = a(\mathbf{x})\boldsymbol{\mathcal{H}}\Big[\cos[\varphi(\mathbf{x})]\Big]. \qquad \blacksquare \tag{6}$$

Proof: Appendix B.

Theorem 1 is a specialized 2-D version of Bedrosian's Hilbert transform product theorem [23],[25]. The result lends insight into how the Hilbert transform extension technique interprets the AM-FM constituents of an image. Under the hypothesis, we have immediately that

$$a(\mathbf{x})\cos[\varphi(\mathbf{x})] + j\boldsymbol{\mathcal{H}}\Big[a(\mathbf{x})\cos[\varphi(\mathbf{x})]\Big] = a(\mathbf{x})\Big\{\cos[\varphi(\mathbf{x})] + j\boldsymbol{\mathcal{H}}\Big[\cos[\varphi(\mathbf{x})]\Big]\Big\}. \tag{7}$$

Let $s(\mathbf{x})$ be a single-component real valued image interpreted as a product $a(\mathbf{x})\cos[\varphi(\mathbf{x})]$ (for multiple components, the following arguments apply on a component-wise basis by linearity of the Hilbert transform). The most useful AM-FM interpretations are those for which slow variations in $s(\mathbf{x})$ are manifested in the amplitude modulation term $a(\mathbf{x})$, while faster variations are manifested in $\cos[\varphi(\mathbf{x})]$, since this gives representations where both

modulating functions are slowly varying.

Consider a particular pair of slowly varying modulating functions $a_1(\mathbf{x}), \nabla\varphi_1(\mathbf{x})$ satisfying $s(\mathbf{x}) = a_1(\mathbf{x})\cos[\varphi_1(\mathbf{x})]$. This gives a complex extension $t_1(\mathbf{x}) = a_1(\mathbf{x})\exp[j\varphi_1(\mathbf{x})]$ that in general differs from the analytic image $t_2(\mathbf{x}) = a_2(\mathbf{x})\exp[j\varphi_2(\mathbf{x})]$ obtained with the Hilbert transform. However, since rapid variations in $s(\mathbf{x})$ are captured in $\cos[\varphi_1(\mathbf{x})]$, while slow, smooth variations are captured in $a_1(\mathbf{x})$, it is reasonable to assume the existence of a horizontal frequency u_0 such that the spectrum of $a_1(\mathbf{x})$ is supported only on the strip $|u| < u_0$, while the spectrum of $\cos[\varphi_1(\mathbf{x})]$ is supported only outside this strip (indeed, if no such u_0 exists, then we may have difficulty arguing that one or the other of $a_1(\mathbf{x}), \cos[\varphi_1(\mathbf{x})]$ should be interpreted as embodying amplitude modulation as opposed to frequency modulation). Then we may use Theorem 1 to formulate $t_2(\mathbf{x})$, the analytic image associated with $s(\mathbf{x})$, in terms of the selected modulations $a_1(\mathbf{x})$, $\nabla\varphi_1(\mathbf{x})$:

$$\begin{aligned} t_2(\mathbf{x}) &= a_1(\mathbf{x})\cos[\varphi_1(\mathbf{x})] + j\boldsymbol{\mathcal{H}}\Big[a_1(\mathbf{x})\cos[\varphi_1(\mathbf{x})]\Big] \\ &= a_1(\mathbf{x})\Big\{\cos[\varphi_1(\mathbf{x})] + j\boldsymbol{\mathcal{H}}\Big[\cos[\varphi_1(\mathbf{x})]\Big]\Big\} \\ &= a_2(\mathbf{x})\exp[j\varphi_2(\mathbf{x})]\,. \end{aligned} \tag{8}$$

From (8) we see that any features of $s(\mathbf{x})$ captured in $a_1(\mathbf{x})$ are also interpreted as amplitude modulation in the analytic image; they do not contribute to the frequency modulation term. In general, there may also be additional features of $s(\mathbf{x})$ that are interpreted as frequency modulation in $t_1(\mathbf{x})$, but as amplitude modulation in the analytic image $t_2(\mathbf{x})$. This establishes that under very reasonable conditions, in addition to providing an unambiguous definition of instantaneous frequency, complex representation of the real image $s(\mathbf{x})$ by the associated analytic image will tend to interpret slowly varying image features as amplitude modulation, while faster variations will tend to be interpreted as frequency modulation.

3 Single Component Demodulation

We now develop an exact demodulation algorithm for a single component analytic image $t(\mathbf{x})$ of the form (1). First observe that

$$|t(\mathbf{x})| = |a(\mathbf{x}) \exp[j\varphi(\mathbf{x})]| = a(\mathbf{x}). \tag{9}$$

Next, take the gradient of $t(\mathbf{x})$ to obtain

$$\begin{aligned} \nabla t(\mathbf{x}) &= \sum_k t^{(k)}(\mathbf{x}) \mathbf{e}_k \\ &= \sum_k \left\{ a^{(k)}(\mathbf{x}) \exp[j\varphi(\mathbf{x})] + jt(\mathbf{x})\varphi^{(k)}(\mathbf{x}) \right\} \mathbf{e}_k \\ &= jt(\mathbf{x})\nabla\varphi(\mathbf{x}) + \exp[j\varphi(\mathbf{x})]\nabla a(\mathbf{x}), \end{aligned} \tag{10}$$

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

$$\frac{\nabla t(\mathbf{x})}{jt(\mathbf{x})} = \nabla\varphi(\mathbf{x}) - j\frac{\nabla a(\mathbf{x})}{a(\mathbf{x})}. \tag{11}$$

This leads immediately to the spatially local nonlinear frequency demodulation algorithm

$$\nabla\varphi(\mathbf{x}) = \mathrm{Re}\left[\frac{\nabla t(\mathbf{x})}{jt(\mathbf{x})}\right], \tag{12}$$

which is *exact* for *any* general single-component n-dimensional complex valued AM-FM signal. The imaginary component of (11) is also of interest because it is related to the instantaneous bandwidth $B(\mathbf{x})$, defined by Cohen in one dimension [30], [31], through

$$\left\{ \mathrm{Im}\left[\frac{\nabla t(\mathbf{x})}{jt(\mathbf{x})}\right] \right\}^2 = \left| \frac{\nabla a(\mathbf{x})}{a(\mathbf{x})} \right|^2 = B(\mathbf{x}). \tag{13}$$

For images that are not monochromatic, a distribution of frequencies is present around the instantaneous frequency $\nabla\varphi(\mathbf{x})$ at every point in the domain, and the quantity $B(\mathbf{x})$ characterizes the tightness of the spread of this distribution.

If out-of-band additive noise or multiple components are present, the nonlinear demodu-

lation algorithm (9), (12) breaks down. The failure results from interference between individual components and noise, as well as cross-component interference between the the multiple components themselves. In each of these cases, it is necessary to isolate the various locally coherent components from each other and from the noise prior to applying the demodulation algorithm. We accomplish this by passing the analytic image through a bank of multiband filters which are sufficiently narrowband to prevent interference between components, but are also capable of isolating the AM-FM components on a *spatially* local basis. This is required, since the components will in general be nonstationary. Hence, even though they are *locally narrowband*, they may in fact be *globally wideband*. The algorithm (9), (12) clearly cannot be applied directly to a filtered component. In Section 4 we present a powerful approximation which will be used in Section 5 to develop a demodulation algorithm applicable directly to the *filtered* components of the analytic image.

4 Quasi-Eigenfunction Approximation

Consider processing $t(\mathbf{x})$ given by (1) with a linear filter $g_\sigma(\mathbf{x})$. The response $t_\sigma(\mathbf{x})$ is given by the convolution integral

$$t_\sigma(\mathbf{x}) = \int_{\mathbb{R}^n} t(\mathbf{x}-\mathbf{p})g(\mathbf{p})d\mathbf{p}. \tag{14}$$

If $t(\mathbf{x})$ is monochromatic ($a(\mathbf{x})$ and $\nabla\varphi(\mathbf{x})$ constant), then it is an eigenfunction of $g_\sigma(\mathbf{x})$ and the response is given exactly by $t_\sigma(\mathbf{x}) = t(\mathbf{x})G_\sigma\left[\nabla\varphi(\mathbf{x})\right]$. Otherwise, without specific knowledge of the filter function, simplifications beyond (14) are generally impossible. However, if the modulating functions are sufficiently slowly varying (*i.e.* if $t(\mathbf{x})$ is sufficiently locally coherent), and if $g_\sigma(\mathbf{x})$ is sufficiently spatially localized, then on a *local* basis we expect that $t(\mathbf{x})$ looks very much like an eigenfunction. This motivates the *quasi-eigenfunction* approximation $t_\sigma(\mathbf{x}) \approx \widehat{t}_\sigma(\mathbf{x}) = t(\mathbf{x})G_\sigma\left[\nabla\varphi(\mathbf{x})\right]$.

Theorem 2, which follows, shows for the general n-D case that the approximation error is tightly bounded by products of the p-energy variance of the filter function $g_\sigma(\mathbf{x})$ with Sobolev p-norms of the modulating functions. This form is in agreement with the intuitive discussion of the preceding paragraph. The p-energy variance quantifies the tightness of a function's distribution in space. For $p = 1$, it is the first absolute moment, and for $p = 2$ it

is the variance. Sobolev p-norms are simply L^p norms of a function's derivatives. In general, the smoother a function is, in the sense of having small derivatives, the smaller will be it's Sobolev norms. Therefore, Theorem 2 states that the quasi-eigenfunction approximation will be best when $g_\sigma(\mathbf{x})$ is highly spatially localized and $t(\mathbf{x})$ is highly coherent. Before stating the theorem, we give formal definitions for the p-energy variances and Sobolev p-norms.

Definition 4.1 (directional p-energy variance) *The* p-energy variance *of* $f : \mathbb{R}^n \to \mathbb{C}$ *in the* i^{th} *direction is*

$$\Delta_i^{(p)}(f) = \left[\int_{\mathbb{R}^n} |x_i|^p |f(\mathbf{x})|^p d\mathbf{x} \right]^{\frac{1}{p}}. \qquad \blacksquare \tag{15}$$

Definition 4.2 (total p-energy variance) *The* total p-energy variance *of* $f : \mathbb{R}^n \to \mathbb{C}$ *is the* ℓ^p*-norm of the sequence* $\{\Delta_i^{(p)}(f)\}$*:*

$$\bar{\Delta}_i^{(p)}(f) = \left[\sum_i \left| \Delta_i^{(p)}(f) \right|^p \right]^{\frac{1}{p}} = \left\| \{\Delta_i^{(p)}(f)\} \right\|_{\ell^p}. \qquad \blacksquare \tag{16}$$

Definition 4.3 (directional p-energy cross-variance) *The* p-energy cross-variance *between the* i^{th} *and* j^{th} *directions of* $f : \mathbb{R}^n \to \mathbb{C}$ *is*

$$\Delta_{i,j}^{(p)}(f) = \left[\int_{\mathbb{R}^n} |x_i x_j|^p |f(\mathbf{x})|^p d\mathbf{x} \right]^{\frac{1}{p}}. \qquad \blacksquare \tag{17}$$

Definition 4.4 (total p-energy cross-variance) *The* total p-energy cross-variance *of* $f : \mathbb{R}^n \to \mathbb{C}$ *is the* ℓ^p*-norm of the sequence* $\{\Delta_{i,j}^{(p)}(f)\}$*:*

$$\bar{\Delta}_{i,j}^{(p)}(f) = \left[\sum_{i,j} \left| \Delta_{i,j}^{(p)}(f) \right|^p \right]^{\frac{1}{p}} = \left\| \{\Delta_{i,j}^{(p)}(f)\} \right\|_{\ell^p}. \qquad \blacksquare \tag{18}$$

Definition 4.5 (directional Sobolev p-norm of orders one and two) *The* Sobolev p-norms *of orders one and two of* $f : \mathbb{R}^n \to \mathbb{C}$ *are*

$$\delta_i^{(p)}(f) = \left[\int_{\mathbb{R}^n} \left| \frac{\partial}{\partial x_i} f(\mathbf{x}) \right|^p d\mathbf{x} \right]^{\frac{1}{p}} \tag{19}$$

and

$$\delta_{i,j}^{(p)}(f) = \left[\int_{\mathbb{R}^n} \left|\frac{\partial^2}{\partial x_i \partial x_j} f(\mathbf{x})\right|^p d\mathbf{x}\right]^{\frac{1}{p}}, \tag{20}$$

respectively. ■

Definition 4.6 (total Sobolev p-norms of orders one and two) *The* total Sobolev p-norms *of orders one and two of $f : \mathbb{R}^n \to \mathbb{C}$ are the ℓ^p-norms of the sequences $\{\delta_i^{(p)}(f)\}$ and $\{\delta_{i,j}^{(p)}(f)\}$, respectively:*

$$\bar{\delta}_i^{(p)}(f) = \left[\sum_i \left|\delta_i^{(p)}(f)\right|^p\right]^{\frac{1}{p}} = \left\|\{\delta_i^{(p)}(f)\}\right\|_{\ell^p} \tag{21}$$

and

$$\bar{\delta}_{i,j}^{(p)}(f) = \left[\sum_{i,j} \left|\delta_{i,j}^{(p)}(f)\right|^p\right]^{\frac{1}{p}} = \left\|\{\delta_{i,j}^{(p)}(f)\}\right\|_{\ell^p}. \quad ■ \tag{22}$$

Theorem 2 (Continuous domain n-dimensional quasi-eigenfunction approximation) *Let $g_\sigma : \mathbb{R}^n \to \mathbb{C}$, $a : \mathbb{R}^n \to \mathbb{R}$, and $\varphi : \mathbb{R}^n \to \mathbb{R}$ be such that a is once-differentiable, φ is twice-differentiable, and $\exists$ two (possibly not distinct) pairs of conjugate exponents p, q, and p', q' with $q, q' > n$ and $1 < p, p' < \frac{n}{n-1}$ for which $|x_i g_\sigma(\mathbf{x})|^p$, $|\frac{\partial}{\partial x_i} a(\mathbf{x})|^q$, $|x_i x_j g_\sigma(\mathbf{x})|^{p'}$, and $|\frac{\partial^2}{\partial x_i \partial x_j}\varphi(\mathbf{x})|^{q'}$ are integrable on $\mathbb{R}^n$ for $1 \leq i, j \leq n$. Let*

$$t(\mathbf{x}) = a(\mathbf{x}) \exp[j\varphi(\mathbf{x})], \tag{23}$$

$$t_\sigma(\mathbf{x}) = t(\mathbf{x}) * g_\sigma(\mathbf{x}) = \int_{\mathbb{R}^n} t(\mathbf{x} - \mathbf{y}) g_\sigma(\mathbf{y}) d\mathbf{y}, \tag{24}$$

and

$$\widehat{t}_\sigma(\mathbf{x}) = t(\mathbf{x}) G_\sigma[\nabla\varphi(\mathbf{x})]. \tag{25}$$

Define

$$\varepsilon_t(\mathbf{x}) = |t_\sigma(\mathbf{x}) - \widehat{t}_\sigma(\mathbf{x})|. \tag{26}$$

Then

$$\varepsilon_t(\mathbf{x}) \quad \leq \quad \varepsilon_{t,a}(\mathbf{x}) + \varepsilon_{t,\varphi}(\mathbf{x})$$

$$= \left(\frac{q'}{q'-n}\right)\bar{\Delta}_i^{(p')}(g_\sigma)\bar{\delta}_i^{(q')}(a) + \frac{q^2}{(q-n)(2q-n)}\bar{\Delta}_{i,j}^{(p)}(g_\sigma)\bar{\delta}_{i,j}^{(q)}(\varphi). \qquad \blacksquare \qquad (27)$$

Proof: Appendix C.

In real engineering applications, where grossly pathalogical mathematical functions that might violate the hypothesis are infrequently encountered, the quasi-eigenfunction approximation is an extremely useful and powerful tool. In Section 5, it enables us to design a filtered demodulation algorithm independent of the specific filter function, and in Section 6 we use it to optimize the impulse responses of the multiband filterbank over the class of all n-D linear functions. Of particular interest is the fact that the dimension n appears in the bound. Various instances of error bounds for related quasi-eigenfunction approximations have appeared. The 1-D continuous domain version was proven in [8], the 2-D version in [19], and the n-D version was proved for the case of *FM-only* signals in [32]. The 1-D discrete domain case was treated in [33]. Proof of the discrete domain n-D theorem remains an important open problem.

The unfiltered demodulation algorithm (9), (12) involves partial derivatives of $t(\mathbf{x})$ in the numerator. Corollary 1, which follows, quantifies the conditions under which the approximation can also be applied to these partials. The hypothesis of the corollary imposes restrictions on the positive and negative parts of the amplitude and phase modulations of $t^{(i)}(\mathbf{x})$. Definition of this decomposition precedes the statement of the corollary.

Definition 4.7 (positive/negative decomposition)

$$f_+(\mathbf{x}) = \begin{cases} f(\mathbf{x}), & f(\mathbf{x}) \geq 0; \\ 0, & \textit{otherwise,} \end{cases}$$

$$f_-(\mathbf{x}) = \begin{cases} -f(\mathbf{x}), & f(\mathbf{x}) < 0; \\ 0, & \textit{otherwise.} \end{cases} \qquad \blacksquare$$

Corollary 1 *Let $g_\sigma(\mathbf{x})$, a, and φ be as in Theorem 2, and suppose that the functions $a_+^{(i)}$, $a_-^{(i)}$, $a\varphi_+^{(i)}$, and $a\varphi_-^{(i)}$ also satisfy the hypothesis conditions on a with conjugate exponent pairs (p,q) and (p',q'). Let $t_\sigma^{(i)}(\mathbf{x}) = t^{(i)}(\mathbf{x}) * g_\sigma(\mathbf{x})$ and $\widehat{t}_\sigma^{(i)}(\mathbf{x}) = t^{(i)}(\mathbf{x})G_\sigma[\nabla\varphi(\mathbf{x})]$. Define*

$\varepsilon(\mathbf{x}) = |t_\sigma^{(i)}(\mathbf{x}) - \widehat{t}_\sigma^{(i)}(\mathbf{x})|$. *Then*

$$\varepsilon(\mathbf{x}) \leq \varepsilon_{a_+^{(i)}}(\mathbf{x}) + \varepsilon_{a_-^{(i)}}(\mathbf{x}) + \varepsilon_{a\varphi_+^{(i)}}(\mathbf{x}) + \varepsilon_{a\varphi_-^{(i)}}(\mathbf{x}). \qquad \blacksquare \tag{28}$$

Proof: Appendix D.

5 Filtered Demodulation Algorithm

An important fact is that Corollary 1 holds independently of whether $t_\sigma^{(i)}(\mathbf{x})$ is obtained by filtering the derivative signal $t^{(i)}(\mathbf{x})$ or by differentiating the filtered signal $t_\sigma(\mathbf{x})$, since differentiation and convolution commute. Furthermore, the images $t^{(i)}(\mathbf{x})$ are exactly the components of $\nabla t(\mathbf{x})$, and hence (28) establishes the validity of applying the quasi-eigenfunction approximation directly to the numerator of the frequency demodulation algorithm (12). The validity of applying it to the denominator follows directly from Theorem 2.

Then, provided that $g_\sigma(\mathbf{x})$ is well localized spatially and $t(\mathbf{x})$ is reasonably locally coherent, applying the approximation yields

$$\begin{aligned}
\mathrm{Re}\left[\frac{\nabla t_\sigma(\mathbf{x})}{j t_\sigma(\mathbf{x})}\right] &\approx \mathrm{Re}\left[\frac{\nabla \widehat{t}_\sigma(\mathbf{x})}{j \widehat{t}_\sigma(\mathbf{x})}\right] \\
&= \mathrm{Re}\left[\frac{G_\sigma\left[\nabla\varphi(\mathbf{x})\right]\nabla t(\mathbf{x})}{j G_\sigma\left[\nabla\varphi(\mathbf{x})\right] t(\mathbf{x})}\right] \\
&= \mathrm{Re}\left[\frac{\nabla t(\mathbf{x})}{j t(\mathbf{x})}\right] \\
&= \nabla\varphi(\mathbf{x}).
\end{aligned} \tag{29}$$

Once an estimate for $\nabla\varphi(\mathbf{x})$ is computed, it follows almost immediately that

$$a(\mathbf{x}) \approx \left|\frac{t_\sigma(\mathbf{x})}{G_\sigma\left[\nabla\varphi(\mathbf{x})\right]}\right|. \tag{30}$$

We refer to (29), (30) as the *filtered demodulation algorithm.*

6 Filterbank Design

The unfiltered demodulation algorithm (9), (12) is exact for a locally coherent nonstationary single-component image of the form (1). Filtering is required as a consequence of the nonlinearity of the frequency demodulation algorithm only when multiple components or additive noise are present. In these cases, the filterbank must cover the frequency plane with a tesselation that provides reasonable transmission at all frequencies assumed by the component of interest, and it must do this with filters that are sufficiently spectrally localized to resolve the individual image components from one another and from the noise.

However, the filters must also be well localized spatially if their responses are to capture the local nonstationarities that may be present in the image under analysis. This is made explicit in the error bound (27), where we see that both terms are proportional to the p-energy variance of the filter function, which quantifies the filter's distribution in space. The more spatially localized the filter is, the smaller will be it's p-energy variance.

We have that the filters must be conjointly spatio-spectrally localized. The degree to which *any* filter can simultaneously satisfy these conflicting criteria is limited by the uncertainty principle lower bound on its time-bandwidth product. Therefore, we construct the filter bank from complex Gabor functions, which realize the lower bound uniquely. To preserve the image energy in each filter passband, we insist that each filter have unity L^2-norm. The isotropic unity L^2-norm baseband filter is $h(\mathbf{x}) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{4}\mathbf{x}^T\mathbf{x}\right]$. Upon adding scaling and frequency modulation (translation) while maintaining unity L^2-norm and aspect ratio, we obtain the scaled, translated channel filter with center frequency $\mathbf{\Omega}_m$:

$$g_m(\mathbf{x}) = \frac{1}{\sigma_m\sqrt{2\pi}} \exp\left[-\frac{1}{4\sigma_m^2}\mathbf{x}^T\mathbf{x}\right] \exp\left[j2\pi\mathbf{\Omega}_m^T\mathbf{x}\right]. \tag{31}$$

In the frequency domain, the filter is the Gaussian

$$G_m(\mathbf{\Omega}) = \mathcal{F}\left[g_m(\mathbf{x})\right] = 2\sqrt{2\pi}\sigma_m \exp\left[-4\pi^2\sigma_m^2(\mathbf{\Omega}-\mathbf{\Omega}_m)^T(\mathbf{\Omega}-\mathbf{\Omega}_m)\right], \tag{32}$$

with radial center frequency $r_m = |\mathbf{\Omega}_m|$ and orientation $\theta_m = \arg[\mathbf{\Omega}_m]$. Figure 2 shows the quantities used to define the bandwidth of the filter. The circle in the (b) part of the figure

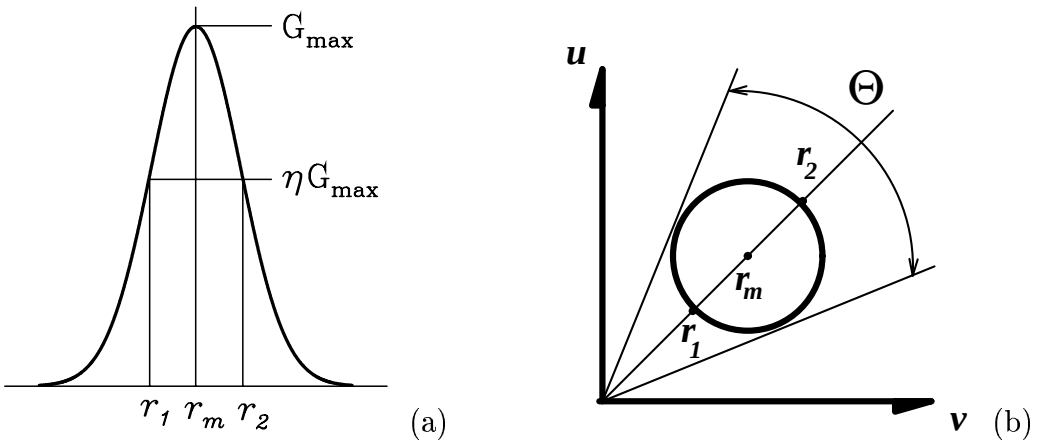


Figure 2: Quantities used in defining the bandwidth B of a Gabor filter. (a) Radial octave bandwidth. The curve shows the filter evaluated on a line from the frequency origin through it's center frequency r_m. The filter magnitude response is at a fraction η of peak response at the radial frequencies r_1 and r_2. (b) Orientation bandwidth. The circle with center r_m is the η-peak contour of the filter in the $\mathbf{\Omega} = [u, v]^T$ plane.

is the η-peak contour of the filter. A line running from the frequency origin through the filter center intercepts the η-peak contour at radial frequencies $r_1 < r_2$, given by $r_1 = \sigma_m - \sqrt{-\ln \eta}$ and $r_2 = \sigma_m + \sqrt{-\ln \eta}$. A section of the filter evaluated along this line is shown in the (a) part of the figure. The η-peak radial octave bandwidth is defined by $B = \log_2[r_2/r_1]$, which is a design parameter of the filter bank. The orientation bandwidth is the angle between two lines tangent to the η-peak contour and passing through the frequency origin. It is given by $\Theta = 2 \arctan \sqrt{\gamma}$, where $\gamma = (2^B - 1)^2/(2^B + 1)^2$.

Since we are using the Hilbert transform complex extension technique, spectra of the analytic images $t(\mathbf{x})$ under analysis are only supported in the right frequency half-plane, where we arrange the filter center frequencies along rays such that any group of four adjacent filters intersect precisely where each is at a fraction η of peak response. With this dense spacing, every point in the right half-plane is covered by a filter responding at η-peak or higher. Along each ray, the filter radial center frequencies progress geometrically with a common ratio R, which is a design parameter of the filterbank. The radial center frequency of the first filter on each ray, r_0, is also a design parameter. The angular spacing between rays is given by $\Lambda = 2 \arcsin \left[(4R)^{-\frac{1}{2}} \left\{ (R^2 + 1)(\gamma - 1) + 2R(\gamma + 1) \right\}^{\frac{1}{2}} \right]$.

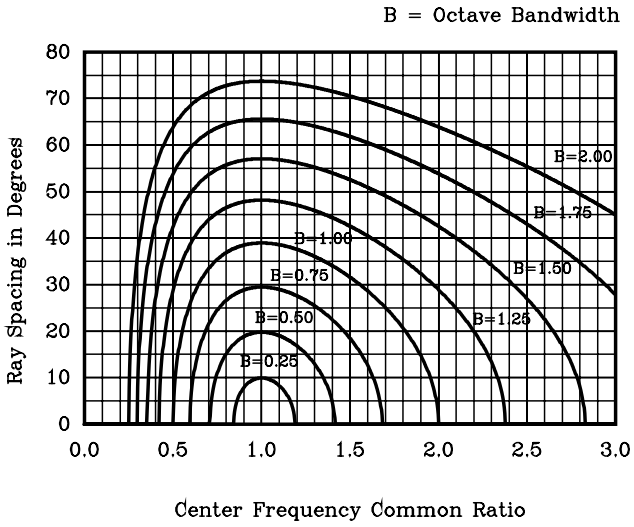


Figure 3: Realizable combinations of the free design parameters R, the common ratio of the geometric progression of filter radial center frequencies along each ray, and B, the filter radial octave bandwidth. The graph shows the ray spacing Λ as a function of R, parameterized by B. Combinations falling below the abscissa of the graph correspond to negative or complex values of the ray spacing Λ, which are not realizable.

The four free design parameters r_0, R, η, and B completely specify the filterbank. However, all arbitrary choices of R and B are not realizable, since certain combinations lead to angular ray spacings Λ that are zero or complex. Figure 3 graphically depicts the realizable combinations by showing Λ as a function of R, parameterized by B. In the figure, values of R less than one correspond to a decreasing progression of filter radial center frequencies on each ray, in which case design parameter r_0 specifies the *highest* radial center frequency which will be assumed by any filter.

For the examples in this paper, we specified the filter bank by $r_0 = 9.6$ cycles per image, $R = 1.8$, $B = 1$ octave, and $\eta = \frac{1}{2}$. In this case, $\gamma = \frac{1}{9}$ and $\Theta = 38.9424°$. Figure 4 shows the entire filter bank in the frequency domain for these parameter choices. The frequency domain coordinates of the figure are as shown in Figure 9 in Appendix A, so that first quadrant is located in the lower right corner of the figure. Under this convention, a vector from the frequency origin to the spectral support of a sinusoidal grating points the *in direction of propagation* of the grating. There are 40 filters arranged on eight rays spaced equally at angles of $\Lambda = 20.6418°$, with five filters per ray, plus one filter at DC. Each filter in the figure

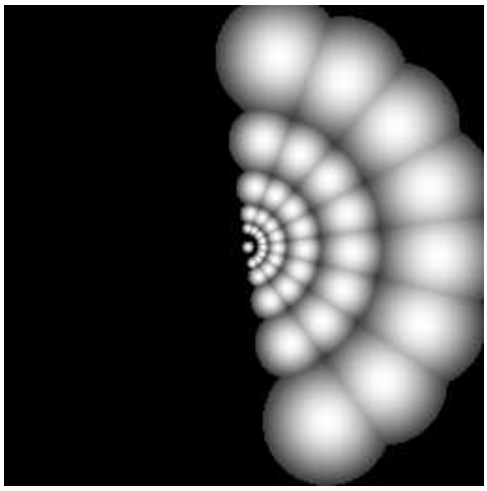


Figure 4: Frequency domain representation of the filter bank for choices of the design parameters $r_0 = 9.6$ cycles per image, $R = 1.8$, $B = 1$ octave, and $\eta = \frac{1}{2}$. There are 40 filters arranged in a polar wavelet-like tesselation on eight rays with five filters per ray, plus one filter centered at DC. Each of the 41 filters in the figure has been independently scaled for maximum dynamic range in the available grey levels.

has been *independently* scaled for maximum dynamic range of grey scales.

Design parameters for all 40 AC filters are given in Tables 1 and 2. The DC filter has the same space constant and bandwidths as filter zero, but has center frequency $\mathbf{\Omega}_{\mathrm{DC}} = 0$. Each row of the tables gives the parameters for one filter. Following is a description of the parameters:

Filter The number of the filter. Along each ray, filters numbers increase with increasing radial center frequency.

Ray The number of the ray the filter is on. Ray zero is at the top of Figure 4, and rays are numbered sequentially proceeding clockwise around the filterbank.

Indx on Ray Index of the filter on it's respective ray. The filter with index 0 on each ray is the one with lowest radial center frequency.

θ_m The orientation of the filter center frequency, $\mathbf{\Omega}_m$, in radians; $\theta_m = \arg[\mathbf{\Omega}_m]$.

r_m Radial center frequency in cycles per pixel; $r_m = |\mathbf{\Omega}_m|$.

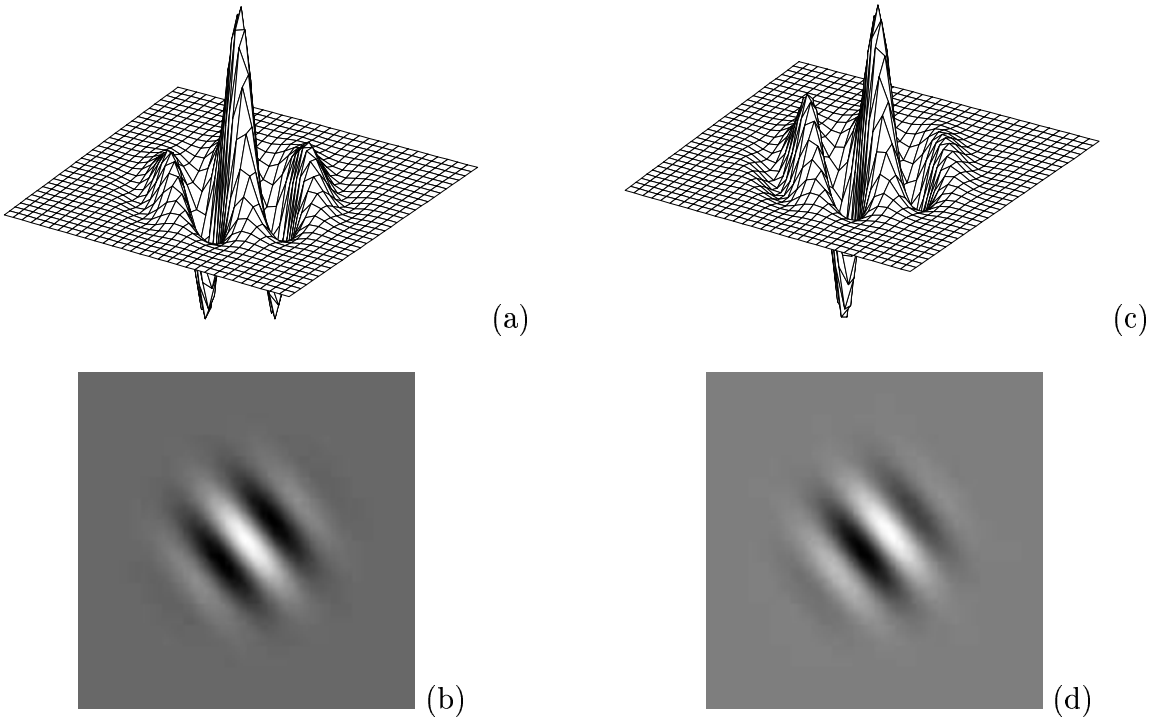


Figure 5: Space domain representation of Filter 10. (a) Real part plotted as a surface. (b) Real part depicted as a grey scale image. (c) Imaginary part plotted as a surface. (d) Imaginary part depicted as a grey scale image.

u_m Horizontal center frequency in cycles per pixel.

v_m Vertical center frequency in cycles per pixel.

σ_m Space constant in pixels.

max $|G(\Omega)|$ Peak value of the filter function in the frequency domain.

As a representative example, Figure 6 shows the real and imaginary parts of Filter 10 in the spatial domain.

7 Dominant Component Demodulation

If only a single component is present, and there is no noise, then the complex extended image $t(\mathbf{x})$ takes the form (1). Theoretically in this case, the filtered demodulation algorithm (29), (30) may be applied to the response of *any* channel in the filter bank to estimate

Filter	Ray	Indx on Ray	Orient θ_m (rad)	r_m $\left(\frac{\text{Hz}}{\text{pix}}\right)$	u_m $\left(\frac{\text{Hz}}{\text{pix}}\right)$	v_m $\left(\frac{\text{Hz}}{\text{pix}}\right)$	Space Const σ_m (pix)	max $\lvert G(\Omega)\rvert$
0	0	0	-1.3907	0.0375	0.0067	-0.0369	10.6004	53.1426
1	0	1	-1.3907	0.0675	0.0121	-0.0664	5.8891	29.5237
2	0	2	-1.3907	0.1215	0.0218	-0.1195	3.2717	16.4020
3	0	3	-1.3907	0.2187	0.0392	-0.2152	1.8176	9.1122
4	0	4	-1.3907	0.3937	0.0705	-0.3873	1.0098	5.0626
5	1	0	-1.0304	0.0375	0.0193	-0.0322	10.6004	53.1426
6	1	1	-1.0304	0.0675	0.0347	-0.0579	5.8891	29.5237
7	1	2	-1.0304	0.1215	0.0625	-0.1042	3.2717	16.4020
8	1	3	-1.0304	0.2187	0.1125	-0.1875	1.8176	9.1122
9	1	4	-1.0304	0.3937	0.2025	-0.3376	1.0098	5.0626
10	2	0	-0.6701	0.0375	0.0294	-0.0233	10.6004	53.1426
11	2	1	-0.6701	0.0675	0.0529	-0.0419	5.8891	29.5237
12	2	2	-0.6701	0.1215	0.0952	-0.0755	3.2717	16.4020
13	2	3	-0.6701	0.2187	0.1714	-0.1358	1.8176	9.1122
14	2	4	-0.6701	0.3937	0.3085	-0.2445	1.0098	5.0626
15	3	0	-0.3099	0.0375	0.0357	-0.0114	10.6004	53.1426
16	3	1	-0.3099	0.0675	0.0643	-0.0206	5.8891	29.5237
17	3	2	-0.3099	0.1215	0.1157	-0.0370	3.2717	16.4020
18	3	3	-0.3099	0.2187	0.2083	-0.0667	1.8176	9.1122
19	3	4	-0.3099	0.3937	0.3749	-0.1200	1.0098	5.0626

Table 1: Design parameters for filters 0 – 19 (see text for descriptions of the individual parameters).

Filter	Ray	Indx on Ray	Orient θ_m (rad)	r_m $\left(\frac{\text{Hz}}{\text{pix}}\right)$	u_m $\left(\frac{\text{Hz}}{\text{pix}}\right)$	v_m $\left(\frac{\text{Hz}}{\text{pix}}\right)$	Space Const σ_m (pix)	max $\lvert G(\Omega)\rvert$
20	4	0	0.0504	0.0375	0.0375	0.0019	10.6004	53.1426
21	4	1	0.0504	0.0675	0.0674	0.0034	5.8891	29.5237
22	4	2	0.0504	0.1215	0.1213	0.0061	3.2717	16.4020
23	4	3	0.0504	0.2187	0.2184	0.0110	1.8176	9.1122
24	4	4	0.0504	0.3937	0.3932	0.0198	1.0098	5.0626
25	5	0	0.4107	0.0375	0.0344	0.0150	10.6004	53.1426
26	5	1	0.4107	0.0675	0.0619	0.0269	5.8891	29.5237
27	5	2	0.4107	0.1215	0.1114	0.0485	3.2717	16.4020
28	5	3	0.4107	0.2187	0.2005	0.0873	1.8176	9.1122
29	5	4	0.4107	0.3937	0.3609	0.1572	1.0098	5.0626
30	6	0	0.7709	0.0375	0.0269	0.0261	10.6004	53.1426
31	6	1	0.7709	0.0675	0.0484	0.0470	5.8891	29.5237
32	6	2	0.7709	0.1215	0.0871	0.0847	3.2717	16.4020
33	6	3	0.7709	0.2187	0.1569	0.1524	1.8176	9.1122
34	6	4	0.7709	0.3937	0.2824	0.2743	1.0098	5.0626
35	7	0	1.1312	0.0375	0.0160	0.0339	10.6004	53.1426
36	7	1	1.1312	0.0675	0.0287	0.0611	5.8891	29.5237
37	7	2	1.1312	0.1215	0.0517	0.1100	3.2717	16.4020
38	7	3	1.1312	0.2187	0.0931	0.1979	1.8176	9.1122
39	7	4	1.1312	0.3937	0.1675	0.3562	1.0098	5.0626

Table 2: Design parameters for filters 20 – 39 (see text for descriptions of the individual parameters).

$a(\mathbf{x})$ and $\nabla\varphi(\mathbf{x})$. However, in practice we would always prefer to perform the demodulation using the response of the channel whose center frequency is closest to the instantaneous frequency $\nabla\varphi(\mathbf{x})$ on a spatially local basis, since this will tend to improve our estimates by affording us an improved signal to noise ratio and improved rejection against other out-of-band components. Then the demodulation problem becomes one of determining, at each pixel, which channel response to apply the demodulation algorithm to. This determination cannot be based on the magnitude of the channel responses alone, since all filters have unit L^2 norm, and hence the peak magnitude of the frequency response of the low frequency channels is much larger than that of the high frequency channels. Therefore, for each channel, we compute at each pixel a metric called the *filter selection criterion* defined by

$$\Psi_m(\mathbf{x}) = \frac{|t_{\sigma_m}(\mathbf{x})|}{\max_{\mathbf{\Omega}} |G_m(\mathbf{\Omega})|}. \tag{33}$$

At each pixel, we apply the filtered demodulation algorithm to the channel that maximizes the selection criterion (33). When only one component is present this technique will tend to perform the estimation using the channel with the best signal to noise ratio.

When multiple components are present, this technique will estimate the quantities $\nabla\varphi(\mathbf{x})$,$a(\mathbf{x})$ corresponding to the component that is dominant on a spatially local basis. The resulting frequency estimates are the emergent frequencies that dominate the local image spectrum. We refer to this as *dominant component analysis*. Figure 7 shows a block diagram of the overall paradigm. The filtered demodulation algorithm is applied in the blocks marked "ESA", which stands for *energy separation algorithm* and is used for historical reasons.

Figure 7 shows the result of applying dominant component analysis to two images. The (a) part of the figure shows a reasonably sophisticated example of a synthetically generated image component with Gaussian phase $\varphi(\mathbf{x}) = (\pi K\sigma^2/N)\exp\left[-\mathbf{x}^T\mathbf{x}/\sigma^2\right]$, where $K = 10$, $\sigma \approx 190$, and $N = 256$. The local instantaneous frequency estimates are shown in the (b) part of the figure, where arrow length is proportional to the instantaneous period, which is the reciprocal of the magnitude of the instantaneous frequency vector. The arrow heads point in the direction of $\nabla\varphi(\mathbf{x})$. With this display convention, shorter arrows correspond to higher frequencies and image features of smaller spatial extent, whereas longer arrows correspond to

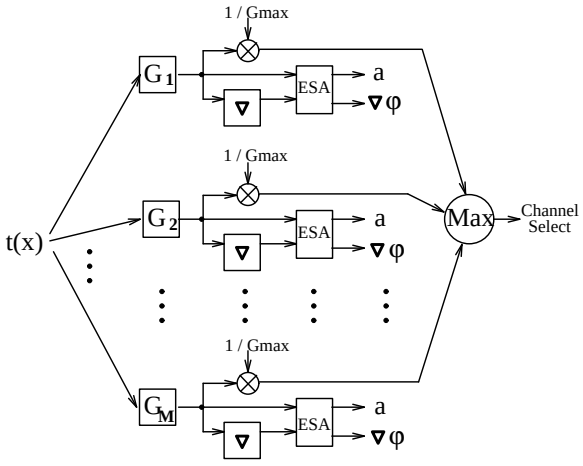


Figure 6: Block diagram of the dominant component paradigm. The filtered demodulation algorithm is performed in the blocks marked ESA. The magnitude of each channel response is divided by the peak value of the channel filter frequency response to obtain the filter selection criterion. On a pixel by pixel basis, the estimation is performed using the ESA results from the channel which maximizes this quantity.

lower spatial frequencies and image feature of larger spatial extent. The amplitude estimate is shown in the (c) part of the figure. Large values of amplitude correspond to regions of high contrast in the image. Compared to the rapid variations that characterize much of the image, both estimated modulating functions are smooth and slowly varying.

Figure 7 (d) shows the texture image raffia. The local instantaneous frequency estimates are given in the (e) part of the figure, where once again the arrow lengths are proportional to the instantaneous period, the arrow heads point in the direction of the instantaneous frequency vector, and only one needle is shown for each block of 144 pixels. Although this image clearly has a multi-partite nature, the strips oriented with the main antidiagonal are particularly strong throughout the image, and this is captured in the dominant component frequency estimates. The amplitude estimates are shown in the (f) part of the figure, where once again high amplitude values correspond to areas of high contrast in the image.

Figure 7 shows the dominant component analysis of two natural images, tree in the (a) part of the figure and flowers in the (d) part. The frequency estimates for tree are shown in the (b) part of the figure. Only one arrow is shown for each block of 16×16 pixels. For this figure, the lengths of the arrows have been squared for display to accentuate the

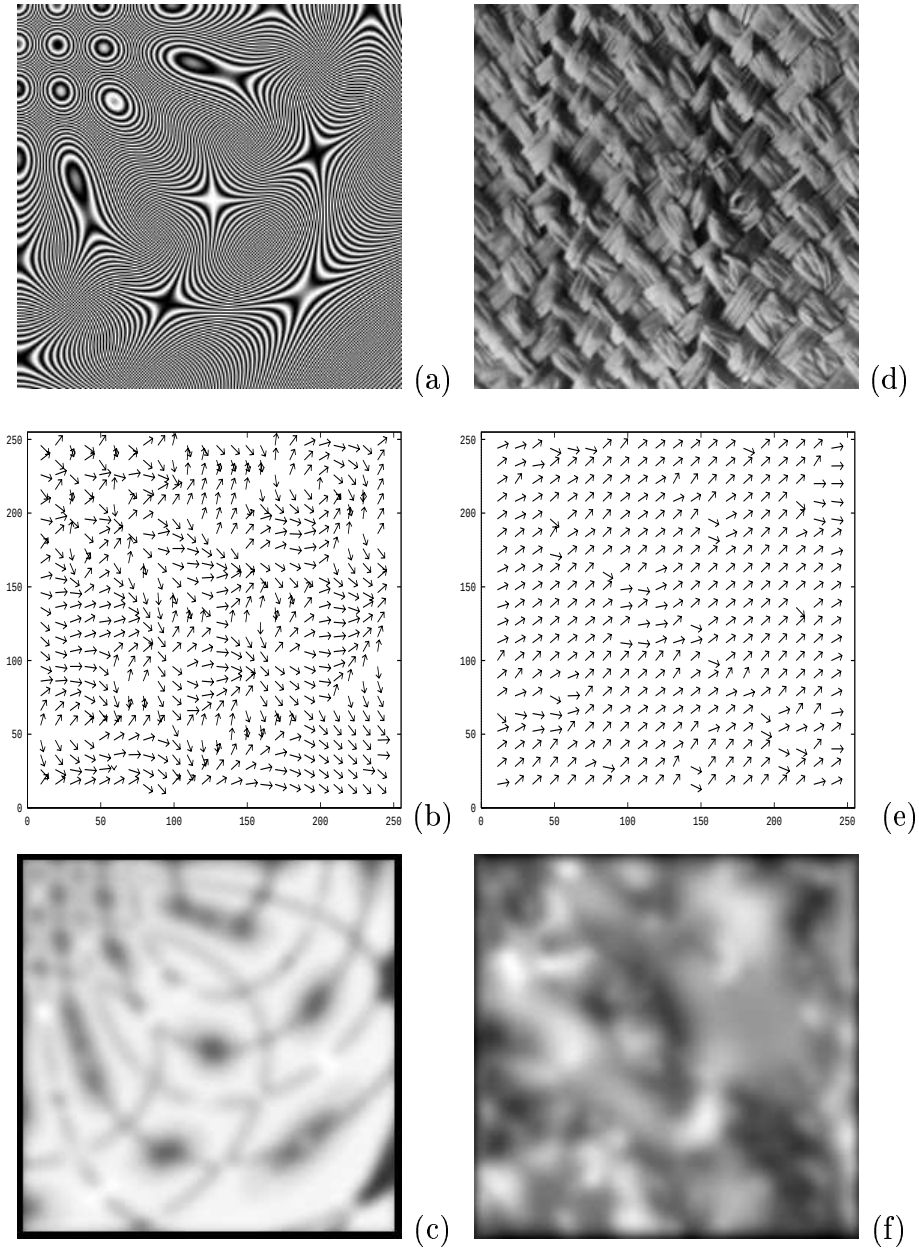


Figure 7: Application of the dominant component demodulation algorithm to the Gaussian and raffia images. (a) The Gaussian image, instantaneous phase is a Gaussian. (b) Needle diagram showing estimates of $\nabla\varphi(\mathbf{x})$ for the Gaussian image. The length of each needle is inversely proportional to to $|\nabla\varphi(\mathbf{x})|$ (proportional to period). (c) Amplitude estimates for the Gaussian image. (d) The raffia image. (e) Needle diagram showing estimates of $\nabla\varphi(\mathbf{x})$ for the raffia image. (f) Amplitude estimates for the raffia image.

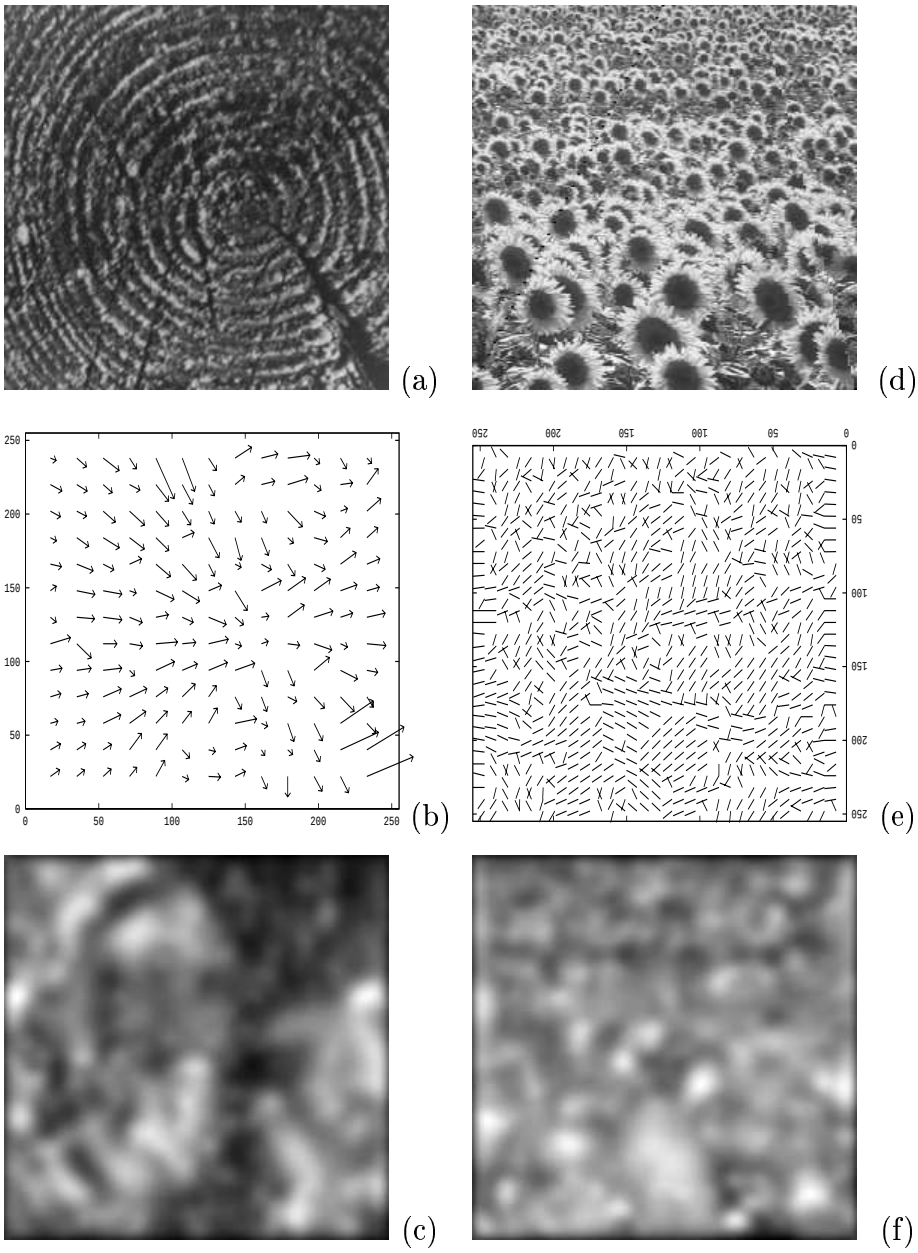


Figure 8: Application of the dominant component demodulation algorithm to the tree and flower images. (a) The tree image. (b) Needle diagram showing estimates of $\nabla\varphi(\mathbf{x})$ for the tree image. The length of each needle is inversely proportional to to $|\nabla\varphi(\mathbf{x})|$ (proportional to period). (c) Amplitude estimates for the tree image. (d) The flower image. (e) Needle diagram showing estimates of $\nabla\varphi(\mathbf{x})$ for the flower image. (f) Amplitude estimates for the flower image.

differences between highest and lowest frequency estimates, and as before shorter needles indicate higher frequencies which correspond to smaller spatial features in the image. While the tree appears to have orientations which could be interpreted as being in all four quadrants of the frequency plane, we see that the estimation algorithm places all arrows in quadrants I and IV. The slowly varying amplitude estimates are shown in the (c) part of the figure, where again we see high values for the amplitude where bright and dark features are juxtaposed in the image. The frequency estimates for the flower image are shown in the (e) part of the figure. In this case we have omitted the arrow heads in order to permit a denser display. We show one needle for every block of 64 pixels in the image, and needle length is directly proportional to the instantaneous period. Finally, the amplitude estimates for the flower image are shown in the (f) part of the figure.

8 Conclusion

In this paper we developed a technique for estimating the dominant instantaneous frequencies and amplitude modulations of an image modeled with AM-FM functions. The estimation is highly localized in space, and is capable of capturing local nonstationarities in the image. When applied to natural images, the technique produces frequency estimates that agree remarkably well with perception.

A unique definition of instantaneous frequency for real valued images does not exist. We argued that it makes sense to consider that the instantaneous frequency of a real image is the same as that of its associated complex analytic image, for which the instantaneous frequency is well defined, and showed that the analytic image can be conveniently generated using the relatively little known multidimensional Hilbert transform.

In general, the local spectrum of any sophisticated image is characterized by a distribution of frequencies that are present around the dominant, or *emergent* frequency. The dominant AM-FM component must be isolated by multiband filtering prior to demodulation. We described in detail the design of a suitable filterbank using Gabor wavelets, which are optimal in their conjoint uncertainty and hence optimize the error in the frequency and amplitude demodulation algorithms.

Characterization of the local instantaneous frequency is fundamentally important to an increasing body of problems in computer vision and image processing, including texture analysis and segmentation, shape from texture, and certain new techniques in texture and frequency based stereopsis. Future and ongoing work in this area concerns techniques for simultaneous demodulation of all the multiple components of a multi-partite image and reconstruction of the image from the estimated modulating functions.

Appendices

A Derivation of the 2D Hilbert Transform

Assume that we have a real image $s(\mathbf{x})$ which contains no Dirac deltas, and further that we wish to generate a complex extension $t(\mathbf{x}) = s(\mathbf{x}) + jw(\mathbf{x})$, $w(\mathbf{x}) \in \mathbb{R}$, in such a way that $\mathcal{F}[t(\mathbf{x})]$ is zero in quadrants II and III of the $\mathbf{\Omega} = [u, v]^T$ frequency plane and equal to $2\mathcal{F}[s(\mathbf{x})]$ in quadrants I and IV. Then $t(\mathbf{x})$ is the unique complex image that results from removing the redundant spectral information from $F[s(\mathbf{x})]$. Let $\mathbf{x} = [x, y]^T$ be a position vector in the spatial domain, $\mathbf{e}_x$ and $\mathbf{e}_y$ be unit vectors in the x and y directions, respectively, and $\mathbf{\Omega} = [u, v]^T$ be a position vector in the frequency domain. In this appendix we will show that the desired result is obtained when $w(\mathbf{x}) = \boldsymbol{\mathcal{H}}[s(\mathbf{x})]$, where $\boldsymbol{\mathcal{H}}[\cdot]$ is defined by (3), with $\mathbf{e}_i = \mathbf{e}_x$ and $\mathbf{e}_j = \mathbf{e}_y$.

Let $S_{CS}(\mathbf{\Omega}) = \mathcal{F}[s(\mathbf{x})]$ and $S_{CA}(\mathbf{\Omega}) = \mathcal{F}[jw(\mathbf{x})]$. From the properties of the two-dimensional Fourier transform, we have that $S_{CS}(\mathbf{\Omega})$ is conjugate symmetric and $S_{CA}(\mathbf{\Omega})$ is conjugate antisymmetric, or

$$S_{CS}(-\mathbf{\Omega}) = S^*_{CS}(\mathbf{\Omega}) \tag{34}$$

$$S_{CA}(-\mathbf{\Omega}) = -S^*_{CS}(\mathbf{\Omega}). \tag{35}$$

Now we have $T(\mathbf{\Omega}) = \mathcal{F}[s(\mathbf{x}) + jw(\mathbf{x})] = S_{CS}(\mathbf{\Omega}) + S_{CA}(\mathbf{\Omega})$. In order that $T(\mathbf{\Omega}) = 0$, $u < 0$, we must have that $S_{CA}(\mathbf{\Omega}) = -S_{CS}(\mathbf{\Omega})$ in quadrants II and III, as shown in Figure 9 (a). The conjugate antisymmetry of $S_{CA}(\mathbf{\Omega})$ then implies that $S_{CA}(\mathbf{\Omega}) = S^*_{CS}(-\mathbf{\Omega})$ in quadrants I and IV. From the conjugate symmetry of $S_{CS}(\mathbf{\Omega})$, we then have that $S_{CA}(u, v) = \text{sgn}\,[u] S_{CS}(u, v)$ on the entire $\mathbf{\Omega}$ plane, as depicted in Figure 9 (b). Then

$$\begin{aligned} jw(\mathbf{x}) &= \mathcal{F}^{-1}\left[\text{sgn}\,[u] S_{CS}(u, v)\right] \\ &= s(\mathbf{x}) * \mathcal{F}^{-1}[\text{sgn}\,[u]]. \end{aligned} \tag{36}$$

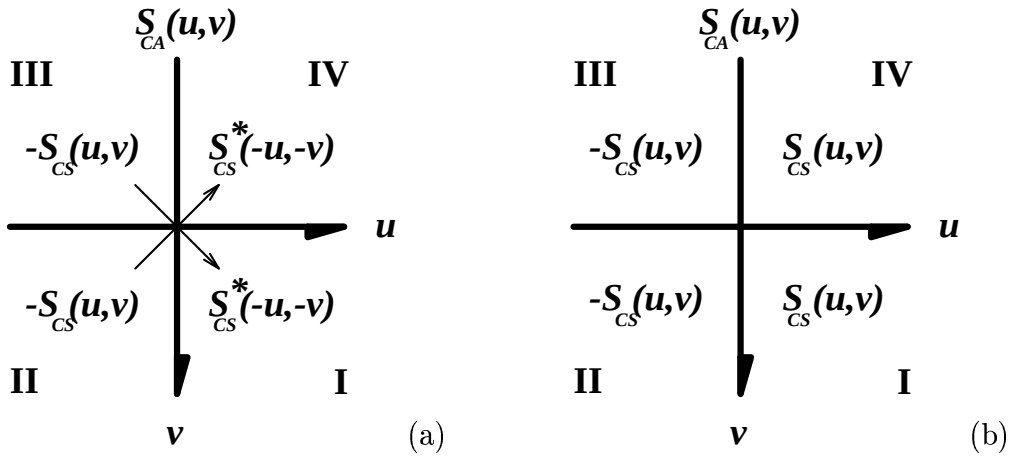


Figure 9: Relationship between the conjugate antisymmetric spectrum of $j\mathcal{H}[s(\mathbf{x})]$ and the conjugate symmetric spectrum of $s(\mathbf{x})$. (a) For cancellation between the spectra in the left-half plane, we must have $S_{CA}(\mathbf{\Omega}) = -S_{CS}(\mathbf{\Omega})$ in quadrants II and III. This implies that $S_{CA}(\mathbf{\Omega}) = S^*_{CS}(-\mathbf{\Omega})$ in quadrants I and IV. (b) The final relation, which results from the conjugate symmetry of $S_{CS}(\mathbf{\Omega})$.

Directly from the definition of the two-dimensional Fourier transform, we have

$$\begin{aligned}
\mathcal{F}^{-1}[\operatorname{sgn}[u]] &= \int_{\mathbb{R}^2} \operatorname{sgn}[u] \exp[j2\pi \mathbf{\Omega}^T \mathbf{x}] d\mathbf{\Omega} \\
&= \int_{\mathbb{R}} \exp[j2\pi vy] dy \int_{\mathbb{R}} \operatorname{sgn}[u] \exp[j2\pi ux] du \\
&= \frac{j\delta(y)}{\pi x}.
\end{aligned} \tag{37}$$

Substituting (37) into (36), we have

$$\begin{aligned}
\mathcal{H}[s(\mathbf{x})] = w(\mathbf{x}) &= s(\mathbf{x}) * \frac{j\delta(\mathbf{x}^T \mathbf{e}_y)}{\pi \mathbf{x}^T \mathbf{e}_x} \\
&= \frac{1}{\pi} \int_{\mathbb{R}^2} s(\mathbf{x} - \mathbf{p}) \frac{\delta(\mathbf{p}^T \mathbf{e}_y)}{\mathbf{x}^T \mathbf{e}_x} d\mathbf{p} \\
&= \frac{1}{\pi} \int_{\mathbb{R}} s(\mathbf{x} - p\mathbf{e}_x) \frac{dp}{p},
\end{aligned}$$

where the integral is interpreted as a Cauchy principle value. ■

B Proof of Theorem 1

Proof: $\mathcal{F}[a(\mathbf{x})\cos[\varphi(\mathbf{x})]] = A(\mathbf{\Omega}) * \Phi(\mathbf{\Omega})$, so the Fourier inversion integral gives us

$$\begin{aligned} a(\mathbf{x})\cos[\varphi(\mathbf{x})] &= \int_{\mathbb{R}^2}\left[\int_{\mathbb{R}^2} A(\mathbf{p})\Phi(\mathbf{\Omega}-\mathbf{p})d\mathbf{p}\right]\exp[j2\pi\mathbf{\Omega}^T\mathbf{x}]d\mathbf{\Omega} \\ &= \int_{\mathbb{R}^2} A(\mathbf{p})\int_{\mathbb{R}^2}\Phi(\mathbf{\Omega}-\mathbf{p})\exp[j2\pi\mathbf{\Omega}^T\mathbf{x}]d\mathbf{\Omega}d\mathbf{p} \\ &= \int_{\mathbb{R}^2} A(\mathbf{p})\int_{\mathbb{R}^2}\Phi(\mathbf{\Omega})\exp[j2\pi(\mathbf{\Omega}+\mathbf{p})^T\mathbf{x}]d\mathbf{\Omega}d\mathbf{p}. \qquad (38) \end{aligned}$$

Taking the Hilbert transform of the kernel of the Fourier inversion integral, we obtain

$$\begin{aligned} \boldsymbol{\mathcal{H}}[\exp[j2\pi\mathbf{\Omega}^T\mathbf{x}]] &= \frac{1}{\pi}\int_{\mathbb{R}}\exp[j2\pi\mathbf{\Omega}^T(\mathbf{x}-p\mathbf{e}_x)]\frac{dp}{p} \\ &= \frac{1}{\pi}\int_{\mathbb{R}^2}\exp[j2\pi\mathbf{\Omega}^T(\mathbf{x}-\mathbf{p})]\frac{\delta(\mathbf{p}^T\mathbf{e}_y)}{\mathbf{p}^T\mathbf{e}_x}d\mathbf{p} \\ &= \frac{1}{\pi}\int_{\mathbb{R}^2}\exp[j2\pi\mathbf{\Omega}^T\mathbf{p}]\frac{\delta[(\mathbf{x}-\mathbf{p})^T\mathbf{e}_y]}{(\mathbf{x}-\mathbf{p})^T\mathbf{e}_x}d\mathbf{p} \\ &= \frac{1}{\pi}\int_{\mathbb{R}^2}\exp[j2\pi(u\xi+v\zeta)]\frac{\delta(y-\zeta)}{x-\xi}d\xi d\zeta \\ &= \frac{1}{\pi}\int_{\mathbb{R}}\exp[j2\pi(u\xi+vy)]\frac{d\xi}{x-\xi} \\ &= \exp[j2\pi vy]\left[\frac{1}{\pi}\int_{\mathbb{R}}\exp[j2\pi u\xi]\frac{d\xi}{x-\xi}\right] \qquad (39) \\ &= -j\mathrm{sgn}\,[u]\exp[\mathbf{\Omega}^T\mathbf{x}], \qquad (40) \end{aligned}$$

where (40) is obtained by recognizing the integral in (39) as the one-dimensional Hilbert transform of $\exp[j2\pi ux]$. Next, we make use of (38) and (40) to compute the Hilbert transform of $a(\mathbf{x})\cos[\varphi(\mathbf{x})]$:

$$\begin{aligned} \boldsymbol{\mathcal{H}}[a(\mathbf{x})\cos[\varphi(\mathbf{x})]] &= \int_{\mathbb{R}^2} A(\mathbf{p})\int_{\mathbb{R}^2}\Phi(\mathbf{\Omega})\boldsymbol{\mathcal{H}}\left\{\exp[j2\pi(\mathbf{\Omega}+\mathbf{p})^T\mathbf{x}]\right\}d\mathbf{\Omega}d\mathbf{p} \\ &= \int_{\mathbb{R}^2}\int_{\mathbb{R}^2} -jA(\xi,\zeta)\Phi(u,v)\mathrm{sgn}\,[u+\xi]\exp[j2\pi\{(u+\xi)x+(v+\zeta)y\}]dudvd\xi d\zeta. \qquad (41) \end{aligned}$$

By hypothesis, the product $A(\xi,\zeta)\Phi(u,v)$ is supported only where $|\xi| > u_0$ and $|u| < u_0$, so $\mathrm{sgn}\,[u+\xi] = \mathrm{sgn}\,[u]$ everywhere that the integrand of (41) is nonzero. Hence

$$\begin{aligned}\boldsymbol{\mathcal{H}}\Big[a(\mathbf{x})\cos\left[\varphi(\mathbf{x})\right]\Big] &= \int_{\mathbb{R}^2} A(\mathbf{p})\exp[j2\pi\mathbf{p}^T\mathbf{e}_x]d\mathbf{p}\int_{\mathbb{R}^2} -j\mathrm{sgn}\,[u]\Phi(u,v)\exp[j2\pi(ux+vy)]dudv \\ &= a(\mathbf{x})\boldsymbol{\mathcal{H}}\Big[\cos\left[\varphi(\mathbf{x})\right]\Big], \qquad (42)\end{aligned}$$

where the last equality follows from (5) and the definition of the inverse Fourier transform. The required result is immediate upon substitution of (42) into (6). ■

C Proof of Theorem 2

We begin with a pair of lemmata that will be used in proving the main result.

C.1 Significant Lemmata

Lemma 1 *Let $g_\sigma : \mathbb{R}^n \to \mathbb{C}$ and $a : \mathbb{R}^n \to \mathbb{R}$ be such that a is once-differentiable and $\exists$ conjugate exponents p, q with $q < n$ and $1 < p < \frac{n}{n-1}$ for which $|x_i g_\sigma(\mathbf{x})|^p$ and $|\frac{\partial}{\partial x_i}a(\mathbf{x})|^q$ are integrable on $\mathbb{R}^n$ for $1 \le i \le n$. Then*

$$\varepsilon_{t,a}(\mathbf{x}) = \int_{\mathbb{R}^n} |g_\sigma(\mathbf{y})||a(\mathbf{x}-\mathbf{y}) - a(\mathbf{x})|d\mathbf{y} \le \left(\frac{q}{q-n}\right)\bar{\Delta}_i^{(p)}(g_\sigma)\bar{\delta}_i^{(q)}(a). \qquad (43)$$

Proof: We first express $a(\mathbf{x}-\mathbf{y}) - a(\mathbf{x})$ in a zeroth-order Taylor series to get

$$|a(\mathbf{x}-\mathbf{y}) - a(\mathbf{x})| = \left|\int_0^1 \mathbf{y}^T\nabla a(\mathbf{x}-s\mathbf{y})ds\right|, \qquad (44)$$

and hence

$$\begin{aligned}\varepsilon_{t,a}(\mathbf{x}) &= \int_{\mathbb{R}^n} |g_\sigma(\mathbf{y})|\left|\int_0^1 \mathbf{y}^T\nabla a(\mathbf{x}-s\mathbf{y})ds\right|d\mathbf{y} \\ &\le \int_{\mathbb{R}^n} |g_\sigma(\mathbf{y})|\int_0^1 \left|\mathbf{y}^T\nabla a(\mathbf{x}-s\mathbf{y})\right|dsd\mathbf{y} \\ &= \int_{\mathbb{R}^n} |g_\sigma(\mathbf{y})|\int_0^1 \left|\sum_i y_i\frac{\partial}{\partial x_i}a(\mathbf{x}-s\mathbf{y})\right|dsd\mathbf{y}\end{aligned}$$

$$
\begin{aligned}
&\leq \int_{\mathbb{R}^n} |g_\sigma(\mathbf{y})| \int_0^1 \sum_i \left| y_i \frac{\partial}{\partial x_i} a(\mathbf{x} - s\mathbf{y}) \right| ds d\mathbf{y} \\
&= \int_{\mathbb{R}^n} |g_\sigma(\mathbf{y})| \int_0^1 \sum_i |y_i| \left| \frac{\partial}{\partial x_i} a(\mathbf{x} - s\mathbf{y}) \right| ds d\mathbf{y} \\
&= \sum_i \int_0^1 \int_{\mathbb{R}^n} |g_\sigma(\mathbf{y})| |y_i| \left| \frac{\partial}{\partial x_i} a(\mathbf{x} - s\mathbf{y}) \right| d\mathbf{y} ds
\end{aligned}
$$

where the last equality results from Fubini's theorem. Application of Hölder's integral inequality with p, q conjugate exponents, $1 < p < \infty$, followed by the change of variable $\mathbf{u} = \mathbf{x} - s\mathbf{y}$ then yields

$$
\begin{aligned}
\varepsilon_{t,a}(\mathbf{x}) &\leq \sum_i \int_0^1 \left[\int_{\mathbb{R}^n} |g_\sigma(\mathbf{y})|^p |y_i|^p d\mathbf{y} \right]^{\frac{1}{p}} \left[\int_{\mathbb{R}^n} \left| \frac{\partial}{\partial x_i} a(\mathbf{x} - s\mathbf{y}) \right|^q d\mathbf{y} \right]^{\frac{1}{q}} ds \\
&= \sum_i \left[\int_{\mathbb{R}^n} |g_\sigma(\mathbf{y})|^p |y_i|^p d\mathbf{y} \right]^{\frac{1}{p}} \int_0^1 \left[\int_{\mathbb{R}^n} \left| \frac{\partial}{\partial x_i} a(\mathbf{x} - s\mathbf{y}) \right|^q d\mathbf{y} \right]^{\frac{1}{q}} ds \\
&= \sum_i \left[\int_{\mathbb{R}^n} |g_\sigma(\mathbf{y})|^p |y_i|^p d\mathbf{y} \right]^{\frac{1}{p}} \int_0^1 \left[s^{-n} \int_{\mathbb{R}^n} \left| \frac{\partial}{\partial x_i} a(\mathbf{y}) \right|^q d\mathbf{y} \right]^{\frac{1}{q}} ds \\
&= \sum_i \left[\int_{\mathbb{R}^n} |g_\sigma(\mathbf{y})|^p |y_i|^p d\mathbf{y} \right]^{\frac{1}{p}} \left[\int_{\mathbb{R}^n} \left| \frac{\partial}{\partial x_i} a(\mathbf{y}) \right|^q d\mathbf{y} \right]^{\frac{1}{q}} \int_0^1 s^{-\frac{n}{q}} ds,
\end{aligned}
$$

which converges only if $q > n$, in which case $1 < p < \frac{n}{n-1}$ and

$$
\int_0^1 s^{-\frac{n}{q}} ds = \left(\frac{q}{q-n} \right).
$$

Finally, upon application of Hölder's inequality for sums with p', q' conjugate exponents, we have

$$
\varepsilon_{t,a}(\mathbf{x}) \leq \left(\frac{q}{q-n} \right) \left[\sum_i \left(\int_{\mathbb{R}^n} |g_\sigma(\mathbf{y})|^p |y_i|^p d\mathbf{y} \right)^{\frac{p'}{p}} \right]^{\frac{1}{p'}} \left[\sum_i \left(\int_{\mathbb{R}^n} \left| \frac{\partial}{\partial x_i} a(\mathbf{y}) \right|^q d\mathbf{y} \right)^{\frac{q'}{q}} \right]^{\frac{1}{q'}}.
$$

The lemma is proved by choosing $p' = p$, $q' = q$, and substituting in the definitions (15), (16), (19), and (21). ■

Lemma 2 *Let $g_\sigma : \mathbb{R}^n \to \mathbb{C}$ and $\varphi : \mathbb{R}^n \to \mathbb{R}$ be such that φ is twice-differentiable and $\exists$ conjugate exponents p, q with $q < n$ and $1 < p < \frac{n}{n-1}$ for which $|x_i x_j g_\sigma(\mathbf{x})|^p$ and $|\frac{\partial^2}{\partial x_i \partial x_j} \varphi(\mathbf{x})|^q$*

are integrable on $\mathbb{R}^n$ *for* $1 \leq i, j \leq n$. *Then*

$$\varepsilon_{t,\varphi}(\mathbf{x}) = \int_{\mathbb{R}^n} |g_\sigma(\mathbf{y})||1 - \exp[jQ_\varphi(\mathbf{x}, \mathbf{y})]|d\mathbf{y} \leq \frac{q^2}{(q-n)(2q-n)} \bar{\Delta}_{i,j}^{(p)}(g_\sigma)\bar{\delta}_{i,j}^{(q)}(\varphi) \tag{45}$$

where

$$Q_\varphi(\mathbf{x}, \mathbf{p}) = \int_0^1 (1-s) \sum_{i,j} p_i p_j \frac{\partial^2}{\partial x_i \partial x_j} f(\mathbf{x} + s\mathbf{p}) ds. \tag{46}$$

Proof: Applying the triangle inequality and the fact that $|1 - e^{jx}| \leq |x| \ \forall x \in \mathbb{R}$, we have

$$\begin{aligned}
\varepsilon_{t,\varphi}(\mathbf{x}) &\leq \int_{\mathbb{R}^n} |g_\sigma(\mathbf{y})||Q_\varphi(\mathbf{x}, \mathbf{y})|d\mathbf{y} \\
&= \int_{\mathbb{R}^n} |g_\sigma(\mathbf{y})| \left| \sum_{i,j} y_i y_j \int_0^1 (1-s) \frac{\partial^2}{\partial x_i \partial x_j} \varphi(\mathbf{x} - s\mathbf{y}) ds \right| d\mathbf{y} \\
&\leq \int_{\mathbb{R}^n} |g_\sigma(\mathbf{y})| \sum_{i,j} |y_i y_j| \left| \int_0^1 (1-s) \frac{\partial^2}{\partial x_i \partial x_j} \varphi(\mathbf{x} - s\mathbf{y}) ds \right| d\mathbf{y} \\
&\leq \int_{\mathbb{R}^n} |g_\sigma(\mathbf{y})| \sum_{i,j} |y_i y_j| \int_0^1 \left| (1-s) \frac{\partial^2}{\partial x_i \partial x_j} \varphi(\mathbf{x} - s\mathbf{y}) \right| ds d\mathbf{y},
\end{aligned}$$

whereupon application of Fubini's theorem followed by Hölder's integral inequality with p, q conjugate exponents, $1 < p < \infty$, we see that

$$\begin{aligned}
\varepsilon_{t,\varphi}(\mathbf{x}) &\leq \sum_{i,j} \int_0^1 (1-s) \int_{\mathbb{R}^n} |y_i y_j| |g_\sigma(\mathbf{y})| \left| \frac{\partial^2}{\partial x_i \partial x_j} \varphi(\mathbf{x} - s\mathbf{y}) \right| d\mathbf{y} ds \\
&\leq \sum_{i,j} \int_0^1 (1-s) \left[\int_{\mathbb{R}^n} |y_i y_j|^p |g_\sigma(\mathbf{y})|^p d\mathbf{y} \right]^{\frac{1}{p}} \left[\int_{\mathbb{R}^n} \left| \frac{\partial^2}{\partial x_i \partial x_j} \varphi(\mathbf{x} - s\mathbf{y}) \right|^q d\mathbf{y} \right]^{\frac{1}{q}} ds.
\end{aligned}$$

Making the change of variable $\mathbf{u} = \mathbf{x} - s\mathbf{y}$,

$$\begin{aligned}
\varepsilon_{t,\varphi}(\mathbf{x}) &\leq \sum_{i,j} \int_0^1 (1-s) \left[\int_{\mathbb{R}^n} |y_i y_j|^p |g_\sigma(\mathbf{y})|^p d\mathbf{y} \right]^{\frac{1}{p}} \left[s^{-n} \int_{\mathbb{R}^n} \left| \frac{\partial^2}{\partial x_i \partial x_j} \varphi(\mathbf{y}) \right|^q d\mathbf{y} \right]^{\frac{1}{q}} ds \\
&= \sum_{i,j} \left[\int_{\mathbb{R}^n} |y_i y_j|^p |g_\sigma(\mathbf{y})|^p d\mathbf{y} \right]^{\frac{1}{p}} \left[\int_{\mathbb{R}^n} \left| \frac{\partial^2}{\partial x_i \partial x_j} \varphi(\mathbf{y}) \right|^q d\mathbf{y} \right]^{\frac{1}{q}} \int_0^1 (1-s) s^{-\frac{n}{q}} ds,
\end{aligned}$$

which converges only if $q > n$, in which case $1 < p < \frac{n}{n-1}$ and

$$\int_0^1 (1-s) s^{-\frac{n}{q}} ds = \frac{q^2}{(q-n)(2q-n)}.$$

Application of Hölder's inequality for sums with p', q' conjugate exponents then yields

$$\varepsilon_{t,\varphi}(\mathbf{x}) \leq \frac{q^2}{(q-n)(2q-n)} \left[\sum_{i,j} \left(\int_{\mathbb{R}^n} |y_i y_j|^p |g_\sigma(\mathbf{y})|^p d\mathbf{y} \right)^{\frac{p'}{p}} \right]^{\frac{1}{p'}} \left[\sum_{i,j} \left(\int_{\mathbb{R}^n} \left| \frac{\partial^2}{\partial x_i \partial x_j} \varphi(\mathbf{y}) \right|^q d\mathbf{y} \right)^{\frac{q'}{q}} \right]^{\frac{1}{q'}}.$$

The proof is completed by choosing $p' = p$, $q' = q$, and simplifying with the definitions (17), (18), (20), and (22). ■

C.2 Proof of Main Result

Proof: From (23) and (24), we have by expanding φ in a first-order Taylor series that

$$\begin{aligned} t_\sigma(\mathbf{x}) &= \int_{\mathbb{R}^n} a(\mathbf{x}-\mathbf{y}) \exp[j\varphi(\mathbf{x}-\mathbf{y})] g_\sigma(\mathbf{y}) d\mathbf{y} \\ &= \int_{\mathbb{R}^n} g_\sigma(\mathbf{y}) a(\mathbf{x}-\mathbf{y}) \exp[j\varphi(\mathbf{x})] \exp[j\mathbf{y}^T \nabla\varphi(\mathbf{x})] \exp[jQ_\varphi(\mathbf{x},\mathbf{y})] d\mathbf{y}, \end{aligned} \tag{47}$$

while combining (23) and (25) we have that

$$\begin{aligned} \widehat{t}_\sigma(\mathbf{x}) &= a(\mathbf{x}) \exp[j\varphi(\mathbf{x})] \int_{\mathbb{R}^n} g_\sigma(\mathbf{y}) \exp[-j\mathbf{y}^T \nabla\varphi(\mathbf{x})] d\mathbf{y} \\ &= \int_{\mathbb{R}^n} g_\sigma(\mathbf{y}) a(\mathbf{x}) \exp[j\varphi(\mathbf{x})] \exp[-j\mathbf{y}^T \nabla\varphi(\mathbf{x})] d\mathbf{y}. \end{aligned} \tag{48}$$

By (26), (47), and (48),

$$\begin{aligned} \varepsilon_t(\mathbf{x}) &= \left| \int_{\mathbb{R}^n} g_\sigma(\mathbf{y}) \exp[j\varphi(\mathbf{x})] \exp[-j\mathbf{y}^T \nabla\varphi(\mathbf{x})] \{a(\mathbf{x}-\mathbf{y}) \exp[jQ_\varphi(\mathbf{x},\mathbf{y})] - a(\mathbf{x})\} d\mathbf{y} \right| \\ &\leq \int_{\mathbb{R}^n} \left| g_\sigma(\mathbf{y}) \exp[j\varphi(\mathbf{x})] \exp[-j\mathbf{y}^T \nabla\varphi(\mathbf{x})] \{a(\mathbf{x}-\mathbf{y}) \exp[jQ_\varphi(\mathbf{x},\mathbf{y})] - a(\mathbf{x})\} \right| d\mathbf{y} \\ &= \int_{\mathbb{R}^n} |g_\sigma(\mathbf{y})| |a(\mathbf{x}-\mathbf{y}) \exp[jQ_\varphi(\mathbf{x},\mathbf{y})] - a(\mathbf{x})| d\mathbf{y}. \end{aligned} \tag{49}$$

Now

$$\begin{aligned} |a(\mathbf{x}-\mathbf{y}) \exp[jQ_\varphi(\mathbf{x},\mathbf{y})] - a(\mathbf{x})| &= |a(\mathbf{x}-\mathbf{y}) \exp[jQ_\varphi(\mathbf{x},\mathbf{y})] - a(\mathbf{x}) + a(\mathbf{x}-\mathbf{y}) - a(\mathbf{x}-\mathbf{y})| \\ &= |\{a(\mathbf{x}-\mathbf{y}) - a(\mathbf{x})\} + \{a(\mathbf{x}-\mathbf{y}) \exp[jQ_\varphi(\mathbf{x},\mathbf{y})] - a(\mathbf{x}-\mathbf{y})\}| \\ &\leq |a(\mathbf{x}-\mathbf{y}) - a(\mathbf{x})| + |a(\mathbf{x}-\mathbf{y})| |1 - \exp[jQ_\varphi(\mathbf{x},\mathbf{y})]| \end{aligned}$$

$$\leq \quad |a(\mathbf{x}-\mathbf{y}) - a(\mathbf{x})| + |1 - \exp[jQ_\varphi(\mathbf{x},\mathbf{y})]|,$$

so

$$\begin{aligned}\varepsilon_t(\mathbf{x}) &\leq \int_{\mathbb{R}^n} |g_\sigma(\mathbf{y})||1-\exp[jQ_\varphi(\mathbf{x},\mathbf{y})]|d\mathbf{y} + \int_{\mathbb{R}^n} |g_\sigma(\mathbf{y})||a(\mathbf{x}-\mathbf{y})-a(\mathbf{x})|d\mathbf{y} \\ &= \varepsilon_{t,a}(\mathbf{x}) + \varepsilon_{t,\varphi}(\mathbf{x}),\end{aligned}$$

and the theorem is proved. ■

D Proof of Corollary 1

Proof: By direct evaluation we have that

$$\begin{aligned}t_\sigma^{(i)}(\mathbf{x}) &= g_\sigma(\mathbf{x}) * t^{(i)}(\mathbf{x}) \\ &= a_+^{(i)}(\mathbf{x})\exp[j\varphi(\mathbf{x})] * g_\sigma(\mathbf{x}) - a_-^{(i)}(\mathbf{x})\exp[j\varphi(\mathbf{x})] * g_\sigma(\mathbf{x}) \\ &\quad + ja\varphi_+^{(i)}(\mathbf{x})\exp[j\varphi(\mathbf{x})] * g_\sigma(\mathbf{x}) - ja\varphi_-^{(i)}(\mathbf{x})\exp[j\varphi(\mathbf{x})] * g_\sigma(\mathbf{x}). \qquad (50)\end{aligned}$$

The result is immediate upon application of Theorem 2 to each term of (50). ■

References

[1] P. Maragos, J. F. Kaiser, and T. F. Quatieri, "Energy separation in signal modulations with application to speech analysis", Tech. Rept. 91-17, Harvard Robotics Laboratory, Harvard University, November 1991.

[2] P. Maragos, T. F. Quatieri, and J. F. Kaiser, "Speech nonlinearities, modulations, and energy operators", in *Proc. IEEE Int'l. Conf. Acoust., Speech, Signal Proc.*, Toronto, Ont., Canada, May 1991, pp. 421–424.

[3] 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.

[4] A. C. Bovik, P. Maragos, and T. F. Quatieri, "Measuring amplitude and frequency modulations in noise using multiband energy operators", in *Proc. IEEE Int'l. Symp. Time-Frequency and Time-Scale Analysis*, Victoria, B.C., Canada, October 4-6 1992.

[5] 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.

[6] P. Maragos, T. F. Quatieri, and J. F. Kaiser, "On amplitude and frequency demodulation using energy operators", in *Proc. IEEE Int'l. Conf. Acoust., Speech, Signal Proc.*, San Francisco, CA, March 1992, pp. II1–II4.

[7] A. C. Bovik, P. Maragos, and T. F. Quatieri, "Demodulation of AM-FM signals in noise using multiband energy operators", in *Proc. IEEE Int'l. Symp. Info. Theory*, San Antonio, TX, January 17-22 1993.

[8] A. C. Bovik, P. Maragos, and T. F. Quatieri, "AM-FM energy detection and separation in noise using multiband energy operators", *IEEE. Trans. Signal Proc.*, vol. SP-41, no. 12, pp. 3245–3265, December 1993.

[9] P. Maragos, J. F. Kaiser, and T. F. Quatieri, "Energy separation in signal modulations with applications to speech analysis", *IEEE. Trans. Signal Proc.*, vol. SP-41, no. 10, pp. 3024–3051, October 1993.

[10] P. Maragos, J. F. Kaiser, and T. F. Quatieri, "On amplitude and frequency demodulation using energy operators", *IEEE Trans. Signal Proc.*, vol. SP-41, no. 4, pp. 1532–1550, April 1993.

[11] 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, pp. I41 – I45.

[12] H. Knutsson, C.F. Westin, and G. Granlund, "Local multiscale frequency and bandwidth estimation", in *Proc. IEEE Int'l. Conf. Image Proc.*, Austin, TX, 1994, pp. I36–I40.

[13] 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, pp. III421–III425.

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

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

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

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

[18] A. C. Bovik, "Variational pattern analysis using Gabor wavelets", in *Proc. IEEE Int'l. Conf. Acoust., Speech, Signal Proc.*, San Francisco, CA, March 1992, pp. IV669 – IV672.

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

[20] 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, Novermber 1991.

[21] 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.

[22] R. M. Lerner, "A matched filter detection system for complicated doppler shifted signals", *IRE Trans. Info. Theo.*, vol. IT-6, pp. 373–385, June 1960.

[23] E. Bedrosian, "A product theorem for Hilbert transforms", *Proc. IEEE*, vol. 51, pp. 868–869, May 1963.

[24] A. W. Rihaczek, "Hilbert transforms and the complex representation of real signals", *Proc. IEEE*, vol. 54, pp. 434–435, March 1966.

[25] A. H. Nuttall, "On the quadrature approximation to the hilbert transform of modulated signals", *Proc. IEEE*, vol. 54, pp. 1458–1459, October 1966.

[26] Alexandros Potamianos and Petros Maragos, "A comparison of the energy operator and the Hilbert transform approach to signal and speech demodulation", Tech. Rept. 92-8, Harvard Robotics Laboratory, Harvard University, July 1992.

[27] A. Reilly, G. Frazer, and B. Boashash, "Analytic signal generation – tips and traps", *IEEE. Trans. Signal Proc.*, vol. SP-42, no. 11, pp. 3241–3245, November 1994.

[28] A. Papoulis, *Systems and Transforms with Applications in Optics*, McGraw-Hill, New York, 1968.

[29] E. M. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton University Press, Princeton, NJ, 1970.

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

[31] L. Cohen and C. Lee, "Instantaneous bandwidth for signals and spectrogram", in *Proc. IEEE Int'l. Conf. Acoust., Speech, Signal Proc.*, Albuquerque, NM, April 1990, pp. 2451–2454.

[32] A. C. Bovik, "A bound involving n-dimensional instantaneous frequency", *IEEE. Trans. Circ. Syst.*, vol. CAS-38, no. 11, pp. 1389–1390, November 1991.

[33] A. C. Bovik, J. P. Havlicek, and M. D. Desai, "Theorems for discrete filtered modulated signals", in *Proc. IEEE Int'l. Conf. Acoust., Speech, Signal Proc.*, Minneapolis, MN, April 27 - 30 1993, pp. III153–III156.