UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

MODULATION DOMAIN IMAGE PROCESSING

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

CHUONG T. NGUYEN
Norman, Oklahoma
2012

MODULATION DOMAIN IMAGE PROCESSING

A DISSERTATION APPROVED FOR THE
SCHOOL OF ELECTRICAL AND COMPUTER ENGINEERING

BY

Joseph P. Havlicek, Chair

J. R. Cruz

Choon Yik Tang

Murad Özaydin

Tomasz Przebinda

To my late loving grandmother.

Acknowledgements

I would like to express my gratitude to five committee members for serving and for offering technical expertise during my studies at the University of Oklahoma. Especially, I would like to honor my academic advisor, Dr. Joebob Havlicek. He has made tremendous impact on my research and professional career. Not only he is a teacher, but also is a friend. His thorough guidance transforms me to a better thinker. I am grateful to him for providing me opportunities to work on interesting research problems and for sponsoring me throughout my graduate study. More importantly, he taught me the phase "quality control".

I have too many friends to thank for. I would like to thank my lab members at the Intelligent Transportation System Laboratory. I have learned many lessons from them, both academic and life. I also want to thank all my friends of the Society of Vietnamese Student (SVS) and the SVS soccer team. They have been a family to me when I am far away from home.

Last but not least, I want to show my appreciation to my parents and brothers. I would not be in this position today without the support and patience of my wife, Anh Nguyen. Anh and my son, Duy, have been my inspiration.

Table of Contents

List of Tables

List of Figures

Abstract

MODULATION DOMAIN IMAGE PROCESSING

Chuong T. Nguyen, Ph.D.
The University of Oklahoma, 2012

Supervisor: Joseph P. Havlicek

The classical Fourier transform is the cornerstone of traditional linear signal and image processing. The discrete Fourier transform (DFT) and the fast Fourier transform (FFT) in particular led to profound changes during the later decades of the last century in how we analyze and process 1D and multi-dimensional signals. The Fourier transform represents a signal as an infinite superposition of stationary sinusoids each of which has constant amplitude and constant frequency. However, many important practical signals such as radar returns and seismic waves are inherently nonstationary. Hence, more complex techniques such as the windowed Fourier transform and the wavelet transform were invented to better capture nonstationary properties of these signals. In this dissertation, I studied an alternative nonstationary representation for images, *viz.*, the 2D AM-FM model. In contrast to the stationary nature of the classical Fourier representation, the AM-FM model represents an image as a finite sum of smoothly varying amplitudes and smoothly varying frequencies. The model has been applied successfully in image processing applications such as image segmentation, texture analysis, and target tracking. However, these applications are limited to *analysis*, meaning that the computed

AM and FM functions are used as features for signal processing tasks such as classification and recognition. For synthesis applications, few attempts have been made to synthesize the original image from the AM and FM components. Nevertheless, these attempts were unstable and the synthesized results contained artifacts. The main reason is that the perfect reconstruction AM-FM image model was either unavailable or unstable. Here, I constructed the first functional perfect reconstruction AM-FM image transform that paves the way for AM-FM image synthesis applications. The transform enables intuitive nonlinear image filter designs in the modulation domain. I showed that these filters provide important advantages relative to traditional linear translation invariant filters.

This dissertation addresses image processing operations in the nonlinear nonstationary modulation domain. In the modulation domain, an image is modeled as a sum of nonstationary amplitude modulation (AM) functions and nonstationary frequency modulation (FM) functions. I developed a theoretical framework for high fidelity signal and image modeling in the modulation domain, constructed an invertible multi-dimensional AM-FM transform (xAMFM), and investigated practical signal processing applications of the transform. After developing the xAMFM, I investigated new image processing operations that apply directly to the transformed AM and FM functions in the modulation domain. In addition, I introduced two classes of modulation domain image filters. These filters produce perceptually motivated signal processing results that are difficult or impossible to obtain with traditional linear processing or spatial domain nonlinear approaches. Finally, I proposed three extensions of the AM-FM transform and applied them in image analysis applications.

The main original contributions of this dissertation include the following.

- I proposed a perfect reconstruction FM algorithm. I used a least-squares approach to recover the phase signal from its gradient. In order to allow perfect reconstruction of the phase function, I enforced an initial condition on the reconstructed phase. The perfect reconstruction FM algorithm plays a critical role in the overall AM-FM transform.

- I constructed a perfect reconstruction multi-dimensional filterbank by modifying the classical steerable pyramid. This modified filterbank ensures a true multi-scale multi-orientation signal decomposition. Such a decomposition is required for a perceptually meaningful AM-FM image representation.

- I rotated the partial Hilbert transform to alleviate rippling artifacts in the computed AM and FM functions. This adjustment results in artifact free filtering results in the modulation domain.

- I proposed the modulation domain image filtering framework. I constructed two classes of modulation domain filters. I showed that the modulation domain filters outperform traditional linear shift invariant (LSI) filters qualitatively and quantitatively in applications such as selective orientation filtering, selective frequency filtering, and fundamental geometric image transformations.

- I provided extensions of the AM-FM transform for image decomposition problems. I illustrated that the AM-FM approach can successfully decompose an image into coherent components such as texture and structural components.

- I investigated the relationship between the two prominent AM-FM computational models, namely the partial Hilbert transform approach (pHT) and

the monogenic signal. The established relationship helps unify these two AM-FM algorithms.

This dissertation lays a theoretical foundation for future nonlinear modulation domain image processing applications. For the first time, one can apply modulation domain filters to images to obtain predictable results. The design of modulation domain filters is intuitive and simple, yet these filters produce superior results compared to those of pixel domain LSI filters. Moreover, this dissertation opens up other research problems. For instance, classical image applications such as image segmentation and edge detection can be re-formulated in the modulation domain setting. Modulation domain based perceptual image and video quality assessment and image compression are important future application areas for the fundamental representation results developed in this dissertation.

Chapter 1

Introduction

The classical Fourier transform is the cornerstone of traditional linear signal and image processing. The fast Fourier transform (FFT) in particular has changed the way we perform signal processing in the last decades of the last century. The Fourier transform represents a signal as a superposition of sinusoids each of which has constant amplitude and constant frequency. These amplitudes and frequencies capture important signal characteristics that are not trivial to obtain from time domain analysis. I will refer to such sinusoids as *stationary* in the sense that they admit amplitudes and frequencies that do not vary across the domain. However, with the Fourier representation, *nonstationary* signal structures can only be obtained by complicated constructive and destructive interferences between stationary Fourier components. For many important practical signals such as radar returns and seismic waves, this stationary representation is counterintuitive. The signal of interest in these applications are nonstationary. Such signals have been treated with advanced techniques designed to capture nonstationary structure such as the windowed Fourier transform and the wavelet transform. In this dissertation, I develop an alternative nonstationary representation for images, *viz.*, the 2D AM-FM model. In contrast to the stationary nature of the classical Fourier representation, the AM-FM model represents an image as a finite sum of smoothly varying amplitudes and smoothly varying frequencies.

This dissertation addresses image processing operations in the nonlinear nonstationary modulation domain. In the modulation domain, an image is modeled as a sum of nonstationary amplitude modulation (AM) functions and nonstationary frequency modulation (FM) functions. In particular, I investigated the theoretical construction and the applications of the AM-FM image transform. First, I developed the functional perfect reconstruction AM-FM transform. Second, I proposed a new image processing framework where filters are designed to operate directly on AM and FM functions. In addition, I introduced two classes of modulation domain image filters. These filters produce perceptually motivated results that are difficult or impossible to obtain with traditional liner and nonlinear filters. Finally, I investigated three derivative works of the AM-FM transform for image analysis applications.

In 1990, Bovik, Clark, and Geisler [13] first proposed an AM-FM image model to analyze textured regions. Since then, the model has been applied successfully in many image processing and computer vision applications such as image segmentation [17, 129], image inpainting [1], fingerprint analysis [63], texture analysis [49, 96], target tracking [86, 90], and biomedical imaging [70]. However, most of these applications are limited to *analysis*, meaning that the computed AM and FM functions are used as features for other signal processing tasks such as classification and recognition. So far, two major attempts have been made to synthesize the original image from the AM and FM functions. Havlicek, Harding, and Bovik [48] proposed an algorithm to reconstruct the dominant component of an image from the dominant AM and FM functions. The authors used a phase difference integration scheme to reconstruct the phase function from the computed FM functions. However, the reconstructed results contain artifacts and errors. Sivley

and Havlicek [109] introduced the first perfect reconstruction AM-FM transform. While the proposed transform is invertible, it is not suitable for practical AM-FM image synthesis applications. The transform requires more than 400 AM-FM components; and the FM reconstruction algorithm is not stable if the FM functions are altered. In this dissertation, I created the first functional perfect reconstruction AM-FM image transform that paves the way for image synthesis applications and a general theory of signal processing the in modulation domain. The AM-FM transform consists of a forward transform using a modified steerable pyramid decomposition and a backward transform using a least-squares phase integration. The transform enables us to design modulation domain image filters and apply them to process the nonstationary amplitude and frequency modulations of an image to realize perceptually motivated signal processing operations.

The main original contributions of this dissertation include the following.

- I propose a least-squares perfect reconstruction FM algorithm that to recover the phase signal from the computed gradient. In order to allow perfect reconstruction of the phase, I enforce an initial condition on the reconstructed phase. This perfect reconstruction FM algorithm plays a critical role in the overall AM-FM transform.

- I construct a modified perfect reconstruction filterbank based on the classical steerable pyramid [38, 106]. The modified filterbank ensures a true multi-scale multi-orientation signal decomposition. Such a decomposition is an important consideration for obtaining a perceptually meaningful AM-FM image representation.

- I adjust the rotation of the partial Hilbert transform axis of action so that the

computed AM and FM functions are artifact free from orientation induced artifacts. This adjustment results in artifact free modulation domain filtering results for the first time.

- I propose the first practical framework for designing filters and performing signal processing directly in the modulation domain. I construct two new classes of practical modulation domain filters. I show that these modulation domain filters outperform traditional linear shift invariant (LSI) filters qualitatively and quantitatively in applications such as selective orientation filtering, selective frequency filtering, image fusion, and fundamental image geometric transformations.

- I provide extensions of the AM-FM transform applicable to certain important image decomposition problems. I illustrate that the AM-FM approach can successfully decompose an image into coherent components such as texture and structural components.

- I investigate the relationship between the two most current AM-FM computational models, namely the partial Hilbert transform approach (pHT) [50, 97] and the monogenic signal [34,67]. The results provide important perspectives on the similarities and differences between the two approaches and suggest powerfully how they can be unified.

This dissertation lays a theoretical foundation for future nonlinear modulation domain image processing applications. For the first time, one can apply modulation domain filters to images to obtain predictable results. The design of modulation domain filters is intuitive and simple, yet these filters produce superior results compared to those of pixel domain LSI filters.

Besides modulation domain image processing, this dissertation opens up other research problems. For instance, classical image applications such as image segmentation and edge detection can be re-investigated in the modulation domain context. In addition, as the AM-FM transform admits important characteristics including perceptual relevance and prefect reconstruction, it can potentially yield performance gains in image and video quality assessment [12] and coding [68].

1.1 Organization

The dissertation contains eight chapters and two appendices.

In Chapter 2, I introduce the 1D AM-FM signal model as a viable representation for nonstationary signals. I discuss the limitations of the classical Fourier representation in cases where the signals of interest are nonstationary. I review practical applications where important nonstationary features of signals can be captured by the instantaneous frequency (IF), here interpreted as frequency modulation (FM). I then discuss three major computational techniques for obtaining the AM and FM for 1D signals, namely the Gabor analytic signal approach (AS) [41], the Teager-Kaiser energy operator (TKEO) approach [16, 79], and the quasi-local method (QL) [43].

In Chapter 3, I explore extensions of the 1D AM-FM signal model into multiple dimensions, with an emphasis on the 2D case. I first reason that the phase and instantaneous frequency play an important role in many image processing applications, such as image restoration, image segmentation, and optical flow estimation. I then discuss the generalization of the 1D AM-FM approaches into 2D. Finally, I evaluate their effectiveness with respect to mean squared error (MSE) via simulations using both synthetic and real images.

In Chapter 4, I introduce the single component perfect reconstruction AM-FM image model. I argue that most previous AM-FM image processing techniques were limited to analysis applications, meaning that the computed AM and FM functions were used for analysis but not to reconstruct images. I then discuss the importance of phase unwrapping in the 2D analysis and reconstruction problems. I review four major approaches to perform 2D phase unwrapping. Then I define the single component perfect reconstruction AM-FM image model. I provide two algorithms that compute AM and FM functions from a single component image and provide prefect reconstruction of the image from the AM and FM functions. I verify the perfect reconstruction property by calculating the MSE and peak signal to noise ratio (PSNR) between the reconstructed and original images.

In Chapter 5, I introduce the perfect reconstruction AM-FM image transform (xAMFM) for general images. I modify the well-known steerable pyramid [38, 106] to create a full multi-scale, multi-orientation perfect reconstruction filterbank. I then discuss the problems associated with the partial Hilbert transform (pHT) where the computed AM and FM functions can show artifacts if the image component has frequency support orthogonal to the pHT filtering axis. For the first time, I overcome this problem by rotating the axis of action of the pHT. Finally, I develop the xAMFM for general images.

In Chapter 6, I introduce the AM-FM image processing framework where filters are designed to operate directly on the AM and FM functions of an image. I define two classes of AM-FM image filters, namely the AM-based filters and the FM-based filters. I show the performance gain of the AM-based filters over traditional LSI filters in applications such as selective orientation attenuation, image enhancement, and image fusion. For the FM-based filters, I performed geometric

image transformation in the AM-FM domain. Qualitative and quantitative measures indicate that the modulation domain image filters outperform traditional spatial domain LSI filters in these applications.

In Chapter 7, I demonstrate the effectiveness of the xAMFM in image analysis. I consider two xAMFM applications in coherent texture decomposition. In addition, I investigate the connection between the AM-FM image model used in this dissertation and the emerging monogenic signal [34, 67].

Conclusions and recommendations for future work are reserved for Chapter 8. I prove for the perfect reconstruction property of the modified steerable pyramid in Appendix A. Derivations of the phase unwrapping algorithm used in this dissertation is given in Appendix B.

1.2 Nomenclature

Mathematical notations used in this dissertation are consistent with those found in typical vector calculus texts. Lower-case bold face is used for vectors, i.e., $\mathbf{x}$ and $\mathbf{n}$. $\mathbf{x}$ refers to continuous variable and $\mathbf{n}$ refers to discrete variable. Upper-case letters such as I and A can either represent images or matrices, depending on the context. $\mathbb{R}$, $\mathbb{Z}$, and $\mathbb{C}$ refer to the set of real, integer, and complex number with the usual notions of addition and multiplication.

The hat($\widehat{\ }$) denotes Fourier transform; e.g., $\widehat{f}(\omega)$ is the Fourier transform of $f(x)$. The tilde($\widetilde{\ }$) denotes a modified version of the original signal; e.g., $\widetilde{f}(x)$ denotes a modified version of $f(x)$.

The meanings of some common symbols, acronyms, terms, and abbreviations used in this dissertation are outlined in Table 1.1 and Table 1.2.

Table 1.1: Symbols used in the dissertation

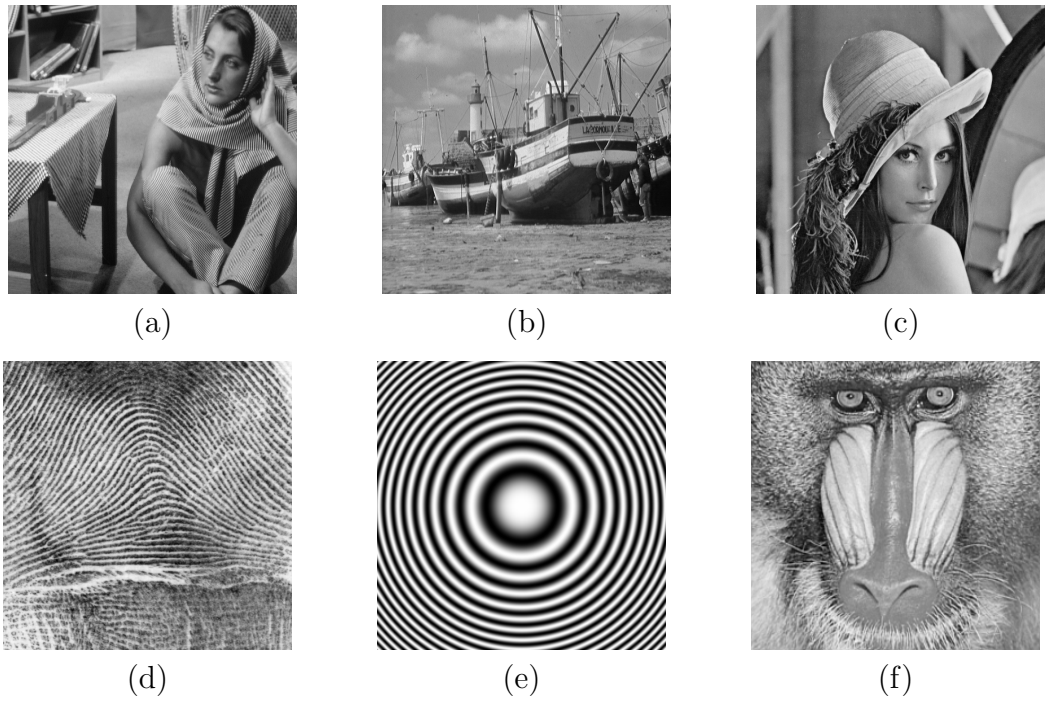
Symbol	**Meaning**
$f(\mathbf{x})$	Input continuous image
$a(\mathbf{x})$	AM function
$\varphi(\mathbf{x})$	Phase modulation function
$\nabla\varphi(\mathbf{x})$	FM vector field
$R(\mathbf{x})$	Magnitude of the FM
$\theta(\mathbf{x})$	Argument of of the FM
$h(\mathbf{x})$	Image filter
$\mathcal{H}\{.\}$	Hilbert transform
$\psi(\mathbf{x})$	Least-squares phase approximation of $\varphi(\mathbf{x})$

Table 1.2: Definitions of terms and acronyms used in the dissertation

Term	**Meaning**
AM	Amplitude Modulation
FM	Frequency Modulation
xAMFM	AM-FM transform
TKEO	Teager-Kaiser energy operator
QL	Quasi-local approximation
AS	Analytic complex image extension
HT	Hilbert transform
pHT	Partial Hilbert transform
MSE	Mean square error
PSNR	Peak signal to noise ratio
PR	Perfect reconstruction
LSI	Linear shift invariant
DCT	Discrete cosine transform
DFT	Discrete Fourier transform
Modulation domain	Refers to representation, analysis, and processing of signals and images in terms of AM and FM functions.
Pixel domain	Refers to representation, analysis, and processing of signals and images in terms of pixel values (intensity) expressed in the image plane.

1.3 Test images

Well-known test images are shown in Fig. 1.1. These test images are used in simulations.



(a) (b) (c)

(d) (e) (f)

Figure 1.1: Well-known test images. (a) Barbara. (b) boat. (c) Lena. (d) fingerprint. (e) Gaussian chirp. (f) mandrill.

Chapter 2

The 1D AM-FM Signal Model

The classical Fourier transform provides an essential tool for time-frequency analysis of stationary signals. Throughout this dissertation, I establish that a *stationary* signal is one that admits a meaningful, intuitive representation as a superposition of amplitude, linear phase, and constant frequency. However, many important signals of practical interest are nonstationary in this sense; i.e., the local amplitude envelope and local frequency content are time-varying. For such signals, the AM-FM model has been widely used to capture their nonstationary characteristics [9, 25]. The AM-FM model represents a real 1D signal as

$$f(x) = a(x)\cos[\varphi(x)], \tag{2.1}$$

where $a(x) : \mathbb{R} \to \mathbb{R}^+$ is a non-negative smoothly varying amplitude modulation function (AM) and $\varphi(x) : \mathbb{R} \to \mathbb{R}$ is a smoothly varying phase modulation function. The frequency modulation function (FM) is defined as the derivative of $\varphi(x)$, e.g., $\varphi'(x) = \frac{\mathrm{d}}{\mathrm{d}x}\varphi(\mathbf{x})$. The FM is also referred to as the *instantaneous frequency* (IF) [25]. The terms IF and FM are used interchangeably in this dissertation.

I discuss the importance of the IF in signal processing applications in Section 2.1.1. The computation of 1D AM-FM signal model is described in Section 2.2. Finally, I evaluate the effectiveness of three major 1D AM-FM computational approaches in Section 2.3.

2.1 Instantaneous Frequency

The IF concept originated in the physics and communications communities where the signals of interest are primarily 1D [25, 41, 75, 124]. The IF of a 1D signal modeled according to (2.1) is given by

$$\mathrm{IF}(x) = \varphi'(x). \tag{2.2}$$

2.1.1 The Importance of Instantaneous Frequency

The IF captures nonstationary characteristics of time-varying signals in many practical applications [9]. In radar applications, the time varying behavior of the IF indicates the variation of frequency and motion of objects [23]. The object motion is modeled as a point target giving rise to time-varying frequency content in the radar returns. In seismic survey data processing, geological structures can be characterized by analyzing the returns of seismic signals [119]. The reflection strength in seismic traces is determined by the amplitude envelope of the return signals. Important geological structures can be deduced from the instantaneous phase even though the return signal has a weak reflection strength. In biomedical applications, the IF is used to identify blood flow [116] and to diagnose diastolic murmurs [130]. In speech processing, important nonlinear features in human speech production can be captured by the AM-FM model [78,120]. In oceanography, the IF measures the kinetic energy dissipation of turbulent water [59]. Middleditch and Wyatt [85] observed that the temporal and spatial inconsistencies within the radar measurement region can lead to distortions in the spectra of the radar returns. They developed an IF filtering technique to remove the first-order modulation from the return signals to improve the quality of oceanographic radar measurements.

2.1.2 Debates about Instantaneous Frequency

Given a 1D continuous signal $f(x)$, there are infinitely many pairs of $a(x)$ and $\varphi(x)$ satisfying the model (2.1). Since the IF is defined as $\varphi'(x)$, it is not unique for a given $f(x)$. Even if the IF is computed, its existence was debated in the research community, based mostly on the physical interpretations of IF [9, 75]. Since IF and Fourier spectrum capture properties of signals in term of frequency, it is natural to wonder if there exists any relationships between these two quantities. Mandel investigated such relationships and showed that the first moment of the IF equals the first moment of the Fourier frequency [75]. He pointed out that this relationship does not hold for higher moments. In addition, he showed that there is no one-to-one relationship between these two descriptors. In other words, there exist signals where the Fourier frequencies manifest while the IF vanishes and vice versa. Indeed, this counter-intuitive fact is the root of many debates.

Cohen [25] noted several paradoxes concerning the IF. First, the IF is generally distinct from the Fourier frequencies. If the signal is a cosine wave oscillating with a single frequency, the IF is a constant and continuous function. In addition, the IF of the analytic signal can be negative, which is counter-intuitive because the analytic signal does not admit negative Fourier frequencies. That is, the IF can lie outside the frequency support of the signal. Some authors [9, 25] interpreted that the IF is the average of Fourier frequency at a given time. This interpretation was originated from first moment relationship between Fourier frequency and IF. Nevertheless, this interpretation is only true for single component signal and is generally false or misleading for general multicomponent AM-FM signals [71]. Despite being a controversial concept, the IF has been used successfully as a descriptor for nonstationary signals [9, 10].

2.2 Computation of the 1D AM-FM Signal Model

The AM-FM model (2.1) is ill-posed in the sense that infinitely many pairs of AM and FM functions satisfy this equation for a given signal $f(x)$. Here I review several major 1D IF computation methods such as analytic signal differentiation, zero-crossing identification, time-frequency distribution, polynomial phase modeling, and the Teager-Kaiser energy operator. Readers are referred to in-depth discussions of the IF by Boashash [9, 10] and Vakman [124].

2.2.1 Analytic Signal

Gabor [41] pioneered a complex signal model to compute the IF. For 1D signal $f(x)$, the complex signal model includes the signal $f(x)$ as the real part and an auxiliary signal $q(x)$ as the imaginary part. The signal $q(x)$ is obtained by applying the Hilbert transform (HT) to $f(x)$.

The HT kernel is characterized by the time domain impulse response

$$h(x) = \frac{1}{\pi x}, \tag{2.3}$$

where $x \in \mathbb{R}$. The function $h(x)$ is neither a $L^1(\mathbb{R})$ nor a $L^2(\mathbb{R})$ function because of the singularity at the origin. Let $f(x) : \mathbb{R} \to \mathbb{R}$. The HT of $f(x)$ is a real signal $q(x)$ given by

$$q(x) = \mathcal{H}\{f(x)\} = f(x) * \frac{1}{\pi x}, \tag{2.4}$$

where $*$ denotes convolution. Since $h(x)$ is undefined at $x = 0$, the convolution operation in (2.4) must be evaluated as a Cauchy principal value according to

$$\begin{aligned} q(x) &= \frac{1}{\pi}\,\mathrm{p.v}\int_{\mathbb{R}} \frac{f(\tau)}{x-\tau}\,\mathrm{d}\tau \\ &= \frac{1}{\pi}\lim_{\epsilon\to 0}\Big(\int_{-\infty}^{x-\epsilon} \frac{f(\tau)}{x-\tau}\,\mathrm{d}\tau + \int_{x+\epsilon}^{\infty} \frac{f(\tau)}{x-\tau}\,\mathrm{d}\tau\Big). \end{aligned} \tag{2.5}$$

The corresponding Fourier spectrum of the 1D HT kernel $h(x)$ is

$$\widehat{h}(\omega) = -j\,\frac{\omega}{|\omega|} = -j\,\mathrm{sgn}(\omega), \tag{2.6}$$

where $\mathrm{sgn}(\omega)$ is defined as

$$\mathrm{sgn}(\omega) = \begin{cases} 1, & \omega > 0 \\ 0, & \omega = 0 \\ -1, & \omega < 0. \end{cases} \tag{2.7}$$

Since $||\widehat{h}(w)||_2^2 = 1$ almost everywhere [1], the HT operator does not affect the magnitude of the Fourier spectrum of the input signal.

The 1D HT possesses three distinct properties that make it an important transform in communication and signal processing.

1. It is a linear time invariant (LTI) operator.

2. It is anti-symmetric and bounded for $1 < p < \infty$ [114][page 49].

3. It is an all-pass filter (except at DC).

All of these properties follow directly from the definition of $\mathcal{H}\{.\}$ in (2.5) and (2.6).

Let $f(x) : \mathbb{R} \to \mathbb{R}$. Then the 1D analytic signal $z(x) : \mathbb{R} \to \mathbb{C}$ is constructed as

$$z(x) = f(x) + j\mathcal{H}\{f(x)\} = f(x) + jq(x), \tag{2.8}$$

[1] Here, I mean *almost everywhere* in the precise mathematical sense, i.e., everywhere except on a set of Lebesgue measure zero.

or equivalently in Fourier domain as

$$\begin{aligned}\widehat{z}(\omega) &= \widehat{f}(\omega) + j[-j\,\mathrm{sgn}(\omega)\widehat{f}(\omega)] \\ &= \widehat{f}(\omega)[1+\mathrm{sgn}(\omega)] \\ &= \begin{cases} 2\widehat{f}(\omega) & \omega > 0 \\ \widehat{f}(\omega) & \omega = 0 \\ 0 & \omega < 0. \end{cases} \end{aligned} \tag{2.9}$$

As the signal $f(x)$ is real, its spectrum is Hermitian, i.e., complex conjugate symmetric. Therefore, negative Fourier frequencies are unnecessary for the unambiguous representation of $f(x)$. The complex signal construction (2.9) using the HT eliminates the spectral information on the negative half of the frequency axis while it doubles it on the positive half.

Vakman [124] showed that the analytic signal constructed by the HT satisfies the following three conditions.

1. Amplitude continuity and differentiability: If a small variation is added to the signal $f(x)$, the variation of the AM must be small.

2. Homogeneity: If the signal $f(x)$ is scaled by a constant $c > 0$, the phase function $\varphi(x)$ and the FM are not changed.

3. Harmonic correspondence: the HT maps pure cosine/sine signals into pure sine/cosine signals. Therefore, the HT can be used to define a complex signal to uniquely solve the AM-FM representation of model (2.1).

Once the analytic signal $z(x)$ is constructed, one can analytically calculate the amplitude modulation (AM) as

$$a(x) = |z(x)|. \tag{2.10}$$

The IF can be computed from different approaches such as complex signal differentiation, zero-crossing identification, phase modeling, and time-frequency peak detection.

Analytic Signal Differentiation

After constructing the analytic signal (2.8), one can directly compute the FM function using differentiation as [41]

$$\varphi'(x) = \mathrm{Im}\left[\frac{z'(x)}{z(x)}\right]. \tag{2.11}$$

The IF in (2.11) is undefined at points where $z(x) = 0$. At such points, the IF can be approximated by the neighborhood IF. It should be noted that the concept of IF is not meaningful on intervals where $z(x)$ or $f(x)$ is identical to zero.

Since the derivative operator in (2.11) is only defined for functions of continuous variable x, this IF computation algorithm is restricted to the continuous time signals. In practical applications where only discrete signals are available, a discrete counterpart of the derivative is required.

Zero-crossing Identification

The nonstationary signal model in (2.1) assumes that $f(x)$ is a locally narrowband signal. Therefore, the number of zero-crossings can describe the oscillating behavior of the underlying signal. Some authors have defined the IF in terms of zero-crossings identification [10]. Let k be the number of sample intervals between two zero-crossings. The IF is then estimated according to

$$\varphi'(x) = \frac{1}{2k}. \tag{2.12}$$

To improve the accuracy of the zero-crossing identification process, a sliding window technique is often applied. The zero-crossing identification is then restricted

to within the support of the sliding kernel [10].

Polynomial Phase Modeling

The phase function $\varphi(x)$ can be modeled by a p-order polynomial [10]

$$\varphi(x) = b_0 + b_1 x + b_2 x^2 + ... + b_p x^p. \tag{2.13}$$

The selection of p depends on the oscillating characteristics of the signal. Regression techniques such as least-squares are used to compute the coefficients b_0, b_1, $\cdots, b_p$. Let $f(n)$ be the true discrete signal of length N. Then $f(n)$ is modeled according to (2.1) as $f(n) = a(n)\cos[\varphi(n)]$. Let $\tilde{f}(n)$ be the measured discrete signal. The coefficients of the polynomial in (2.13) are obtained by minimizing the sum of squared error E as

$$E = \sum_{n=0}^{N-1} |\tilde{f}(n) - f(n)|^2. \tag{2.14}$$

A least square minimization approach was given by Boashash [10]. However, the solution depends on initial estimate of the polynomial coefficients b_k.

Time Frequency Distribution Techniques

In the study of time-frequency analysis, it was realized that important characteristics of a nonstationary signal are captured by the IF [9, 10, 24, 25]. Specifically, the IF can be estimated by temporally tracking peaks in the time-frequency distribution [10]. For example, IF can be estimated from tracking peaks of the time-frequency distribution computed by the short time Fourier transform (STFT) [11]. However, IF estimation from the STFT method depends on the choice of the window. Martin and Flandrin used the Wigner-Ville distribution (WVD) as an alternative to STFT to extract the IF [81]. Recently, Sedjić et al. [103] proposed a

hybrid approach where they combined the STFT and wavelets for time frequency analysis. They showed that the hybrid approach performs better than the STFT and the WVD.

2.2.2 Teager-Kaiser Energy Operator

Maragos, Kaiser, and Quatieri [78, 79] used the Teager-Kaiser energy operator (TKEO) to compute the AM and FM functions of the signal (2.1). The Teager energy of the signal $f(x)$ in (2.1) is defined by

$$\begin{aligned} \Psi[f(x)] &= [f'(x)]^2 - f(x)f''(x) \\ &= [a(x)]^2[\varphi'(x)]^2. \end{aligned} \tag{2.15}$$

They applied TKEO to the derivative signal $f'(x)$ to obtain

$$\begin{aligned} \Psi[f'(x)] &= [f''(x)]^2 - f'(x)f'''(x) \\ &= [a(x)]^2[\varphi'(x)]^4. \end{aligned} \tag{2.16}$$

From (2.15) and (2.16), the AM and FM functions may be computed according to [79]

$$a(x) = \frac{\Psi[f(x)]}{\sqrt{\Psi[f'(x)]}}, \tag{2.17}$$

$$\varphi'(x) = \sqrt{\frac{\Psi[f'(x)]}{\Psi[f(x)]}}. \tag{2.18}$$

In (2.17), the AM function is undefined at points where when $\Psi[f'(x)]$ is zero; similarly, the FM function in (2.18) is undefined where $\Psi[f(x)]$ is zero. Based on these observations, Vakman [124] provided examples where the TKEO fails to produce intuitive AM and FM functions. In addition, $\Psi[f(x)]$ can be negative [15]. However, for many signals of practical interest, $\Psi[f(x)]$ is non-negative [78].

Approximate discrete counterparts for the continuous algorithms in (2.17) and (2.18) were proposed in [78]. Here I will restrict attention to the first discrete energy separation algorithm, DESA-1. The closely related DESA-2 algorithm is substantially similar. The DESA-1 can demodulate AM and FM functions with maximum IF up to half the sampling frequency. Let $f(n)$ be a real discrete signal with N samples where $n = 0, 1, \cdots, N-1$. Let $y(n) = x(n) - x(n-1)$ be the first order finite backward asymmetric difference of $x(n)$. Then the AM and FM functions may be estimated by

$$a(n) \approx \sqrt{\frac{\Psi[f(n)]}{1 - \left(1 - \frac{\Psi[y(n)]+\Psi[y(n+1)]}{4\Psi[f(n)]}\right)^2}}, \tag{2.19}$$

$$\varphi'(n) \approx \arccos\left(1 - \frac{\Psi[y(n)] + \Psi[y(n+1)]}{4\Psi[f(n)]}\right). \tag{2.20}$$

The FM computed by (2.20) is always non-negative. Hence, the sign of the FM computed by TKEO is ambiguous if the signal contains both negative and positive IF.

In contrast to the analytic signal approach of Section 2.2.1, the 1D TKEO approach is a local method. The TKEO approach assumes that the signal $f(x)$ is locally narrowband. For locally wideband signals, a multiband decomposition must be performed prior to the demodulation process [16]. The 1D TKEO has been used successfully to extract features from human speech signals [47, 78].

2.2.3 Quasi-Local AM-FM Estimation

As mentioned in Section 2.2.1 and 2.2.2, the analytic signal approach is a global method and the TKEO is a local method. Girolami and Vakman [43] proposed a quasi-local method to compute the AM and FM functions. Let $f(x)$ be the real

signal in (2.1). Let "$\langle\cdot\rangle$" denote a Bessel lowpass filter. If the lowpass filter can be assumed to efficiently eliminate signal components of the form $\cos[2\varphi(x)]$, then the AM may be obtained according to [43]

$$\begin{aligned} 2\langle f^2(x)\rangle &= 2\langle a^2(x)\cos^2[\varphi(x)\rangle \\ &= \langle a^2(x)\left(1+2\cos[2\varphi(x)]\right)\rangle \\ &\approx a^2(x), \end{aligned} \tag{2.21}$$

in which $\langle a^2(x)\cos[2\varphi(x)]\rangle \approx 0$ and $\langle a^2(x)\rangle \approx a^2(x)$. This assumes that $a^2(x)$ and $\cos[2\varphi(x)]$ are spectrally disjoint, which is reasonable for many signals of practical interest [6]. Equivalently, the AM can be written as

$$a(x) \approx \sqrt{2\langle f(x)f(x)\rangle}. \tag{2.22}$$

In order to compute the FM, Girolami and Vakman defined an auxiliary function

$$R(x,\tau) = \frac{2\langle f(x+\tau)x(x-\tau)\rangle}{\langle f(x)f(x+\tau)+f(x)f(x-\tau)\rangle}. \tag{2.23}$$

They then obtained obtained the FM by

$$\varphi'(x) \approx \lim_{\tau\to 0}\left(\frac{1}{\tau}\arccos\left(\frac{R+\sqrt{R^2+8}}{4}\right)\right). \tag{2.24}$$

In practice, a small value of τ is used in (2.23) to approximate (2.24).

The method requires three local points, $f(x)$, $f(x-\tau)$, and $f(x+\tau)$, which makes it a local method. However, as a lowpass filter is involved in the computation process, the method is not strictly local. Hence, it is classified as quasi-local. The authors showed that the AM and FM functions computed by this approach also satisfy the same three conditions as those computed by the analytic signal approach in Section 2.2.1. Nevertheless, the quasi-local method depends on the bandwidth

selection of the lowpass filter, which is signal dependent. The selection of the constant τ also affects the demodulated AM and FM functions. In addition, the quasi-local approach can only compute the absolute value of the FM function. Therefore, the signed FM is ambiguous in the QL method.

2.3 Comparison of 1D AM-FM Techniques

In this section, I quantitatively compare the AM-FM demodulation effectiveness of the analytic signal (AS) approach given in Section 2.2.1, the Teager-Kaiser energy operator (TKEO) approach given in Section 2.2.2, and the quasi-local (QL) method given in Section 2.2.3. The test signal is taken from [79] and is given by

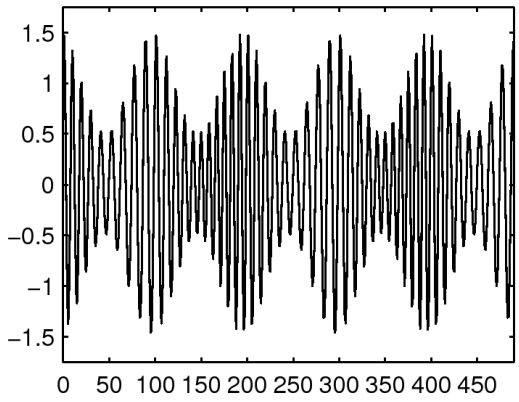
$$f(n) = \left(1 + 0.5\cos\left[\frac{\pi}{50}n\right]\right)\cos\left[\frac{\pi}{5}n + 4\sin\left(\frac{\pi}{100}n + \frac{\pi}{4}\right)\right], \tag{2.25}$$

where the AM is a 0.01 Hz sinusoid and the FM is given by

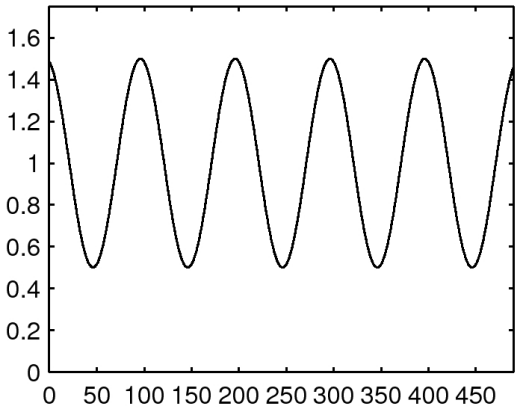
$$\varphi'(n) = \frac{\pi}{5} + \frac{\pi}{25}\cos\left[\frac{\pi}{100}n + \frac{\pi}{4}\right], \tag{2.26}$$

which is sinusoidally varying in the range $0.502 \leq \varphi'(n) \leq 0.754$ radians. Here, it is understood that $\varphi'(n)$ is obtained by interpolating (2.25) to obtain $f(x)$ where $x \in \mathbb{R}$, differentiating $\varphi(x)$ to obtain $\varphi'(x)$, and finally sampling $\varphi'(x)$ to obtain $\varphi'(n)$. The original signal, the AM, and the FM are illustrated in Fig. 2.1(a)-(c).

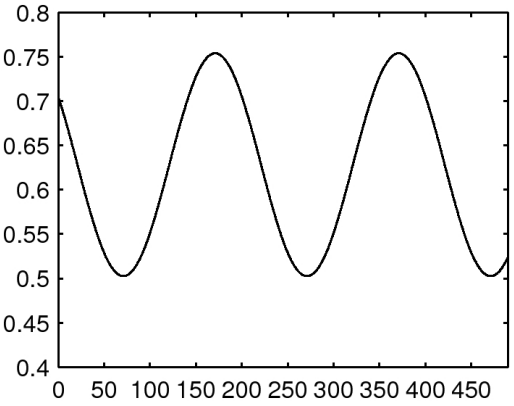
While all three approaches were formulated in continuous settings, I used discrete implementations to evaluate their effectiveness. For the TKEO approach, I first used the DESA-1 algorithm [79] to computed the modulation functions. The AM and FM functions computed using (2.19) and (2.20); and they are shown in Fig. 2.2. The AM and FM functions obtained by the AS method are depicted in Fig. 2.3. In the AS approach, the AM is computed as in (2.10) and the FM is



(a)

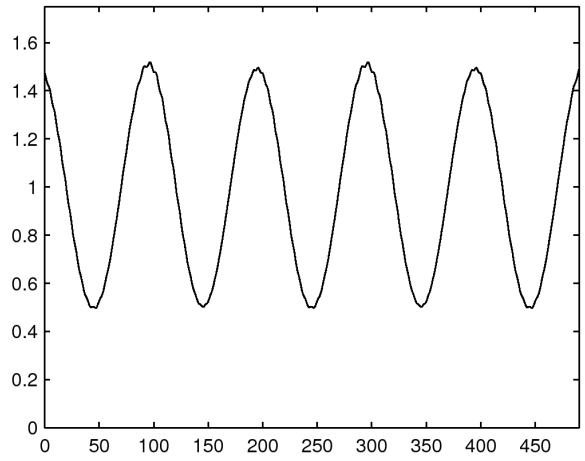


(b)

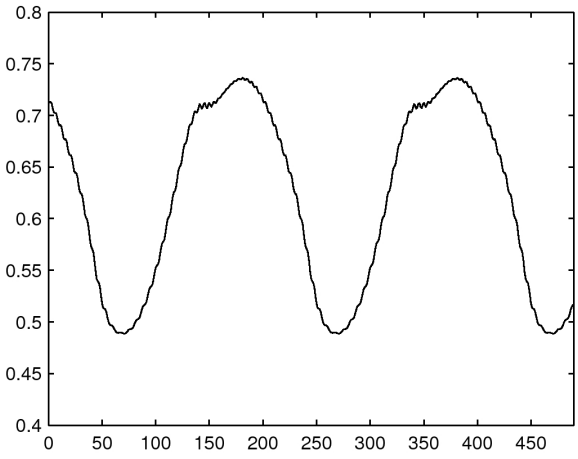


(c)

Figure 2.1: Test signal f(n) and theoretical values of the AM and FM functions. (a) Original AM-FM test signal given in (2.25). (b) Theoretical AM function obtained analytically from (2.25). (c) Theoretical FM function obtained analytically from (2.25), as given in (2.26)

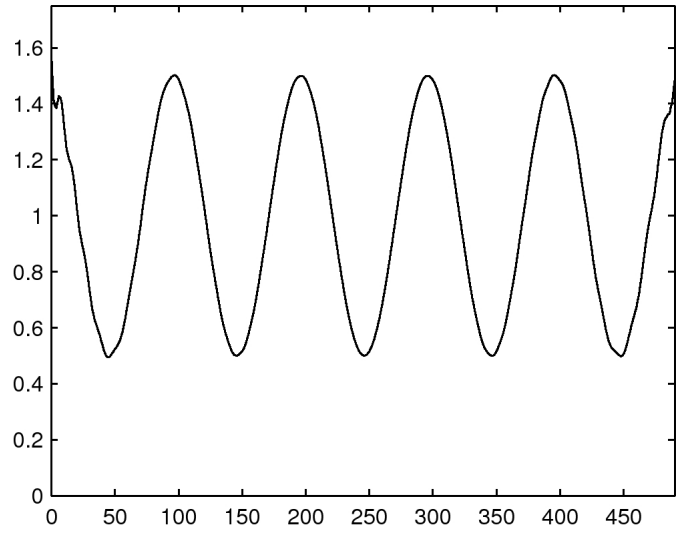


(a)

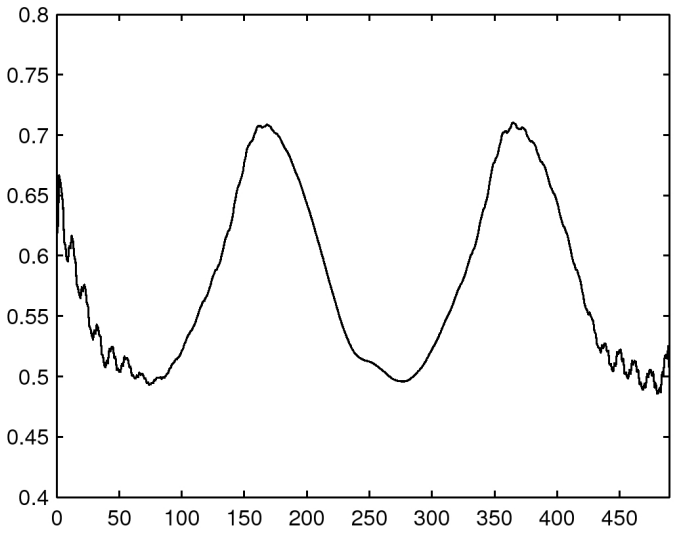


(b)

Figure 2.2: Computed AM and FM functions of the TKEO approach. (a) The computed AM function using (2.19). (b) The computed FM function using (2.20)

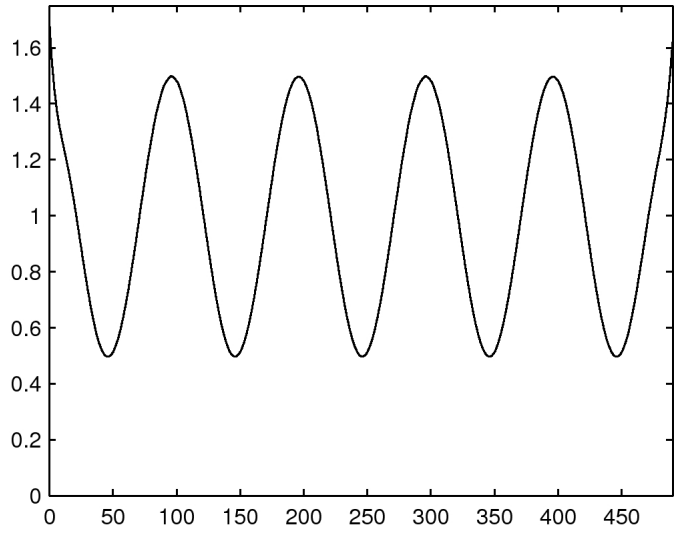


(a)

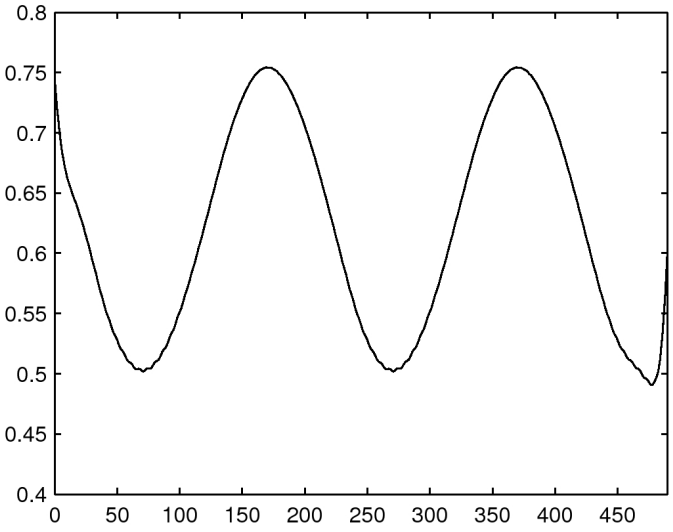


(b)

Figure 2.3: Computed AM and FM functions of the AS approach. (a) The computed AM function using (2.10). (b) The computed FM function by discritizing (2.11).



(a)



(b)

Figure 2.4: Computed AM and FM functions of the QL approach. (a) The computed AM function using (2.22). (b) The computed FM function using (2.24).

obtained as in (2.11) using a 2-point forward asymmetric difference as the discrete counterpart of the continuous derivative operator. Notice that direct discretization of (2.11) can lead to poor performance if the signal $f(n)$ is corrupted. The AM and FM functions obtained by the quasi-local approach are shown in Fig. 2.4. The lowpass filter is a 2^{nd} order Butterworth filter with normalized cutoff frequency 0.04. In this experiment, τ in (2.23) is set to 1.

To evaluate the performances of these three methods in the presence of additive noise, I corrupted the signal (2.25) with additive white Gaussian noise with standard deviation $\sigma = 0.1$ (40.33 dB) and $\sigma = 0.2$ (26.47 dB). I measured the accuracy of the three algorithms by computing the mean square error (MSE) between the demodulated signals and their theoretical values as shown in Fig. 2.1(b) and (c). The MSE comparison of the AM is given in Table 2.1 and the MSE comparison of the FM given in Table 2.2. From the MSE results, the quasi-local method [43] yields the best MSE performance. However, the performance of this method depends on the bandwidth of the lowpass filter. For this experiment, the lowpass filter was designed with a priori knowledge of the true FM function. In particular, the cutoff frequency of the lowpass filter was set to 0.04 Hz so that the lowpass filter captures the AM function, which is oscillated at 0.01 Hz, and rejects the FM frequency, which is in the range [0.079 0.12] Hz. Therefore, I restricted the comparison to the TKEO and AS.

While producing a lower MSE for the FM function in the noise-free case, the TKEO is relatively more sensitive to noise. As the noise power increases, the accuracy of the demodulated FM by the TKEO degrades faster than that of the AS. The accuracy of the TKEO can be improved by filtering the input signal with a multiband filterbank and performing the demodulation algorithm on each

Table 2.1: Comparison of MSE in estimated AM obtained by three competing discrete demodulation algorithms in presence of additive white Gaussian noise. ∞ denotes the noise-free signal.

PSNR	TKEO	AS	Quasi-Local
∞	5.0399×10^{-4}	7.9848×10^{-5}	2.8751×10^{-4}
40.33 dB	0.2480	0.0103	0.0017
26.47 dB	0.5604	0.0377	0.0056

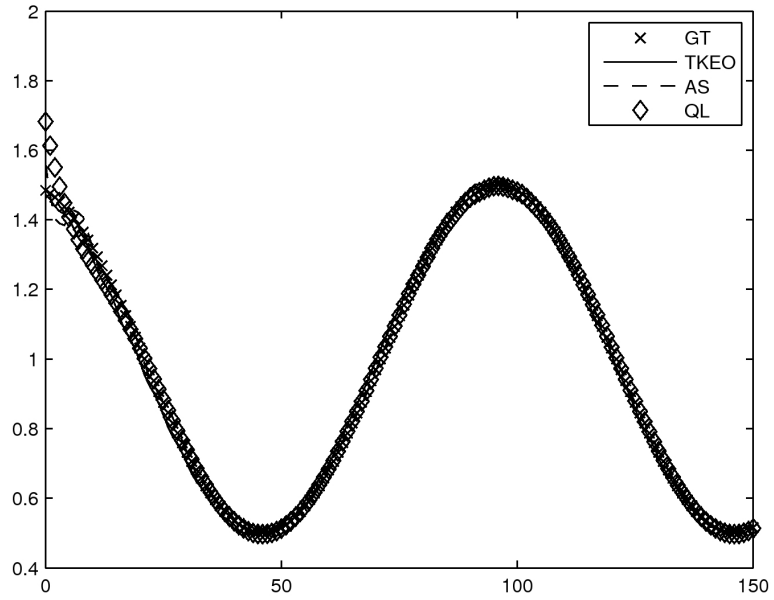
Table 2.2: Comparison of MSE in estimated FM obtained by three competing discrete demodulation algorithms in presence of additive white Gaussian noise. ∞ denotes the noise-free signal.

PSNR	TKEO	AS	Quasi-Local
∞	3.1906×10^{-4}	0.0022	3.8563×10^{-5}
40.33 dB	0.4487	0.0213	8.183×10^{-5}
26.47 dB	1.0093	0.1686	0.0023

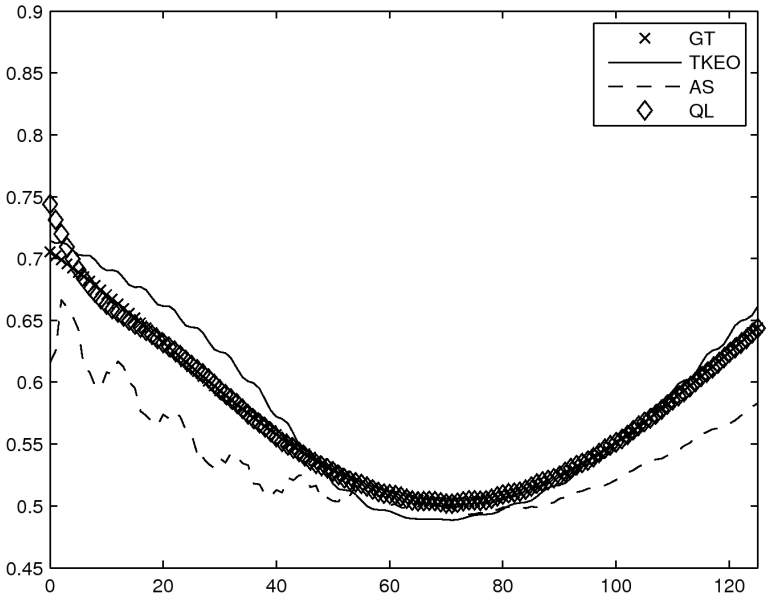
subband separately as suggest in [16]. In addition, improvements in the accuracy of the TKEO can be achieved by median filtering the demodulated outputs [78].

The AS consistently yields better AM demodulation than the TKEO. As illustrated in Fig. 2.5 and Fig. 2.6, the FM of the AS is more resilient to noise than that of the TKEO. Nevertheless, the FM of the AS is also prone to noise, because it is computed using a 2-point forward difference as the discrete counterpart of the continuous derivative operator. This poor performance against noise is related to the derivative filter which is generally a highpass filter. Therefore, noise in the signal is amplified whenever the derivative filter is applied the signal.

While computing the AM and FM functions of the same model, the three approaches are fundamentally different in term of localization. While the discrete unfiltered demodulation algorithm associated with the AS method is temporally localized, the HT requires the entire signal to construct the analytic signal $q(x)$ in (2.4). Consequently, this approach is decidedly global in this respect. The

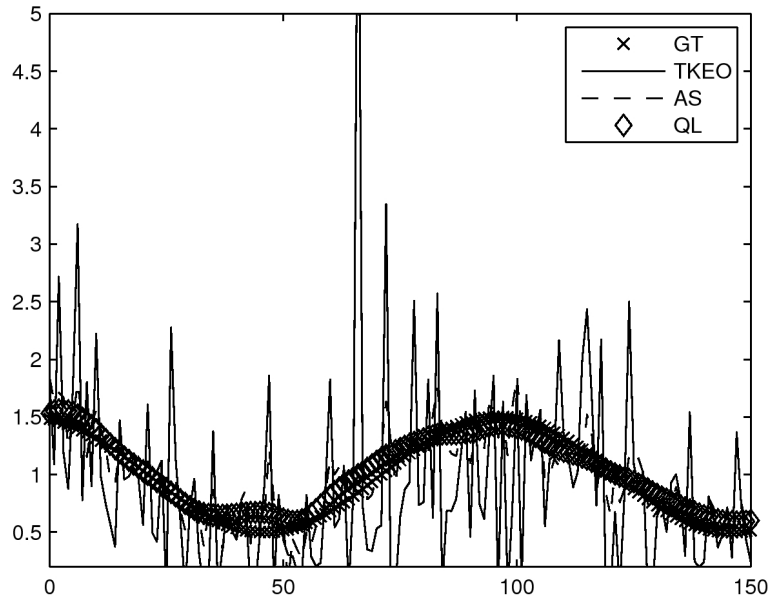


(a)

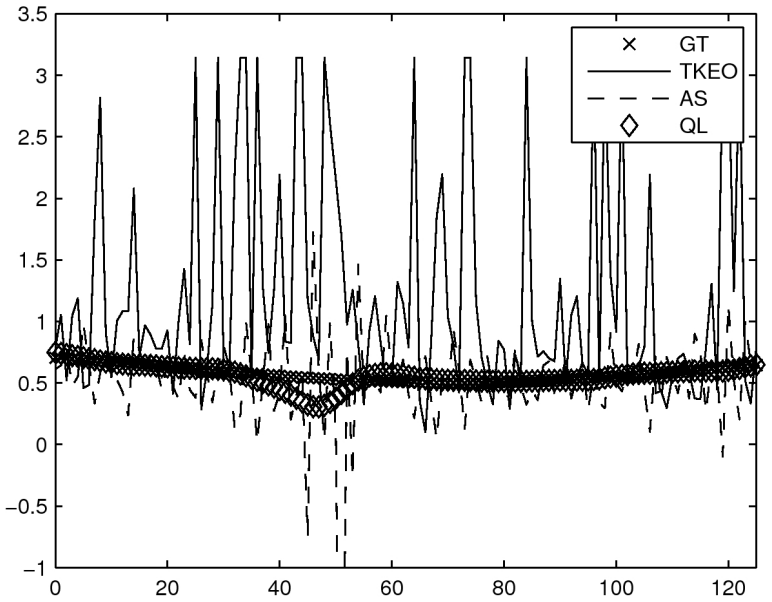


(b)

Figure 2.5: Detail view of computed AM and FM functions from the noise-free signal $f(n)$ in (2.25) using all three methods. (a) Computed AM function. (b) Computed FM functions.



(a)



(b)

Figure 2.6: Detail view of computed AM and FM functions from the signal $f(n)$ in (2.25) corrupted by additive white Gaussian noise with standard deviation $\sigma = 0.2$ using all three methods. (a) Computed AM function. (b) Computed FM functions.

TKEO is regarded as a local method because the AM and FM functions are computed from local samples only, e.g., generally only five consecutive samples are required by the DESA-1 algorithm. The quasi-local method is a mixture of local and global techniques as it involves both a lowpass filtering process and a local operator.

The TKEO and the quasi-local method put a restriction on the time-varying characteristics of the input signal. The TKEO DESA-1 algorithm is applicable to signal with IF limited to half the sampling frequency. The maximum IF requirement for the quasi-local is stricter than that of the TKEO. The quasi-local method works well for signal having IF less one fourth of the sampling frequency. In contrast to these two approaches, the AS applies for any input signal without a priori knowledge of its IF range.

2.4 Summary

In this chapter, I introduced the concept of the 1D AM-FM signal model. I discussed the importance of the instantaneous frequency (IF), or FM, in practical applications and reviewed three major approaches that compute the AM and FM functions from a given real signal.

I discussed the mathematical construction of the Gabor analytic signal (AS) approach, the Teager-Kaiser energy operator (TKEO) approach , and the quasi-local (QL) approach. I evaluated the effectiveness of the three approaches in terms of the accuracy of the demodulated AM and FM functions. The adopted quality metric is the mean square error of the demodulated functions with respect to their theoretical values. Of the three methods, the AS has an advantage over the TKEO and the QL, because it does not make any a priori assumption about the IF range

of the signal and it is more robust to noise compared to the TKEO. The QL can provide the best performance if the lowpass filter is correctly designed, i.e., a priori knowledge of true FM signal is required. The noise immunity and general accuracy of the TKEO approach can be improved if the input signal is first decomposed into subbands and the TKEO is then applied to the subband signals separately.

While the IF is well-understood and has been a useful descriptor for 1D nonstationary signals in many practical applications, its generalization to multiple dimensions is not straightforward. For instance, the Gabor analytic signal can not be naturally extended to higher dimensions because the HT is strictly defined for 1D signals. As there exist multiple definitions for the nD extension of the HT, the multidimensional extensions of the Gabor analytic signal are not unique. In addition, the physical interpretation of the multidimensional IF poses a challenging problem, e.g., what information does IF carry if the signal is an image? These issues are the topic of the next chapter.

Chapter 3

The 2D AM-FM Image Model

While the IF is well-understood and has been proven to be a useful descriptor for 1D nonstationary signals in many practical applications [9], its multidimensional counterpart remains an open research problem [34, 49]. The main reason is that the multidimensional extension of the 1D IF is not unique. For instance, the Gabor analytic signal can not be naturally extended to higher dimensions because there is no standard accepted definition of the HT in dimensions greater than one. Consequently, multiple multidimensional extensions of Gabor's analytic signal have been proposed. In addition, even if the nD IF is computed, especially for signals with dimension higher than two, whether it provides any correlation with physical interpretations is largely unexplored. In this chapter, I focus my attention on the 2D case where the IF captures structural properties of the image texture that are meaningful in terms of visual perception and interpretation.

Traditionally image filters operate on the pixel intensity, *viz.*, pixel domain. However, in many applications such as motion detection, edge detection, and image segmentation, the phase function and the IF can provide more accurate and robust solutions [36]. Oppenheim and Lim [95] argued that for images, the phase captures more important information to human visual perception than the magnitude frequency. Fleet and Jepson [36] showed that the phase is more resilient to image deformation than pixel intensity in the sense that phase is approximately

linear on sufficiently small neighborhood, whereas the intensities are not. Moreover, the phase is invariant to image contrast changes. The authors used the IF to compute optical flow in images [5]. Bovik, Clark, and Geisler [13] used the phase in image segmentation. They showed that homogeneous regions can be separated by detecting discontinuities in the local phase. In addition, they showed that the IF carries important characteristics of the texture, namely orientation and pattern granularity.

I discuss the 2D AM-FM model in Section 3.1.1. The computational approaches are reviewed in Section 3.2 and 3.3. I compare the performance of 2D AM-FM demodulation algorithms in Section 3.4. Finally, the summary is given in 3.5.

3.1 The AM-FM Image Model: A Review

3.1.1 The AM-FM Image Model

Let $f(\mathbf{x}) : \mathbb{R}^2 \rightarrow \mathbb{R}$ be a real image. The AM-FM image model represents the image as

$$f(\mathbf{x}) = a(\mathbf{x}) \cos[\varphi(\mathbf{x})], \tag{3.1}$$

where $a(\mathbf{x}) : \mathbb{R}^2 \rightarrow \mathbb{R}^+$ is the amplitude modulation function (AM) and $\varphi(\mathbf{x}) : \mathbb{R}^2 \rightarrow \mathbb{R}$ is the phase modulation function. The frequency modulation function (FM) is defined by $\nabla\varphi(\mathbf{x})$, which I will sometimes write in terms of its components as $[U(\mathbf{x})\ V(\mathbf{x})]^T$. The AM and FM are assumed to be smoothly varying signals.

The AM and FM functions can provide an intuitive interpretation of an image $f(\mathbf{x})$ [52]. The AM captures local contrast of the image where larger values of AM imply higher local contrast. The magnitude of the FM, $|\nabla\varphi(\mathbf{x})|$, describes texture spacing or granularity. A large value of $|\nabla\varphi(\mathbf{x})|$ indicates high frequency

structures, e.g., closer texture spacing. The argument of the FM, $\arg\nabla\varphi(\mathbf{x})$, characterizes the local orientation of image textures.

3.1.2 Computation of the AM-FM Image Model

As described in Section 3.2 and 3.3, the AM and FM functions in (3.1) can be computed by different approaches. Peyrin, Zhu, and Goutte [97] defined a 2D complex signal by performing a 1D Hilbert transform along a specified direction; such transform is frequently referred to as *partial Hilbert transform* (pHT) [46,52]. All Hilbert transforms belong to a class known as multiplier transform; Havlicek, Havlicek, and Bovik proposed an *adjusted* multiplier in [53] to enforce harmonic correspondence [124] of the multidimensional pHT and used it to define a multidimensional analytic image [49]. Other 2D AM-FM developments include Hahn's single orthant complex signal [45] and the hypercomplex signal of Bülow and Sommer [19]. Felsberg and Sommer [34] introduced the multidimensional monogenic signal where the 1D HT is replaced by a nD Riesz transform. Independently, Larkin, Bone, and Oldfield [67] used the same signal model to study fingerprints.

Most of the nD AM-FM approaches in the literature can be categorized into two groups, namely the complex signal extensions and the approaches using separable implementations. I discuss 2D complex signal extension in Section 3.2 and the separable approaches in Section 3.3. While all of these approaches a valid for signal with dimension greater than two, I restrict the discussion to 2D signals as image processing is the primary goal in this dissertation. I compare, contrast, and evaluate these techniques in terms of mathematical construction and meaningful physical interpretation.

3.2 The 2D Complex Image Extension

Let $\mathbf{x} = [x_1\ x_2]^T$ be the 2D position vector, $f(\mathbf{x}) : \mathbb{R}^2 \to \mathbb{R}$ be the 2D real signal, $q(\mathbf{x}) : \mathbb{R}^2 \to \mathbb{R}$ be the 2D Hilbert transform of $f(\mathbf{x})$, and $z(\mathbf{x}) : \mathbb{R}^2 \to \mathbb{C}$ be the 2D complex signal obtained by adding an imaginary part to $f(\mathbf{x})$.

3.2.1 Partial Hilbert Transform

The 2D partial Hilbert transform (pHT) extends the 1D HT by performing the 1D HT in a certain direction defined by $\mathbf{e} = [e_1\ e_2]$. The spectrum of the 2D pHT is given by

$$\widehat{q}_{\mathbf{e}}(\boldsymbol{\omega}) = -j\mathrm{sgn}(\boldsymbol{\omega}^T\mathbf{e})\widehat{f}(\boldsymbol{\omega}). \tag{3.2}$$

The 2D complex signal shares the same construction as the 1D case, i.e., the real part is the original signal $f(\mathbf{x})$ and the imaginary part is the pHT of the real part. The Fourier representation of the complex signal $z_{\mathbf{e}}(\mathbf{x})$ is given by

$$\begin{aligned} \widehat{z}_{\mathbf{e}}(\boldsymbol{\omega}) &= \widehat{f}(\boldsymbol{\omega})[1 + \mathrm{sgn}(\boldsymbol{\omega}^T\mathbf{e})] \\ &= \begin{cases} 2\widehat{f}(\boldsymbol{\omega}) & \boldsymbol{\omega}^T\mathbf{e} > 0 \\ \widehat{f}(\boldsymbol{\omega}) & \boldsymbol{\omega}^T\mathbf{e} = 0 \\ 0 & \boldsymbol{\omega}^T\mathbf{e} < 0. \end{cases} \end{aligned} \tag{3.3}$$

Fig. 3.1 shows the spectral support of the 2D complex signal generated by the pHT acting in direction specified by $\mathbf{e}$. The spectrum is zero in the white regions and is doubled in the shaded region.

The 2D pHT retains most the properties of the 1D HT, e.g., it is linear and bounded. It also reduces the redundancy inherent in the conjugate symmetric spectrum of a real signal. Because of the pHT spectral multiplier $\mathrm{sgn}(\boldsymbol{\omega}^T\mathbf{e})$ defined in (3.2), the transform fails to satisfy the harmonic correspondence property. Particularly, the pHT of a pure cosine or sine signal with frequency vector lying in the

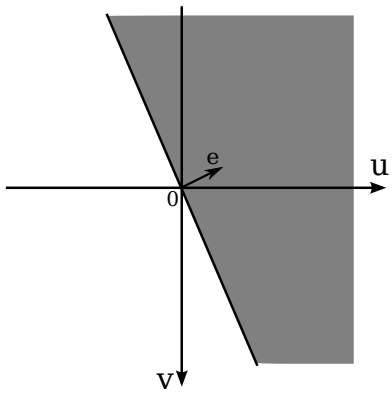


Figure 3.1: Spectral support of the 2D spectrum of the complex signal constructed using the pHT acting in the direction **e**.

axis perpendicular to **e** is suppressed because the signum function is zero at these frequencies. This limitation is addressed by the adjusted pHT in Section 3.2.4.

3.2.2 Total Hilbert Transform

While the pHT in Section 3.2.1 acts in one direction, the 2D total Hilbert transform (tHT) admits true 2D action. The 2D tHT is given by

$$q_{\text{tot}}(\mathbf{x}) = \mathcal{H}\{f(\mathbf{x})\} = \frac{1}{\pi^2}\,\text{p.v.}\int_{\mathbb{R}^2} \frac{f(\boldsymbol{\tau})}{\prod_{i=1}^{2}(x_i - \tau_i)}\,\mathrm{d}^2\boldsymbol{\tau}, \tag{3.4}$$

or equivalently in the Fourier domain as

$$\widehat{q}_{\text{tot}}(\boldsymbol{\omega}) = \mathcal{F}\{f(\mathbf{x}) * h(\mathbf{x})\} = (-j)^2\widehat{f}(\boldsymbol{\omega})\prod_{i=1}^{2}\text{sgn}(\omega_i), \tag{3.5}$$

where $\boldsymbol{\omega} = [w_1\ w_2]$. The spectrum of the complex 2D signal $z_{\text{tot}}(\mathbf{x})$ constructed by using the tHT approach is then given by

$$\widehat{z}_{\text{tot}}(\boldsymbol{\omega}) = \widehat{f}(\boldsymbol{\omega}) + j\widehat{q}_{\text{tot}}(\boldsymbol{\omega}) = \widehat{f}(\boldsymbol{\omega})\left[1 - j\prod_{i=1}^{2}\text{sgn}(\omega_i)\right]. \tag{3.6}$$

Similar to previous approaches, the real signal $f(\mathbf{x})$ can be recovered simply by taking the real part of the $z_{\text{tot}}(\mathbf{x})$.

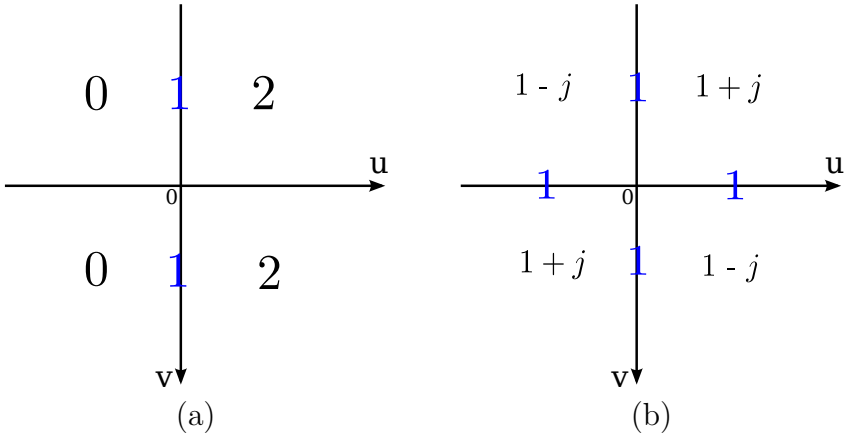


Figure 3.2: Spectral multiplier of the 2D complex signal: (a) of the pHT with $\mathbf{e} = [1\,0]^T$. (b) of the tHT.

I illustrate the distinction between the 2D spectral multipliers of the pHT and of the tHT in Fig. 3.2. The spectral multiplier for the 2D pHT with $\mathbf{e} = [1\,0]^T$ is shown in Fig. 3.2(a). The spectrum of $z_{\mathbf{e}}(\mathbf{x})$ is zero on the left half plane, one on the vertical axis, and two on the right half of the frequency plane. The multiplier of the tHT approach, on the other hand, retains full spectral support. The spectral multiplier of the tHT is depicted in Fig. 3.2(b). The unity gain symbols are exaggerated in these two figures as they occur only at locations where $\text{sgn}(\omega_1) = 0$ (for the pHT) and $\text{sgn}(\omega_1)\text{sgn}(\omega_2) = 0$ (for the tHT).

Compared to the pHT, the tTH does not provide an intuitive AM-FM representation [34]. Nevertheless, it plays an important role in the development of the single orthant complex signal model in Section 3.2.3 and the hypercomplex signal model in Section 3.2.5.

3.2.3 Single Orthant Complex Signal

Hahn [45] proposed the single orthant complex signal as a 2D extension of the 1D complex signal, where *orthant* is equivalent to *quadrant* in the 2D case. He

modified the HT spectral multiplier so that the spectrum of the 2D complex signal is zero in all but one orthant. The spectrum of the single orthant complex signal is defined as

$$\widehat{z}_{\rm so}(\boldsymbol{\omega}) = \widehat{f}(\boldsymbol{\omega}) \prod_{i=1}^{2} \Big(1 + \mathrm{sgn}(\omega_i)\Big). \tag{3.7}$$

Expanding (3.7), the Fourier spectrum of $\widehat{z}_{\rm so}(\boldsymbol{\omega})$ is

$$\begin{aligned} \widehat{z}_{\rm so}(\boldsymbol{\omega}) &= \widehat{f}(\boldsymbol{\omega})\big(1 + \mathrm{sgn}(\omega_1)\big)\big(1 + \mathrm{sgn}(\omega_2)\big) \\ &= \widehat{f}(\boldsymbol{\omega}) + \mathrm{sgn}(\omega_1)\widehat{f}(\boldsymbol{\omega}) + \mathrm{sgn}(\omega_2)\widehat{f}(\boldsymbol{\omega}) + \mathrm{sgn}(\omega_1)\mathrm{sgn}(\omega_2)\widehat{f}(\boldsymbol{\omega}) \\ &= \Big(\widehat{f}(\boldsymbol{\omega}) - \widehat{q}_{tot}(\boldsymbol{\omega})\Big) + j\Big(\widehat{q}_{\rm e_1}(\boldsymbol{\omega}) + \widehat{q}_{\rm e_2}(\boldsymbol{\omega})\Big), \end{aligned} \tag{3.8}$$

where $\widehat{q}_{\rm e_1}$ and $\widehat{q}_{\rm e_2}$ are defined in (3.2) with $\mathrm{e}_1 = [1\ 0]^T$ and $\mathrm{e}_2 = [0\ 1]^T$. From (3.8), we see that the spectrum of the complex signal in Hahn's approach contains contributions from both partial and total Hilbert transform terms. The 2D spectral support of $\widehat{z}_{\rm so}(\boldsymbol{\omega})$ is shown in Fig. 3.3, where the spectrum is quadrupled in the shaded region and is zero in the remaining white regions.

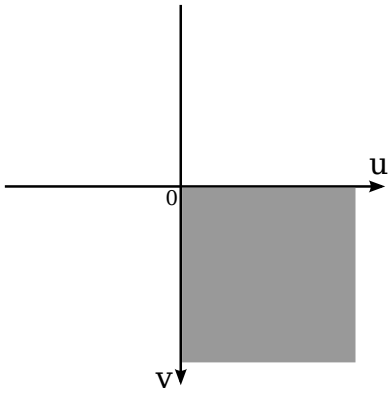


Figure 3.3: The 2D spectrum of single orthant complex signal

As Hahn's single orthant approach eliminates $2^n - 1$ orthants of the nD complex signal spectrum, the model is efficient in terms of digital storage. However,

the single orthant representation is insufficient to reconstruct the original real signal $f(\mathbf{x})$. In particular, $f(\mathbf{x})$ can not be obtained by taking the real part of (3.8). To recover the original real signal $f(\mathbf{x})$, an additional single orthant complex signal $z_{\mathrm{so}_1}(\mathbf{x})$ must be defined with the following spectrum

$$\begin{aligned}\widehat{z}_{\mathrm{so}_1}(\boldsymbol{\omega}) &= \widehat{f}(\boldsymbol{\omega})\Big(1-\mathrm{sgn}(\omega_1)\Big)\Big(1+\mathrm{sgn}(\omega_2)\Big) \\ &= \Big(\widehat{f}(\boldsymbol{\omega})+\widehat{q}_{tot}(\boldsymbol{\omega})\Big)+j\Big(-\widehat{q}_{\mathrm{e}_1}(\boldsymbol{\omega})+\widehat{q}_{\mathrm{e}_2}(\boldsymbol{\omega})\Big). \end{aligned} \tag{3.9}$$

Combining $\widehat{z}_{\mathrm{so}}(\boldsymbol{\omega})$ and $\widehat{z}_{\mathrm{so}_1}(\boldsymbol{\omega})$, the original signal is then recovered as

$$f(\mathbf{x}) = \frac{1}{2}\mathrm{Re}\{\mathcal{F}^{-1}\{\widehat{z}_{\mathrm{so}}(\boldsymbol{\omega})+\widehat{z}_{\mathrm{so}_1}(\boldsymbol{\omega})\}\}. \tag{3.10}$$

As for physical interpretation, the single orthant complex signal can not capture signal which admit orientations outside the support of $\widehat{z}_{\mathrm{so}}(\boldsymbol{\omega})$. For an image, oriented textures with spectral support lying in quadrant one and quadrant two can not be simultaneously represented by the single orthant signal.

3.2.4 Adjusted Hilbert Transform

Although the 2D pHT in Section 3.2.1 preserves most of the important properties of the 1D HT, it fails to retain the harmonic correspondence property for some pure sinusoidal signals. Let $\mathcal{Z} = \{\boldsymbol{\omega} : \boldsymbol{\omega}^T\mathbf{e} = 0\}$ with $\mathbf{e} = [1,0]^T$ be a subset of the 2D frequency plane. If a signal $f(\mathbf{x})$ has spectral support in $\mathcal{Z}$, the *ordinary* pHT suppresses it, because $\mathrm{sgn}(\boldsymbol{\omega}^T\mathbf{e}) = 0$ when $\boldsymbol{\omega}^T\mathbf{e} \in \mathcal{Z}$. Havlicek, Havlicek, and Bovik [50] proposed an *adjusted* multiplier for the 2D pHT as

$$\mathcal{F}\{\mathcal{H}_{\mathrm{adj}}\{f(\mathbf{x})\}\} = -j\mathrm{sgn}_{\mathrm{adj}}(\boldsymbol{\omega}^T\cdot\mathbf{e})\widehat{f}(\boldsymbol{\omega}), \tag{3.11}$$

where they defined $\mathrm{sgn}_{\mathrm{adj}}(\boldsymbol{\omega}^T\mathbf{e})$ as

$$\mathrm{sgn}_{\mathrm{adj}}(\boldsymbol{\omega}) = \sum_{\ell=1}^{2}\mathrm{sgn}(\omega_\ell)\prod_{k=1}^{\ell-1}(1-|\mathrm{sgn}(\omega_k)|). \tag{3.12}$$

In this definition, the adjusted multiplier $\mathcal{H}_{\text{adj}}$ differs from that in (3.2) only at frequencies in $\mathcal{Z}$. The *adjusted* HT spectral multiplier is further simplified as

$$\begin{aligned} \text{sgn}_{\text{adj}}(\omega_1, \omega_2) &= \text{sgn}(\omega_1) + \text{sgn}(\omega_2)[1 - |\text{sgn}(\omega_1)|] \\ &= \begin{cases} 1 & \omega_1 > 0 \\ 1 & \omega_1 = 0 \,\text{and}\, \omega_2 > 0 \\ -1 & \omega_1 < 0 \\ -1 & \omega_1 = 0 \,\text{and}\, \omega_2 < 0 \\ 0 & \omega_1 = \omega_2 = 0. \end{cases} \end{aligned} \tag{3.13}$$

I illustrate the difference between the *ordinary* HT 3.2 and the adjusted HT in Fig. 3.4. Notice that the two multipliers differ only on the vertical axis where the ordinary multiplier is zero and the adjusted multiplier takes on $+j$, 0, and $-j$. This modification allows the adjusted pHT to retain the harmonic correspondence property for all pure sinusoidal signals [50].

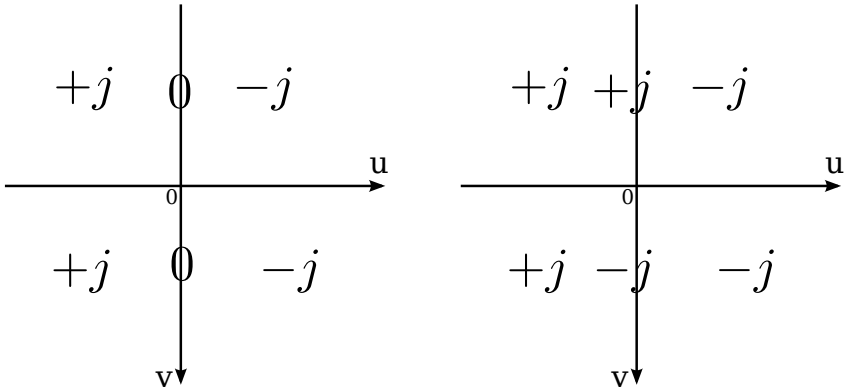


Figure 3.4: Illustration of the frequency multiplier of the ordinary pHT in (3.2) and frequency multiplier of the adjusted HT in (3.12). (a) Frequency multiplier of the pHT with $\mathbf{e} = [1\,0]^T$. (b) Frequency multiplier of the adjusted HT.

Like to complex signal (3.3) generated with the 2D pHT, the signal (3.11) constructed using the adjusted pHT fails to satisfy the multidimensional Cauchy

Riemann equations [50]. However, the adjusted pHT retains most of important properties of the 1D analytic signal:

1. $\mathcal{H}_{\text{adj}}$ is LTI, bounded, and has unity gain.
2. The complex signal generated with $\mathcal{H}_{\text{adj}}$ has frequency support only in the right half of the frequency plane.
3. $\mathcal{H}_{\text{adj}}$ satisfies Vakman's three conditions [124], e.g., amplitude continuity, amplitude homogeneity, and harmonic correspondence.

3.2.5 Hypercomplex Signal

Bülow and Sommer [19] proposed an extension of the 1D analytic signal for higher dimensions using the hypercomplex Fourier transform. The hypercomplex Fourier transform is an extension of the traditional Fourier transform in combination with hypercomplex number theory. The authors combined the hypercomplex Fourier transform and Hahn's single orthant approach [45] to construct a hypercomplex signal. In 2D, the hypercomplex signal is in the form of a quaternionic signal

$$p(\mathbf{x}) = f(\mathbf{x}) + ir(\mathbf{x}) + ju(\mathbf{x}) + kv(\mathbf{x}), \tag{3.14}$$

where $f(\mathbf{x})$, $r(\mathbf{x})$, $u(\mathbf{x})$, and $v(\mathbf{x})$ are real signals, and i, j, and k are the quaternionic imaginary units. The basic properties of the quaternionic units are $ij = k$, $jk = i$, $ki = j$, and $i^2 = j^2 = k^2 = -1$. These relations do not commute, i.e., $ji = -ij = -k$.

The spectrum of the hypercomplex signal $p(\mathbf{x})$ is computed using the quaternionic Fourier transform (QFT) [19] defined by

$$\widehat{f^q}(\boldsymbol{\omega}) = \int_{\mathbb{R}^2} e^{-i2\pi\omega_1 x_1} f(\mathbf{x}) e^{-j2\pi\omega_2 x_2} \mathrm{d}^2\mathbf{x}. \tag{3.15}$$

Since the QFT of a real signal $f(\mathbf{x})$ is quaternionic Hermitian [19], it is sufficient to retain only the spectral information from one orthant in the hypercomplex signal $p(\mathbf{x})$. Therefore, Bülow and Sommer [19] adopted Hahn's single orthant idea to construct the hypercomplex signal. The spectrum of the hypercomplex signal is given by

$$\widehat{p}^{q}_{\mathrm{so}}(\boldsymbol{\omega}) = \widehat{f}^{q}(\boldsymbol{\omega}) \prod_{i=1}^{n} \Big(1 + \mathrm{sgn}(\omega_i)\Big). \tag{3.16}$$

For a 2D signal, (3.16) may be rearranged to simplify the spectrum of the quaternionic signal according to

$$\begin{aligned} \widehat{p}^{q}_{\mathrm{so}}(\boldsymbol{\omega}) &= \widehat{f}^{q}(\boldsymbol{\omega})[1 + \mathrm{sgn}(\omega_1)][1 + \mathrm{sgn}(\omega_2)] \\ &= \widehat{f}^{q}(\boldsymbol{\omega})[1 + \mathrm{sgn}(\omega_1) + \mathrm{sgn}(\omega_2) + \mathrm{sgn}(\omega_1)\mathrm{sgn}(\omega_2)], \end{aligned} \tag{3.17}$$

whereupon the inverse QFT is applied to arrive at

$$p_{\mathrm{so}}(\mathbf{x}) = f(\mathbf{x}) + iq_{\mathrm{e}_1}(\mathbf{x}) + jq_{\mathrm{e}_2}(\mathbf{x}) + kq_{\mathrm{tot}}(\mathbf{x}). \tag{3.18}$$

Here, $q_{\mathrm{e}_1}(\mathbf{x})$, $q_{\mathrm{e}_2}(\mathbf{x})$ are pHT's (3.2) of $f(\mathbf{x})$ acting in the direction specified by the unit vectors $\mathbf{e}_1 = [1\,0]^T$ and $\mathbf{e}_2 = [0\,1]^T$, and $q_{\mathrm{tot}}(\mathbf{x})$ is the tHT (3.4). Unlike Hahn's single orthant approach [45], the hypercomplex representation provides direct access to the original signal $f(\mathbf{x})$ as its real part.

Once the hypercomplex signal $p_{\mathrm{so}}(\mathbf{x})$ is constructed, the AM function of $f(\mathbf{x})$ is then defined as

$$a(\mathbf{x}) = ||p_{\mathrm{so}}(\mathbf{x})|| = \sqrt{f^2(\mathbf{x}) + q^2_{\mathrm{e}_1}(\mathbf{x}) + q^2_{\mathrm{e}_2}(\mathbf{x}) + q^2_{\mathrm{tot}}(\mathbf{x})}. \tag{3.19}$$

However, the phase of the quaternionic signal is complicated and is not defined for all orientations [19]. The phase functions can be retrieve from the polar representation $p(\mathbf{x}) = |a(\mathbf{x})|e^{i\phi}e^{j\theta}e^{k\psi}$. Let $\tilde{p} = p(\mathbf{x}_0)/|a(\mathbf{x}_0)| = p_0 + ip_1 + jp_2 + kp_3$ be a

normalized quaternionic number at $\mathbf{x}_0$. The phase of the hypercomplex signal has three components that are point-wise given by

$$\psi = -\arcsin[2(p_1p_2 - p_0p_3)]/2 \tag{3.20}$$

$$\phi = \begin{cases} 0.5\arctan[2(-p_2p_3 + p_0p_1), p_0^2 - p_1^2 - p_2^2 + p_3^2], & \psi = \pm\pi/4, \\ 0.5\arctan[2(p_2p_3 + p_0p_1), p_0^2 - p_1^2 + p_2^2 - p_3^2], & \text{otherwise}, \end{cases} \tag{3.21}$$

$$\theta = \begin{cases} 0 & \psi = \pm\pi/4 \\ 0.5\arctan[2(p_1p_3 + p_0p_2), p_0^2 + p_1^2 - p_2^2 - p_3^2], & \text{otherwise}, \end{cases} \tag{3.22}$$

where I have dropped the spatial argument for brevity.

The hypercomplex signal addresses the shortcoming of Hahn's single orthant approach where the real signal $f(\mathbf{x})$ can not be directly reconstructed from the complex single orthant signal $z_{\text{so}}(\mathbf{x})$. In other words, the hypercomplex signal construction allows perfect reconstruction of $f(\mathbf{x})$ from $z_{\text{so}}(\mathbf{x})$ without defining an auxiliary hypercomplex signal.

While the mathematical foundations of the hypercomplex signal are solid and interesting, the hypercomplex signal has practical limitations. First, the phase functions computed from the hypercomplex signals are not defined for signals with dimension greater than two [19]. Second, even when the phase functions are defined for 2D signal, the computation of the phase is complicated and the computed phase functions do not correlate well with human visual perception of the image texture structure. Finally, it is not clear that the hypercomplex AM function satisfies the basic requirement that the AM should capture contrast as opposed to the local texture granularity and orientation.

3.2.6 Monogenic Signal

Felsberg and Sommer [34] introduced the Riesz transform as a natural 2D extension of the 1D HT. They reasoned that the 1D HT can be derived from 2D

harmonic fields and the 2D Riesz transform can be derived from 3D harmonic fields. Therefore, they claimed that this generalization is both natural and valid.

The Riesz transform maps a signal into a vector of signals. Let $\mathbf{h}_R(\mathbf{x})$ be the kernel of the Riesz transform. Its Fourier transform is given by

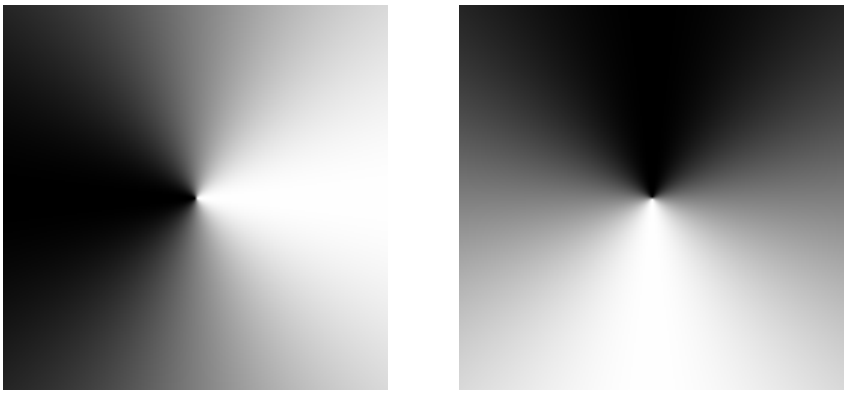
$$\widehat{\mathbf{h}}_R(\boldsymbol{\omega}) = \mathcal{R}\{\mathbf{h}_R(\mathbf{x})\}(\boldsymbol{\omega}) = -j\frac{\boldsymbol{\omega}}{||\boldsymbol{\omega}||}. \tag{3.23}$$

For the 2D case, using the Fourier transform differentiation identity, one can express the kernel of the Riesz transform (RT) in the spatial domain as [34, 123]

$$\mathbf{h}_R(\mathbf{x}) = \begin{pmatrix} h_1(\mathbf{x}) \\ h_2(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} \frac{x}{2\pi(x^2+y^2)^{3/2}} \\ \frac{y}{2\pi(x^2+y^2)^{3/2}} \end{pmatrix}, \tag{3.24}$$

where $\mathbf{x} = [x\, y]^T$ and $h_1(\mathbf{x})$ and $h_2(\mathbf{x})$ are two filters acting in the x and y direction respectively. The 2D RT $\mathcal{R}\{.\}$ of the real signal $f(\mathbf{x})$ is then obtained by time domain convolution according to

$$\mathcal{R}\{f\}(\mathbf{x}) = \begin{pmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} h_1(\mathbf{x}) * f(\mathbf{x}) \\ h_2(\mathbf{x}) * f(\mathbf{x}) \end{pmatrix}. \tag{3.25}$$



(a) (b)

Figure 3.5: Riesz transform frequency response: (a) $\widehat{h}_1(\boldsymbol{\omega})$. (b) $\widehat{h}_2(\boldsymbol{\omega})$.

I plotted the frequency responses of $h_1(\mathbf{x})$ and $h_2(\mathbf{x})$ in Fig. 3.5. The black corresponds to -1 and the white indicates $+1$. Similar to the 1D HT, the Riesz

kernels are not defined at the origin. However, the frequency responses of the Riesz transform are *smoother* compared to that of the partial Hilbert transform in Chapter 3.2.1. In other words, the HT multiplier contains sharp transitions from $-j$ to $+j$ in regions where the signum function changes sign, while the frequency responses of the Riesz filters vary smoothly as shown in Fig. 3.5. This difference is an important property that allows the Riesz transform to produce better AM function than the HT-based approaches.

Felsberg and Sommer [34] named the complex signal extension using Riesz transform the *monogenic* signal. In 2D, they defined the monogenic image in the hypercomplex plane as

$$z_{\mathrm{mo}}(\mathbf{x}) = f(\mathbf{x}) + if_1(\mathbf{x}) + jf_2(\mathbf{x}), \tag{3.26}$$

where i, j are a pair of orthogonal hypercomplex imaginary units with basic operations defined in Section 3.2.5.

The hypercomplex components of the monogenic signal are given by

$$\begin{aligned} f(\mathbf{x}) &= a_{\mathrm{mo}}(\mathbf{x})\cos[\varphi_{\mathrm{mo}}(\mathbf{x})], \\ f_1(\mathbf{x}) &= a_{\mathrm{mo}}(\mathbf{x})\sin[\varphi_{\mathrm{mo}}(\mathbf{x})]\cos[\phi(\mathbf{x})], \\ f_2(\mathbf{x}) &= a_{\mathrm{mo}}(\mathbf{x})\sin[\varphi_{\mathrm{mo}}(\mathbf{x})]\sin[\phi(\mathbf{x})]. \end{aligned} \tag{3.27}$$

From this model, the AM and FM functions of $f(\mathbf{x})$ can be calculate analytically. The AM function is computed as

$$a_{\mathrm{mo}}(\mathbf{x}) = |z_{\mathrm{mo}}(\mathbf{x})| = \sqrt{f^2(\mathbf{x}) + f_1^2(\mathbf{x}) + f_2^2(\mathbf{x})}. \tag{3.28}$$

The imaginary signal in the RT approach is then computed by

$$q_{\mathrm{mono}}(\mathbf{x}) = f_1(\mathbf{x})\cos[\phi(\mathbf{x})] + jf_2(\mathbf{x})\sin[\phi(\mathbf{x})], \tag{3.29}$$

where $\phi(\mathbf{x})$ is given by $\phi(\mathbf{x}) = \arctan[f_2(\mathbf{x})/f_1(\mathbf{x})]$. The monogenic signal can be rewritten in the regular complex signal form as

$$z_{\text{mono}}(\mathbf{x}) = f(\mathbf{x}) + jq_{\text{mono}}(\mathbf{x}) \tag{3.30}$$

Similar to the 1D HT, the RT preserves harmonic correspondence [34, 123]. In addition, it is an all-pass filter with unity gain. However, the key property of the RT is that the Riesz operator is isotropic [34, 123]. I illustrate in Section 3.4 that this property allows the monogenic image approach to produce artifact-free AM functions. On the other hand, the AM function obtained by the pHT approach contains rippling artifacts along a line normal to the direction of action of the pHT.

The computation of $\phi(\mathbf{x})$ involves the arctan operation. Therefore, $\phi(\mathbf{x})$ is wrapped between $-\pi/2$ and $\pi/2$. As a result, the imaginary image $q_{\text{mono}}(\mathbf{x})$ contains discontinuous of magnitude π at branch cuts in the arctan function. These discontinuous points can lead to artifacts in the computed FM function of the monogenic signal.

3.3 The 2D AM-FM Computation: Direct Approaches

While the complex signal extension approach is prevalent in the literature, there exist approaches that compute the AM and FM functions directly from the real signal $f(\mathbf{x})$, i.e., without constructing an explicit complex signal. Here, I discuss two important approaches, namely the 2D Teager-Kaiser energy operator (TKEO) by Maragos and Bovik [77] and the 2D extension of the 1D quasi-local approach [87].

3.3.1 The 2D Teager-Kaiser Energy Operator

Maragos and Bovik [77] proposed a 2D extension of the 1D TKEO described in Section 2.2.2 to extract the AM and FM functions from images. The AM and FM functions are computed by the continuous energy separation algorithm (CESA) as

$$a(\mathbf{x}) \approx \frac{\Psi[f(\mathbf{x})]}{\sqrt{\Psi\left[\frac{\partial f}{\partial x_1}\right] + \Psi\left[\frac{\partial f}{\partial x_2}\right]}}, \tag{3.31}$$

$$\frac{\partial \varphi}{\partial x_1} \approx \sqrt{\frac{\Psi\left[\frac{\partial f}{\partial x_1}\right]}{\Psi[f(\mathbf{x})]}}, \tag{3.32}$$

$$\frac{\partial \varphi}{\partial x_2} \approx \sqrt{\frac{\Psi\left[\frac{\partial f}{\partial x_2}\right]}{\Psi[f(\mathbf{x})]}}, \tag{3.33}$$

where $\mathbf{x} = [x_1\ x_2]^T$ and $\nabla\varphi(\mathbf{x}) = \left(\frac{\partial \varphi}{\partial x_1}, \frac{\partial \varphi}{\partial x_2}\right)$ is the FM function. $\Psi[f(\mathbf{x})]$ denotes the 2D Teager-Kaiser energy of the image $f(\mathbf{x})$ given by

$$\Psi[f(\mathbf{x})] = ||\nabla f(\mathbf{x})||^2 - f(\mathbf{x})\nabla^2 f(\mathbf{x}), \tag{3.34}$$

where $\nabla f(\mathbf{x}) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}\right)$.

Similar to the 1D algorithm, the discrete 2D DESA is a local method where the AM and FM function are computed from neighborhood points. Let $y_v(m,n) = f(m+1,n) - f(m-1,n)$ and $y_h(m,n) = f(m,n+1) - f(m,n-1)$ be two auxiliary discrete signals. The discrete version (DESA) of the CESA is given as

$$\begin{aligned} a(m,n) &\approx \frac{2\Psi_d[f(m,n)}{\sqrt{\Psi_d[y_h(m,n)] + \Psi_d[y_v(m,n)]}}, \\ \left|\frac{\partial \varphi}{\partial x_1}\right| &\approx \arcsin\left(\sqrt{\frac{\Psi_d[y_v(m,n)]}{4\Psi[f(m,n)]}}\right), \\ \left|\frac{\partial \varphi}{\partial x_2}\right| &\approx \arcsin\left(\sqrt{\frac{\Psi_d[y_h(m,n)]}{4\Psi[f(m,n)]}}\right). \end{aligned} \tag{3.35}$$

The 2D DESA can estimate AM and FM function up to one fourth of the sampling frequency [77], i.e., $0 \leq |\frac{\partial \varphi}{\partial x_1}|, |\frac{\partial \varphi}{\partial x_2}| \leq \pi/2$. Similar to the 1D DESA, the 2D DESA can not deduce the sign of the FM function.

3.3.2 The 2D Quasi-local Approach

Murray, Rodríguez, and Pattichis [87] extended the 1D quasi-local method originally proposed by Girolami and Vakman [43] to 2D to compute AM and FM functions for images. The 2D method is a separable implementation of the 1D quasi-local method in Section 2.2.3, i.e., the FM in the horizontal direction is obtained by treating the 2D image as a set of 1D signals along rows of the image and applying the 1D quasi-local algorithm to these 1D signals.

As noted in the 1D case, the quasi-local method can estimate frequencies up to one fourth of the sampling frequency. This restriction holds for the 2D case as well. For signals with maximum IF above one fourth of the sampling frequency, it is possible to resample the input signal before applying the quasi-local method [43]. Murray, Rodríguez, and Pattichis [87] provided an extension to the 2D quasi-local method to account for frequencies higher than one fourth of the sampling frequency. They used two filters with disjoint passbands to compute the FM for entire discrete frequency range $[-\pi\ \pi]$.

The 2D quasi-local method faces the same limitations as in the 1D case because it depends on the bandwidth of the lowpass filter. In order to get good demodulation results, the lowpass filter must filter the AM function and completely reject the FM frequencies . This requirement is difficult to satisfy in practice because the variation of the FM has to be known a priori. As a result, for an arbitrary image, the 2D quasi-local method has limited use.

3.4 Comparison of the Complex Extension Approaches

I evaluated the effectiveness of the four complex extension approaches, e.g., single orthant complex signal by Hahn [45], adjusted Hilbert transform by Havlicek, Havlicek, and Bovik [49], hypercomplex signal by Büllow and Sommer [19], and monogenic signal by Felsberg and Sommer [34]. I did not consider the direct approaches because of the limitations of the quasi-local method and the difficulty in determining the relative signs of the FM components with the TKEO. Although the complex extension approaches can all be generalized to arbitrary dimensions, I restricted my attention to the 2D case. Since the phase unwrapping process was not discussed in this chapter, the phase functions are not compared. In particular, I computed the AM function and FM function of model 3.1 using (2.10) and (2.11).

Fig. 3.6 shows the AM and FM representation of the test image *chirp*. The AM and FM functions are given by Fig. 3.6(b) and (c). This image has a Gaussian AM function, a circularly symmetric quadratic phase, and a circularly symmetric linear FM function. In 2D, the FM function is computed by taking the gradient of the phase function. I plotted the FM frequency vector as a needle diagram. The magnitude of each needle is normalized for display purpose. The orientation of each needle represents the orientation of the FM vector in the original image.

As discussed in Section 3.2.5, the hypercomplex signal model does not lead to direct intuitive interpretation of the model (3.1). Therefore, results for the hypercomplex signal are shown separately in Fig. (3.7). The computed AM function is depicted in Fig. 3.7(a) and its corresponding three phase components are illustrated in Fig. 3.7(b),(c),(d) respectively. For the remaining three approaches, I compared the computed AM and FM functions of each approach side by side. The computed AM functions are shown Fig. 3.8 and the computed FM functions are

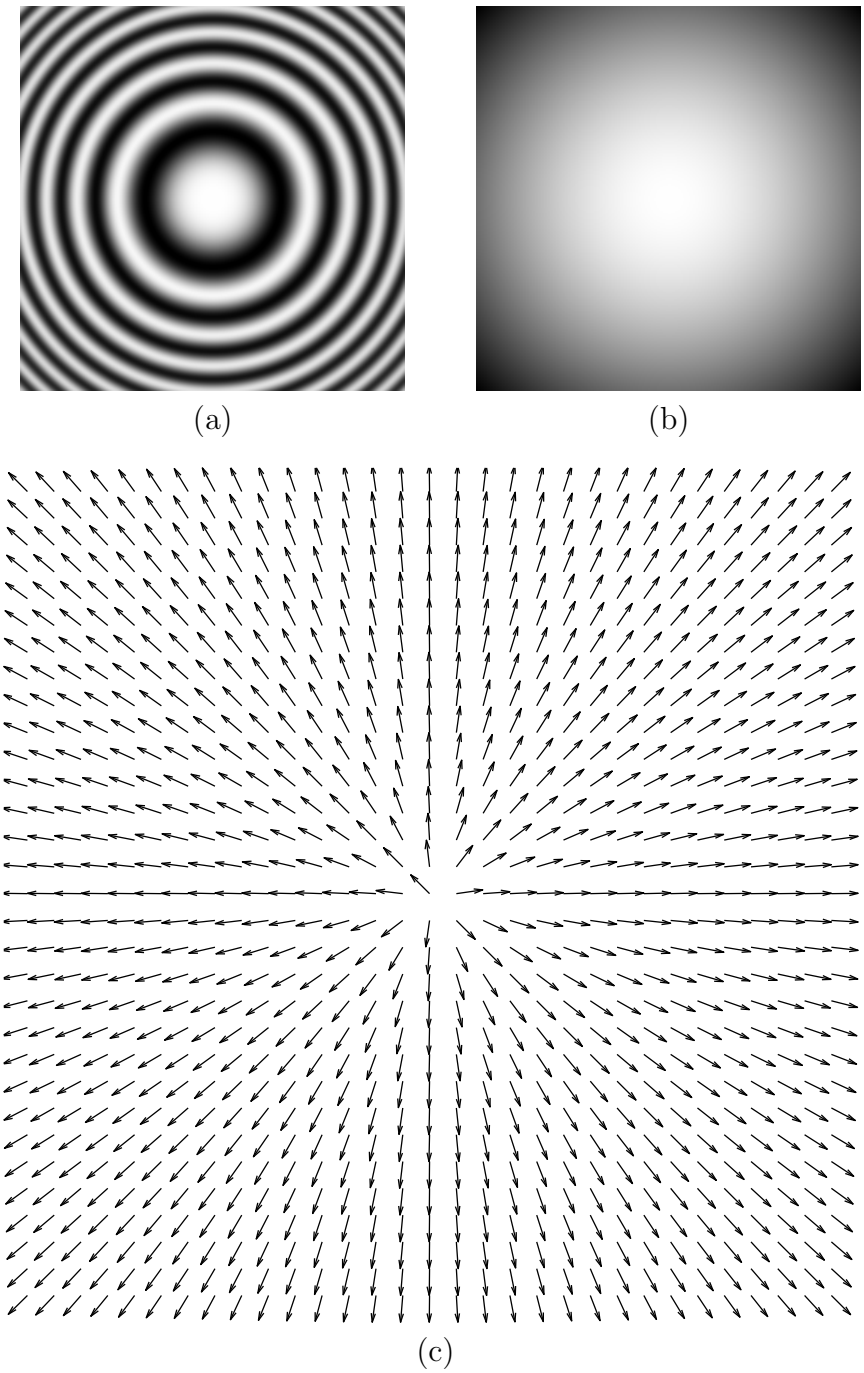


Figure 3.6: Chirp image: (a) Original image. (b) AM function. (c) FM function.

shown in Fig. 3.9. For quantitative evaluation, I provide the mean squared error (MSE) comparison for the computed AM in Table. 3.1.

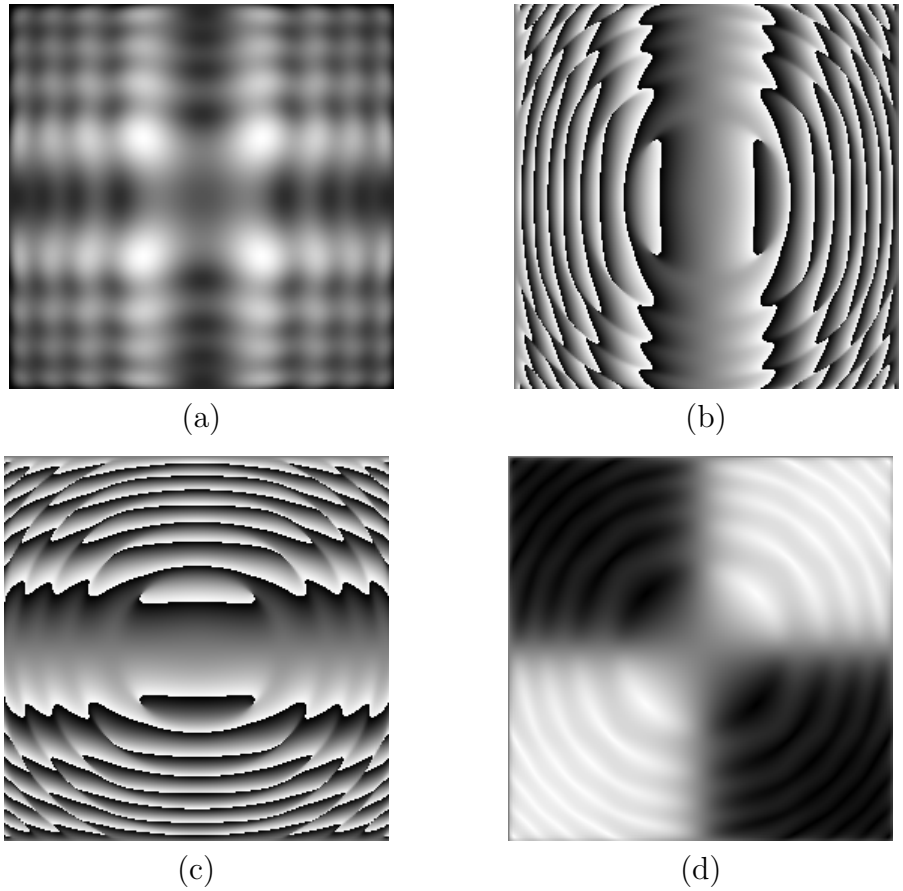


Figure 3.7: AM and phase of hypercomplex signal: (a) AM function. (b) First phase component $\phi(\mathbf{x})$ in (3.22). (c) Second phase component $\theta(\mathbf{x})$ in (3.22). (d) Third phase component $\psi(\mathbf{x})$ in (3.22).

Observe that for the single orthant approach, both the computed AM function in Fig. 3.8(b) and the computed FM function in Fig. 3.9(b) fail to fully capture the original AM and FM functions. Both the estimated AM function and the estimated FM function are valid in some parts of the image, e.g. regions where the AM is large in Fig. 3.8(b).

Among the four complex extension approaches, the adjusted Hilbert trans-

Table 3.1: Mean squared error (MSE) comparison of the computed AM with respect to the original AM in Fig. 3.6(b) of four complex extension approaches.

	Single Orthant	Hypercomplex	Adjusted HT	Monogenic
MSE	0.4449	0.1398	0.0295	0.0440

form and the monogenic signal yield the best results in term of meaningful interpretations of the AM and FM. While the AM of the adjusted HT in Fig. 3.8(c) yields smaller error with respect to the original AM function (0.0295 versus 0.0440), the computed AM function of monogenic signal in Fig. 3.8(d) is visually closer to that of the original signal. In addition, both approaches also produce satisfactory FM results. Over most of the image, the FM vectors in Fig. 3.9(b) and Fig. 3.9(c) are correctly oriented normal to important structures of the chirp image. However, both fail to provide meaningful FM vectors along the vertical shear line running down the center of the image.

I repeated the same experiment with the well-known test image *Lena* shown in Fig. 3.10(a). I computed the AM function and wrapped phase function using all four complex extension approaches. Observe that none of the approaches deliver an intuitively satisfying AM functions and phase. This is because the model (3.1) is a valid representation only for locally narrowband signals, while the 2D signal like Lena image in Fig. 3.10(a) is decidedly wideband. In practice, wideband signals must be decomposed into locally narrowband components, e.g., by a filterbank, before one can apply any of these complex signal extension approaches.

The 2D pHT extends the notion of 1D HT by performing the 1D HT in a specified direction. Similar to the 1D case, the spectral redundancy of the partial HT is reduced by 50%. The 2D pHT, however, fails to satisfy the harmonic correspondence property if the real signal has frequency support in the region $\mathcal{Z}$

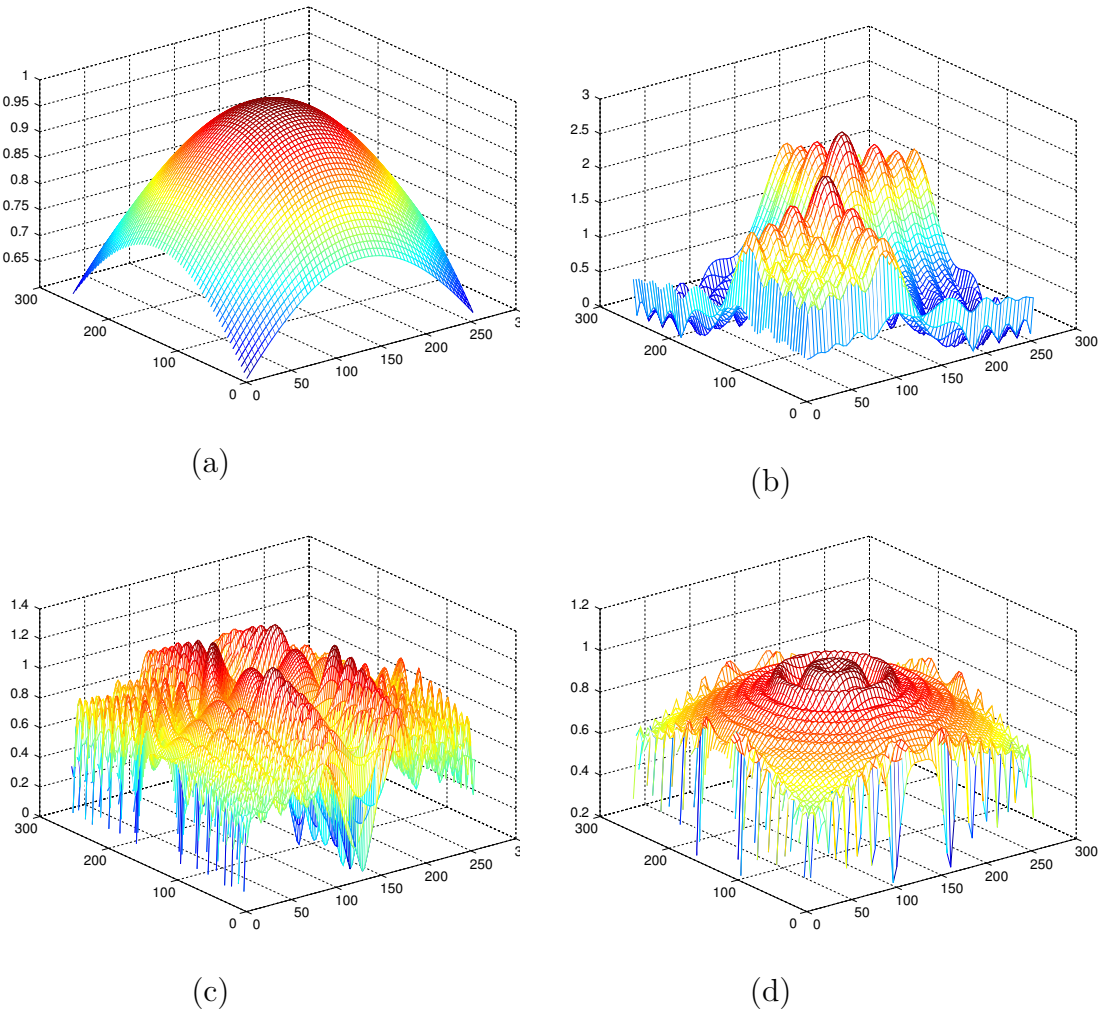


Figure 3.8: AM results for the *chirp* image computed using three different complex signal extensions. (a) Original AM. (b) Single orthant AM. (c) Adjusted HT AM. (d) Monogenic AM.

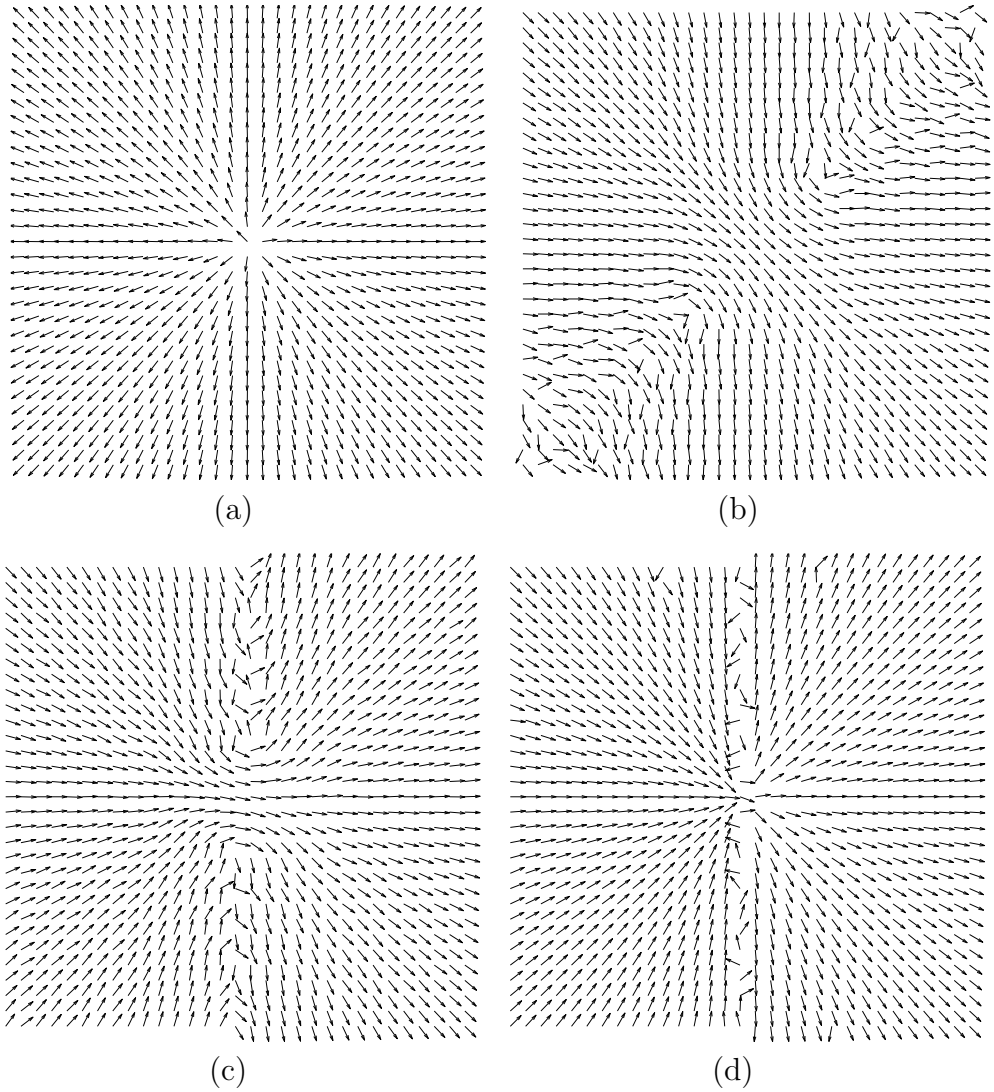


Figure 3.9: FM results for the *chirp* image computed using three different complex signal extensions. (a) Original chirp FM. (b) Single orthant FM. (c) Adjusted HT FM. (d) Monogenic FM.

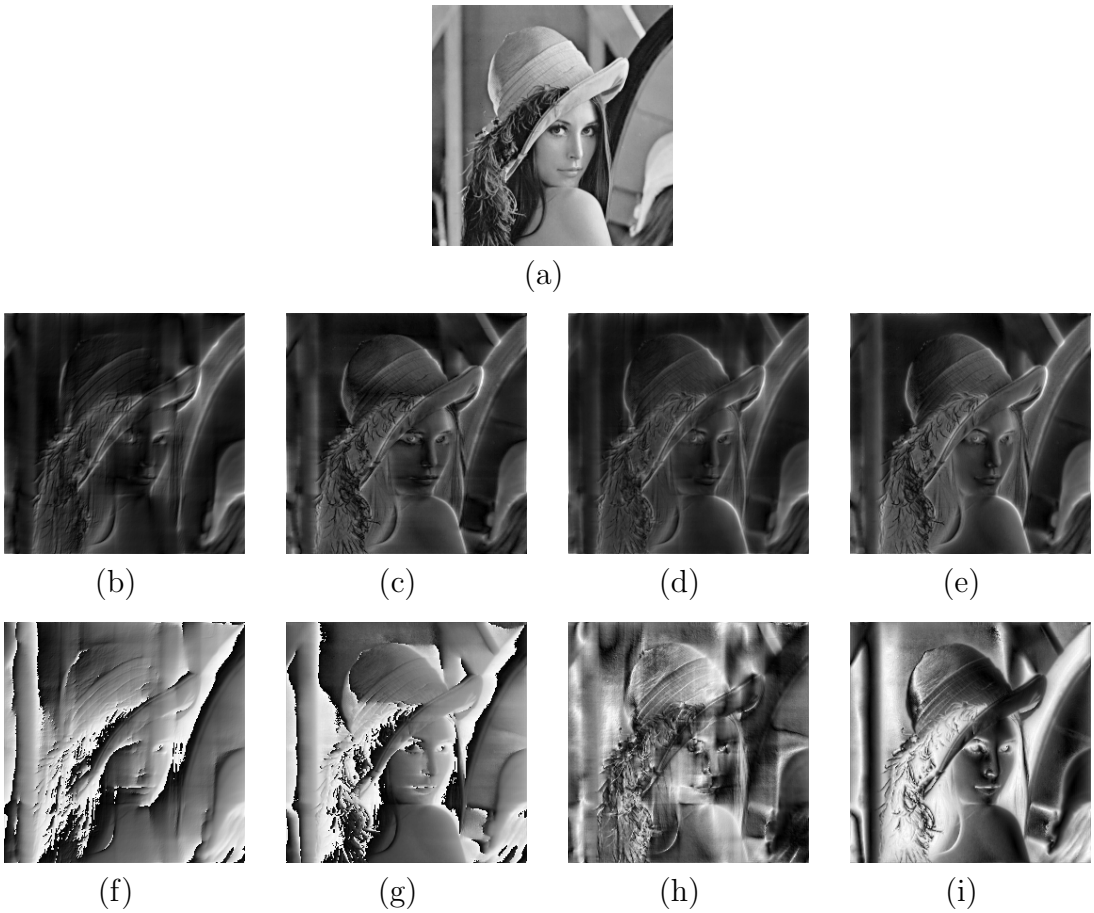


Figure 3.10: AM-FM representations of *Lena* image: (a) Original image. (b) Single orthant AM. (c) Adjusted HT AM. (d) Hypercomplex AM. (e) Monogenic AM. (f) Single orthant phase. (g) Adjusted HT phase. (h) Hypercomplex third phase component $\psi(\mathbf{x})$. (i) Monogenic phase.

as described in Section 3.2.4. The 2D tHT extends the 1D HT by applying a true 2D transform. Nevertheless, this transform fails to produce visually meaningful AM and FM functions [19]. The single orthant complex signal approach proposed by Hahn [45] eliminates all but one orthant of the signal's spectrum. However, the representation is not complete, i.e., the real signal can not be directly recovered from the single orthant complex signal. For a 2D signal, the single orthant complex signal model is limited to signals whose frequencies lie in the first and third quadrant of the frequency plane, which is the regions in Fig. 3.2 where $\text{sgn}(\omega_1)\text{sgn}(\omega_2) \geq 0$. Consequently, the model can not be used to represent arbitrary images as demonstrated in Fig. 3.8(b) and Fig. 3.9(b).

The 2D adjusted HT proposed by Havlicek, Havlicek, and Bovik [50] preserves most of the properties of the 1D analytic signal. More importantly, it satisfies the harmonic correspondence property for all pure sinusoidal signals. The harmonic correspondence is made possible by an adjustment in the frequency multiplier of the transform. The adjustment is illustrated by the spectral multiplier shown in Fig. 3.2. Furthermore, the adjusted HT is able to produce visually meaningful phase function. The computed AM function from this approach suffers from discontinuities and rippling artifacts around the operating axis of the pHT filter. While the AM function obtained with this approach has the lowest MSE with respect to the true AM of the original signal, it is not the most visually satisfying AM result.

The hypercomplex signal approach extends the 1D analytic signal to higher dimensions using the hypercomplex Fourier transform. For a 2D signal, the hypercomplex signal is constructed using the quaternionic Fourier transform (QFT). Since the 2D QFT of a real signal is quaternionic Hermitian, the entire spectrum

can be recovered from a single orthant. The hypercomplex signal construction uses Hahn's single orthant approach to reduce the spectral redundancy to one orthant in the hypercomplex frequency plane. In the image domain, the hypercomplex signal is a combination of the pHT and the tHT . The computed AM and FM functions of hypercomplex signal do not correspond with human visual perception. In addition, the phase is not well-defined for signals with dimensions greater than two. In the 2D case, even when the phase is mathematically defined, the phase calculation is not straight forward. Despite the fact that the hypercomplex signal retains only one orthant of spectral information, the amount of data for storage in this orthant is two times larger than that of the pHT.

The monogenic signal uses the Riesz transform as an extension of the 1D HT into higher dimensions. In contrast to the pHT where there is only one imaginary component in the complex signal, the monogenic signal has two. The monogenic representation provides meaningful interpretation of the AM function because the Riesz transform is an isotropic operator. In addition, the Riesz transform also satisfies the harmonic correspondence property. In addition, the monogenic FM function generally gives a visually meaningful representation of the image structure. Similar to the adjusted HT case, the monogenic FM has problems in the computation of the FM function. In particular, the monogenic signal requires an auxiliary orientation estimation step in order to compute the correct FM function. In term of spectral reduction, the monogenic signal does not offer any spectral cancellation compared to other approaches.

Mathematical construction and important properties of complex extension approaches are summarized in Table. 3.2 and Table. 3.3.

3.5 Summary

In this chapter, I discussed the computational aspect of the 2D AM-FM image model. I argued for the importance of the phase and the instantaneous frequency in two dimensional image processing applications. I discussed major approaches to compute the 2D AM-FM image model and compared their performance. The performance comparison was measured in term of the mean squared error (MSE) between the computed AM functions and their theoretical values.

Among the approaches I discussed for computing 2D AM-FM image models, the complex signal extension approaches are more popular in the literature than the direct approaches. The pHT and the monogenic signal yield satisfactory results as they compute low MSE demodulated functions and offer meaningful interpretations of the underlying image. Compared to the monogenic signal, the pHT has a simpler representation, i.e., two components versus three components. I adopted the pHT as the AM-FM computation algorithm for the perfect reconstruction AM-FM transform, which I will discussed in Chapter 4. While these two approaches use different signal transformations, there is a connection between them as I will show in Chapter 7.

I concluded this chapter by arguing that the quality of computed AM and

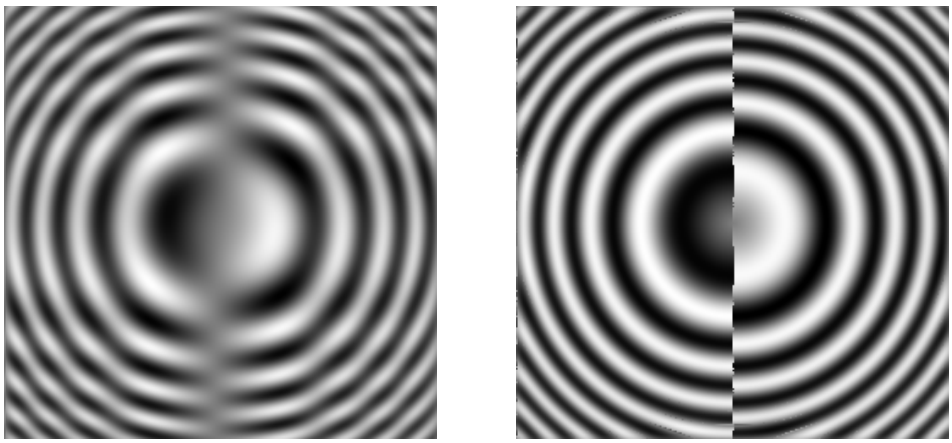
Table 3.2: Construction comparison of different n-D complex signal extensions

	Frequency Multiplier	$\mathrm{Re}\,[z(\mathbf{x})] = f(\mathbf{x})$	Spectral Reduction
Partial Hilbert	$-j\mathrm{sgn}(\boldsymbol{\omega}^T \cdot \mathbf{e})$	Yes	50%
Total Hilbert	$(-j)^n \prod_{i=1}^{n} \mathrm{sgn}(\omega_i)$	Yes	0%
Single Orthant	$\prod_{i=1}^{n} \big(1 + \mathrm{sgn}(\omega_i)\big)$	No	75%
Adjusted Hilbert	$-j\mathrm{sign}_{\mathrm{adj}}(\boldsymbol{\omega}^T \cdot \mathbf{e})$	Yes	50%
Hypercomplex Signal	$\prod_{i=1}^{n} \big(1 + \mathrm{sgn}(\omega_i)\big)$	Yes	*75%
Monogenic Signal	$-j\frac{\boldsymbol{\omega}}{\|\|\boldsymbol{\omega}\|\|}$	Yes	N/A

Table 3.3: Comparison of main properties of different approaches including isotropic, meaningful phase interpretation, and harmonic correspondence

	Isotropic	Meaningful Phase	Harmonic Correspondence
Partial Hilbert	No	Almost	Almost
Total Hilbert	No	No	No
Single Orthant	No	Inadequate	No
Adjusted Hilbert	No	Almost	Yes
Hypercomplex Signal	No	Inadequate	No
Monogenic Signal	Yes	Almost	Yes

FM functions using the complex extension approaches depends on the construction of the imaginary component. I illustrate the imaginary components of the adjusted HT approach in Fig. 3.11(a) and of the monogenic approach in Fig. 3.11(b). Observe that the locations where the imaginary signals exhibit discontinuities coincide with where the FM of these approaches have artifacts. The question of how to construct a complex-valued image without introducing such discontinuities remains an open research problem.



(a) (b)

Figure 3.11: Imaginary image of the complex signal models for the chirp image: (a) of the adjusted HT approach (b) of the monogenic signal approach.

Chapter 4

The Perfect Reconstruction AM-FM Image Model

In Chapter 3, I discussed different algorithms for computing the AM and FM functions of an image. In theory, these modulating functions can be used to define a complete AM-FM image representation. However, they have primarily been used as features for higher level image processing applications such as target tracking, image segmentation, stereo vision, and texture analysis [52]. In these applications, visually important image features are characterized in terms of AM-FM functions and used as input to higher level processes such as recognition or classification. Such applications are not concerned with recovering the original image from the computed modulating functions. In other words, most AM-FM image processing applications have been limited to image analysis. The image *synthesis* process, where the image is recovered from the modulating functions, has not been widely investigated. The main reason is that a perfect reconstruction AM-FM transform suitable for image synthesis did not exist. I first address this problem by developing the perfect reconstruction AM-FM image transform for a one component image.

I begin by discussing the importance of the 2D phase unwrapping problem in image synthesis in Section 4.1. I then review major approaches for computing the unwrapped phase in 2D in Section 4.2, with an emphasis on the 2D least-squares phase unwrapping technique in Section 4.3. I discuss the splined-based FM perfect reconstruction technique in Section 4.4.3. In addition, I propose the

least-squares FM perfect reconstruction technique in Section 4.4.4. Finally, the perfect reconstruction AM-FM transform for single component images is the focus of Section 4.5. A summary of the chapter is given in Section 4.6.

4.1 The Importance of Phase Unwrapping

The AM-FM image synthesis process aims to compute the original signal $f(\mathbf{x})$ from the AM and FM functions. Since the original image is defined as $f(\mathbf{x}) = a(\mathbf{x})\cos[\varphi(\mathbf{x})]$, this process requires reconstructing the phase function $\varphi(\mathbf{x})$ from the computed FM function $\nabla\varphi(\mathbf{x})$. In particular, $\varphi(\mathbf{x})$ is obtained by integrating the gradient field $\nabla\varphi(\mathbf{x})$. However, the true phase function $\varphi(\mathbf{x})$ can not be obtained from the measured gradient easily in 2D. Missing data samples, noisy measurements, and quantization errors at the receivers cause the measured gradient $\nabla\varphi_{\text{measured}}(\mathbf{x})$ to be different from the true gradient $\nabla\varphi(\mathbf{x})$. Hence, the integrated phase $\varphi_{\text{interg}}(\mathbf{x})$ does not equal the true phase function $\varphi(\mathbf{x})$. Consequently, the reconstructed image $f_{\text{interg}}(\mathbf{x}) = a(\mathbf{x})\cos[\varphi_{\text{interg}}(\mathbf{x})]$ differs from the original image $f(\mathbf{x})$. This problem has prevented the development of AM-FM synthesis applications.

One possible way to avoid phase integration is to obtain $\varphi(\mathbf{x})$ directly from the complex image $z_{\mathbf{e}}(\mathbf{x})$ as

$$\varphi(\mathbf{x}) = \arctan\left[\frac{q_{\mathbf{e}}(\mathbf{x})}{f(\mathbf{x})}\right]. \tag{4.1}$$

Because arctan is a multi-valued function, the phase $\varphi(\mathbf{x})$ in (4.1) is always computed as the principal value, which lies in $[-\pi, \pi]$. This principal value phase function is referred to as $\varphi_{\text{wrap}}(\mathbf{x})$. As discussed in Section 3.2.1, the complex function $z_{\mathbf{e}}(\mathbf{x})$ is the complex image constructed using the partial Hilbert trans-

form (pHT) approach. The real part of the complex image is the original image $f(\mathbf{x})$ and the imaginary part $q_{\mathbf{e}}(\mathbf{x})$ of the complex image is obtained by the pHT. The AM is computed by taking the magnitude of the complex image

$$a(\mathbf{x}) = \sqrt{f^2(\mathbf{x}) + q_{\mathbf{e}}^2(\mathbf{x})}. \tag{4.2}$$

With the computed AM and the computed phase $\varphi_{\text{wrap}}(\mathbf{x})$ defined as in (4.2) and (4.1), the original image $f(\mathbf{x})$ may be perfectly reconstructed. However, the FM functions calculated from $\varphi_{\text{wrap}}(\mathbf{x})$ do not correspond to human visual perception of the image. This problem is caused by the multi-valued nature of the arctangent where arguments differing by integer multiples of π yield the same value. As a result, the computed FM functions $\nabla\varphi_{\text{wrap}}(\mathbf{x})$ contain discontinuities at locations where the branch cuts in $\varphi_{\text{wrap}}(\mathbf{x})$ occur. These discontinuities arise solely from the branch cuts in the arctangent function and do not correspond to any visual information in the image. Moreover, because they are unbounded, they introduce large errors in the computed FM function and violate the smoothly varying assumption of the phase function in (3.1). Fig. 4.1 illustrates some of the problems associated with the wrapped phase $\varphi_{\text{wrap}}(\mathbf{x})$. Fig. 4.1(c) shows the phase function $\varphi_{\text{wrap}}(\mathbf{x})$ computed using (4.1) for a narrowband component of the well-known test image Barbara. The component is given in Fig. 4.1(a). Observe that $\varphi_{\text{wrap}}(\mathbf{x})$ in Fig. 4.1(c) contains many 2D discontinuous (e.g., bifurcations) introduced by the branch cuts in the arctangent function. The FM field computed from the wrapped phase is shown in Fig. 4.1(e), where many mathematically correct but visually meaningless needles with large magnitudes appear. In contrast, the unwrapped phase is shown in Fig. 4.1(d). The FM field computed from the unwrapped phase is depicted Fig. 4.1(f), where the needle length and orientation

are consistent visual perception of the image component. This example demonstrates that the 2D phase unwrapping process is a crucial step in constructing a perfect reconstruction FM transform.

4.2 The 2D Phase Unwrapping Problem

Phase unwrapping is an important step in many applications that use the phase information. In synthetic aperture radar (SAR) interferometric imaging, the phase at a given point indicates the terrain evaluation height [115]. In fiber-optic interferometry, the phase represents the depth of the imaged object [44]. In magnetic resonance (MR), the phase contains information about flow or inhomogeneities in the magnetic field. However, in these applications direct measurement of the phase is not possible. For example, in compensated imaging, one can only obtain phase differences from the receivers [39, 56]. In other applications, only the wrapped phase is measurable. However, the wrapped phase does not provide an intuitive way to present or perform analysis on the observed phenomena. Due to the image processing emphasis of this dissertation, I limited the discussion of the phase unwrapping problem to 2D signals.

2D phase unwrapping aims to find the unwrapped phase function from its principal (wrapped) values in the range $[-\pi, \pi]$. In other words, given a continuous function obtained by an inversion of a trigonometric function, e.g., $\arccos[\varphi(\mathbf{x})]$, the goal is to find a smooth function $\varphi(\mathbf{x})$ such that its range is no longer restricted in $[-\pi, \pi]$. Fig. 4.2 shows the difference between the wrapped phase and the unwrapped phase. The wrapped phase in Fig. 4.2(a) contains discontinuous points because its range is restricted to $[-\pi, \pi]$. Fig. 4.2(b) is the unwrapped version of the phase in Fig. 4.2(a). Notice that the unwrapped phase is *smoother* compared

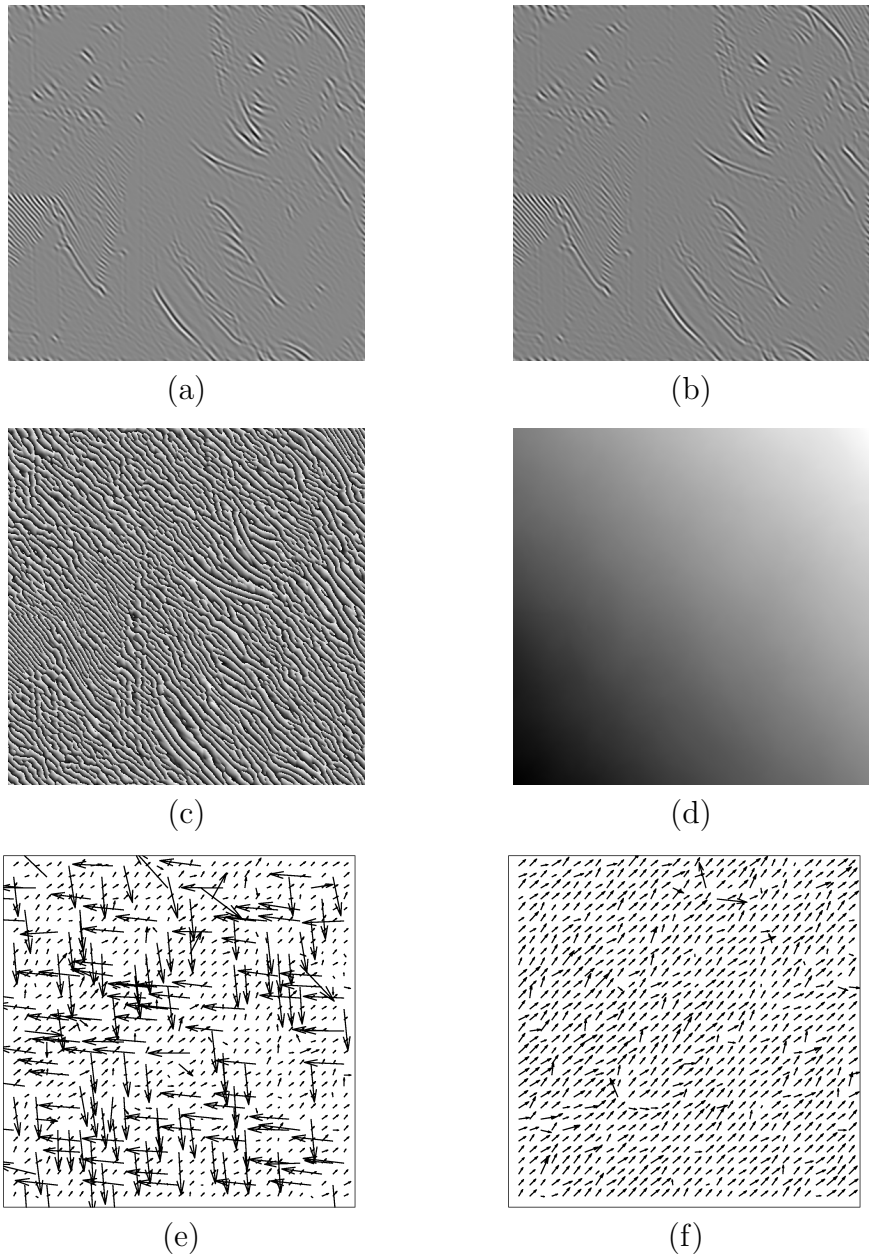
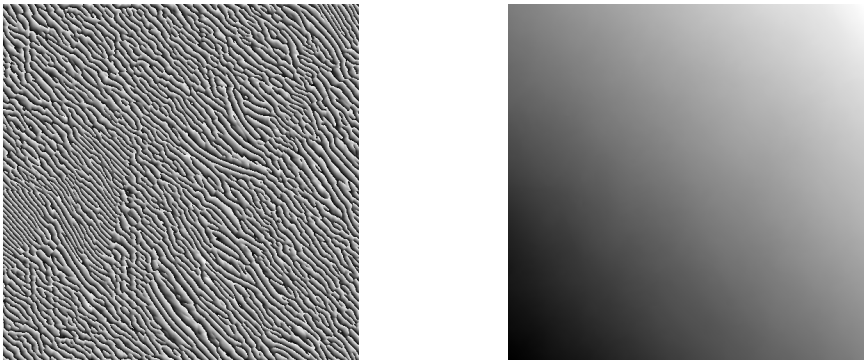


Figure 4.1: 2D Phase unwrapping of one narrowband component of the *Babara* image. (a) Real image component. (b) Imaginary image component. (c) Wrapped phase function. (d) Unwrapped phase function. (e) Wrapped FM field. (f) Unwrapped FM field.

the wrapped phase while still mapping through the cosine function.



(a) (b)

Figure 4.2: Wrapped phase *vs.* unwrapped phase. (a) Wrapped phase. (b) Unwrapped phase.

Formally, let $\varphi(\mathbf{x})$ be a continuous phase function. Let $\mathcal{W}[.]$ be the wrap operator such that $\mathcal{W}[\varphi(\mathbf{x})] \in [-\pi, \pi]$. Given $\mathcal{W}[\varphi(\mathbf{x})]$, the task is to compute the unwrapped phase $\varphi(\mathbf{x})$ such that

$$\varphi(\mathbf{x}) = \mathcal{W}[\varphi(\mathbf{x})] + p(\mathbf{x})\pi, \tag{4.3}$$

where $p(\mathbf{x}) \in \mathbb{Z}$. Unfortunately, 2D phase unwrapping is an ill-posed problem [39,117]. Therefore, phase unwrapping is usually formulated as an optimization problem. In this chapter, I give a short summary of the major 2D phase unwrapping approaches. I organize the 2D phase unwrapping algorithms into four main categories: path integration, energy norm minimization, model-based estimation, and Bayesian-based estimation.

4.2.1 Numerical Path Integration

Goldstein, Zebker, and Werner [44] proposed a path integration phase unwrapping algorithm. They provided an algorithm, *branch cut*, to detect local errors

caused by large phase discontinuities in order to prevent discontinuous points from contributing to the global phase reconstruction. The unwrapped phase is then obtained by performing path integration by knowing local horizontal and vertical derivatives. The numerical integration process must not cross the cut boundaries detected by the branch cuts algorithm.

4.2.2 Least-squares Energy Minimization

Fried [39] and Hudgin [56] formulated the least-squares phase reconstruction problem for the wave-front sensor application. They aimed to minimize the sum of errors between the phase differences and gradient of the unwrapped phase. Hunt [57] cast the phase reconstruction problem in a linear algebra setting and proposed a method to improve the convergence rate of the phase solution. Takajo and Takahashi [117] studied the least-squares phase reconstruction and introduced conditions where the solution is unique. They then proposed a closed-form non-iterative algorithm in the frequency domain to solve for the phase function [118]. Ghiglia and Romero [42] extended the least-squares phase reconstruction approach to facilitate weighted contributions of measured phase differences. They proposed two iterative algorithms to solve for the unwrapped phase. Strand, Taxt, and Jain proposed a block-based phase unwrapping algorithm [115]. Bioucas-Dias and Valadão [8] proposed an energy minimization framework for 2D phase unwrapping based on graph cuts. Spagnolini [112] used the IF estimated directly from the signal instead of from the wrapped phase. The estimated IF is then used in 2D phase unwrapping and the phase is solved by the least-squares framework. Because the unwrapped phase values differ from those of the wrapped by multiples of 2π, Costantini [26] formulated the phase unwrapping problem as an energy

minimization with integer variables.

4.2.3 Model-based Parameter Estimation

Friedlander and Francos [40] proposed a parametric model for phase unwrapping. First, they used a 2D polynomial model to fit the observed phase. The estimated phase was then used to guide the phase unwrapping process where the phase of each sample was corrected by adding or subtracting a multiple of 2π based on the difference between the principle value of the phase and the estimated phase. For general 2D phase unwrapping, the observed wrapped phase signal was segmented prior to the model fitting process.

4.2.4 Bayesian Phase Unwrapping

Nico, Palubinskas, and Datcu [93] applied the Bayesian framework to the phase unwrapping problem. Because measurement noise and phase aliasing caused inaccuracies in the least-squares phase reconstruction solution, they advocated the use of one and two regularization terms to enforce the phase prior models. However, it is not easy to find a suitable prior models given an arbitrary phase function.

4.3 Least-squares Phase Unwrapping for the 2D AM-FM Image Model

Let $\varphi(\mathbf{x})$ be the true unwrapped phase function. Let $\nabla\rho(\mathbf{x})$ be the measured phase gradient. The least-squares phase unwrapping approach finds the unwrapped phase $\varphi(\mathbf{x})$ by minimizing the mean squared error between the gradient of $\varphi(\mathbf{x})$ and the measured phase differences $\nabla\rho(\mathbf{x})$. In other words, the relationship between the

measured gradient and the true gradient is given by

$$\nabla\varphi(\mathbf{x}) = \nabla\rho(\mathbf{x}) + d(\mathbf{x}), \tag{4.4}$$

where $d(\mathbf{x})$ models the errors in the IF measurement or estimation process.

In practice, the acquired measurements are discrete. Therefore, the finite difference is often used to approximate the derivative operator. For example, $[\varphi(m,n) - \varphi(m-1,n)]$ and $[\varphi(m,n) - \varphi(m,n-1)]$ are approximations of the vertical and horizontal derivatives of $\varphi(m,n)$ at pixel location (m,n) in the $[0\cdots M-1]\times[0\cdots N-1]$ rectangular grid. Specifically, let $\rho_m(m,n)$ and $\rho_n(m,n)$ be the measured phase gradient, i.e., $\rho_m(m,n)$ and $\rho_n(m,n)$ are discrete approximations of the vertical and the horizontal derivatives of the measured phase $\rho(m,n)$. Let $\nabla\varphi(m,n) = [\varphi_m(m,n), \varphi_n(m,n)]$ be the true gradient field of the unwrapped phase. The unwrapped phase $\varphi(m,n)$ is the solution to the L_2 norm minimization

$$\mathcal{E}[\varphi(m,n)] = ||\varphi_m(m,n) - \rho_m(m,n)||^2 + ||\varphi_n(m,n) - \rho_n(m,n)||^2. \tag{4.5}$$

Hunt [57] formulated (4.5) as a matrix multiplication. He constructed the matrix $\mathbf{A}$ which acts like a phase difference operator. In this formulation, the phase functions $\varphi(m,n)$, $\rho_m(m,n)$, and $\rho_n(m,n)$ are vectorized into 1D vector. Concretely, let φ be a 1D vectorization of $\varphi(m,n)$ and let $\gamma = [\rho_m\ \rho_n]^T$ be a 1D vector consisting of two stacked 1D vectors ρ_m and ρ_n. The energy minimization equation (4.5) is equivalent to

$$\begin{aligned}
\mathcal{E}(\varphi) &= ||\mathbf{A}\varphi - \gamma||^2 \\
&= (\mathbf{A}\varphi - \gamma)^T(\mathbf{A}\varphi - \gamma) \\
&= (\mathbf{A}\varphi)^T\mathbf{A}\varphi - (\mathbf{A}\varphi)^T\gamma - \gamma^T\mathbf{A}\varphi + \gamma^T\gamma \\
&= \varphi^T\mathbf{A}^T\mathbf{A}\varphi - 2\varphi^T\mathbf{A}^T\gamma + \gamma^T\gamma.
\end{aligned} \tag{4.6}$$

The error in (4.6) is a quadratic function of φ. The optimal solution for φ is computed by equating the derivative of $\mathcal{E}[\varphi(m,n)]$ to zero according to

$$\frac{\partial \mathcal{E}}{\partial \varphi} = 2\mathbf{A}^T\mathbf{A}\phi - 2\mathbf{A}^T\gamma = \mathbf{0},$$
$$\text{or equivalently, } \mathbf{A}^T\mathbf{A}\varphi = \mathbf{A}^T\gamma. \tag{4.7}$$

Notice that (4.7) is the well-known least-squares solution of an overdetermined linear system. One can solve (4.7) for φ using matrix inversion provided that the matrix $\mathbf{A}^T\mathbf{A}$ is not singular. In the phase unwrapping problem, the matrix $\mathbf{A}$ computes the discrete approximation to the gradient. As a result, $\mathbf{A}^T\mathbf{A}\varphi$ can be interpreted as the Laplacian of φ and $\mathbf{A}^T\rho$ represents the derivative of the measured gradient ρ, which is the Laplacian of measured phase ρ. With this representation, equation (4.7) is the discretization of the classical Poisson equation which can be solved exactly using the fast discrete cosine transform (DCT) [118]. A detailed derivation of this is given in Appendix B. Let Φ be the 2D DCT transform of φ and Γ be the 2D DCT transform of γ. We have the in DCT domain relationship [118].

$$\Phi(i,j) = \frac{\Gamma(i,j)}{2\cos\left(\frac{\pi}{M}i\right) + 2\cos\left(\frac{\pi}{N}j\right) - 4}, \tag{4.8}$$

where i and j are indexes of the 2D grid. The least-squares unwrapped phase $\varphi_{\text{LS}}(m,n)$ is obtained by taking the inverse 2D discrete cosine transform as

$$\varphi_{\text{LS}}(m,n) = \text{IDCT}\{\Phi(i,j)\}. \tag{4.9}$$

Notice that $\Phi(0,0)$ in (4.8) is not defined because the denominator is zero. In practice, I set $\Phi(0,0) = 0$ which results in a zero mean unwrapped phase function. I emphasize that $\varphi_{\text{LS}}(m,n)$ is not necessarily equal to the true phase $\varphi(m,n)$; rather, because it is the least-squares approximation of $\varphi(m,n)$. As the unwrapped phase is obtained by integrating the gradient, the computed solution can differ from the true phase by a global constant in the best case.

4.4 Perfect Reconstruction FM Algorithm

4.4.1 Arguments for the Least-squares Phase Unwrapping

I adopted the 2D least-squares phase unwrapping algorithm for the AM-FM perfect reconstruction model for several reasons. The unwrapped phase produced by the least-squares phase unwrapping algorithm is usually smooth and agrees with the smoothly varying assumption of the AM-FM image model. In addition, as the measurement noise is inherently modeled in the least-squares formulation, the algorithm is robust to phase discrepancies. As a result, the least-squares approach does not need to explicitly determine the phase discontinuities like other approaches such as path integration discussed in Section 4.2.1. Moreover, the least-squares formulation provides an efficient and stable algorithm to compute the unwrapped phase. Finally, additional constraints can be enforced on the reconstructed phase because regularization terms can be conveniently integrated into the least-squares formulation.

4.4.2 Enforcing Phase Congruence

Let $\mathbf{n} = [n \ m] \in \mathbb{Z}^2$. As noted in Section 4.3, the least-squares phase $\varphi_{\text{LS}}(\mathbf{n})$ computed from (4.9) is not guaranteed to be equal to the true phase $\varphi(\mathbf{n})$, i.e., $\varphi_{\text{LS}}(\mathbf{n}) \neq \varphi(\mathbf{n})$. As a result, the original image can not be recovered exactly because $\cos[\varphi_{\text{LS}}(\mathbf{x})] \neq \cos[\varphi(\mathbf{x})]$ in general.

In the best case, the unwrapped phase differs from the true phase by a global constant c. The constant c can be estimated by minimizing the energy between the wrapped phase and the unwrapped phase. In other words, c is obtained by minimizing the following norm

$$\mathcal{E}(c) = || \cos(\mathcal{W}\{\varphi(m,n)\}) - \cos[\varphi_{\text{LS}}(m,n) + c]||^2. \tag{4.10}$$

Even in cases where the unwrapped phase differs from a true phase by a function $d(\mathbf{n})$, we can still use this least-squares approach to improve the accuracy of the unwrapped phase. For instance, Pritt [100] computed (4.10) with different c over the range of $[0, 2\pi]$ and chose the c with the lowest error. The constant c can also be found using an iterative approach such as the gradient descent method. For instance, in the phase unwrapping example in Fig. 4.1, I found that $c = 0.9105$ using the gradient descent algorithm with step size 0.3 and 1000 iteration. The mean squared error is improved from 0.9773 to 0.9626.

In order to obtain FM perfect reconstruction, Sivley and Havlicek [108] enforced congruence on the least-squares unwrapped phase. For perfect reconstruction, they required that $\cos[\varphi(\mathbf{n})] = \cos[\mathcal{W}\{\varphi(\mathbf{n})\}]$. Therefore, they introduced an auxiliary function $b(\mathbf{n}) \in \mathbb{Z}$ and computed $b(\mathbf{n})$ as

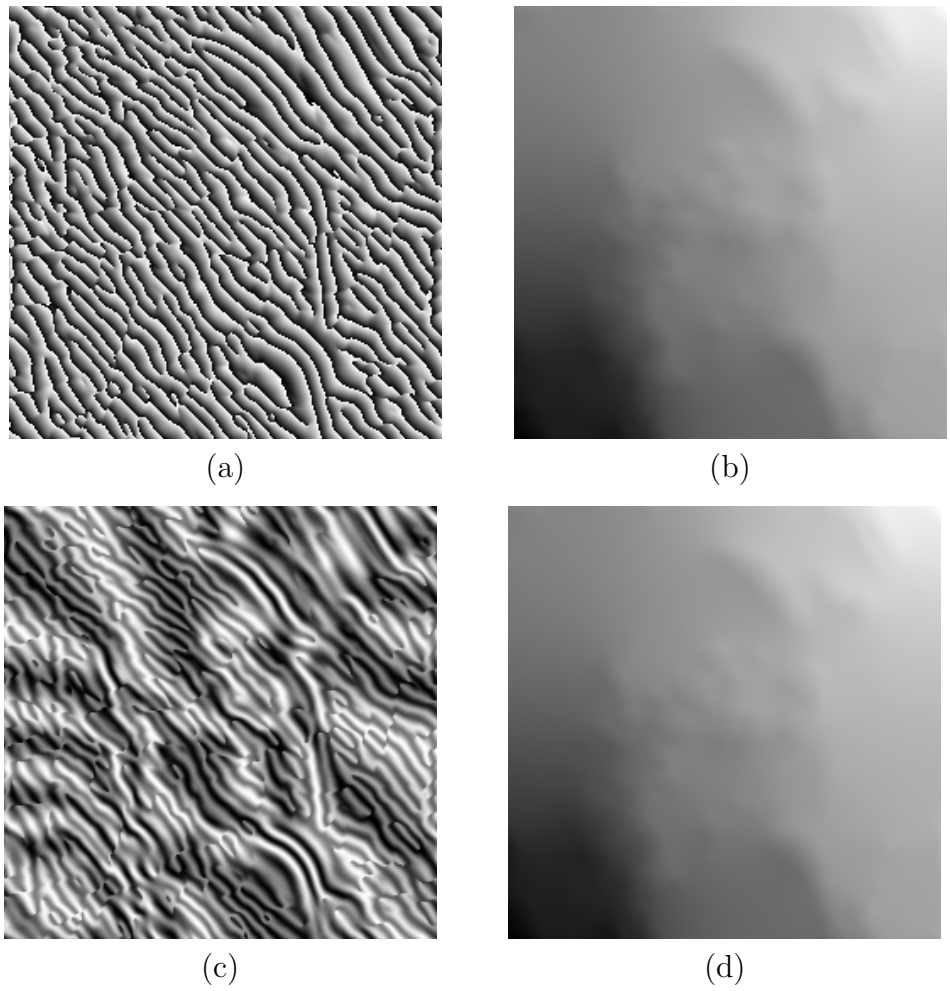
$$b(\mathbf{n}) = \left\lfloor \frac{\beta\varphi_{\mathrm{LS}}(\mathbf{n}) - \mathcal{W}\{\varphi(\mathbf{n})\}}{2\pi} \right\rfloor, \tag{4.11}$$

where "$\lfloor\ \rfloor$" denotes the floor function and β is a predefined positive real constant. The constant β was set to 300 according to [108]. The constant β alleviates discontinuous jumps in the final unwrapped phase function. The final unwrapped phase $\varphi(\mathbf{x})$ is given by

$$\varphi(\mathbf{n}) = \mathcal{W}\{\varphi(\mathbf{n})\} + 2\pi b(\mathbf{n}). \tag{4.12}$$

I illustrate the importance of enforcing the phase congruence in Fig. 4.3. Fig. 4.3(a) shows the wrapped phase of a locally narrowband component. Fig. 4.3(b) shows the unwrapped phase $\varphi(\mathbf{n})$. The unwrapped phase is smoother and does not contain discontinuities compared to the wrapped phase in Fig. 4.3(a). The error between the least-squares phase and the wrapped phase after mapping through the cosine function is shown in Fig. 4.3(c). The congruency term $b(\mathbf{n})$ is shown in

Fig. 4.3(d). The error between the least-squares phase and the final unwrapped phase $\varphi(\mathbf{n})$ are illustrated with 3D plots in Fig. 4.4(a). This error which is shown in Fig. 4.4(b) is effectively zero after the congruency term is enforced. Similar example of phase congruency enforcement is shown in Fig. 4.5 and 4.6

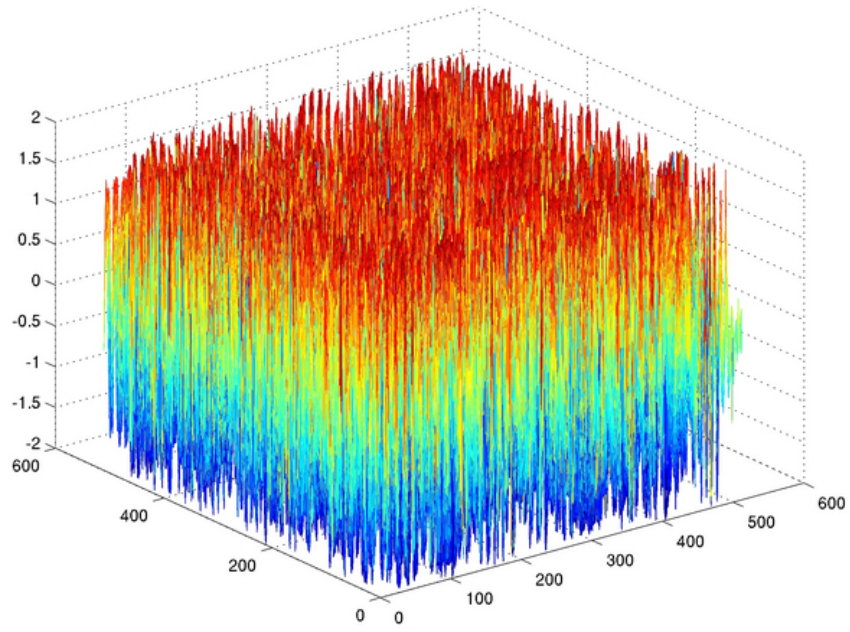


(a) (b)

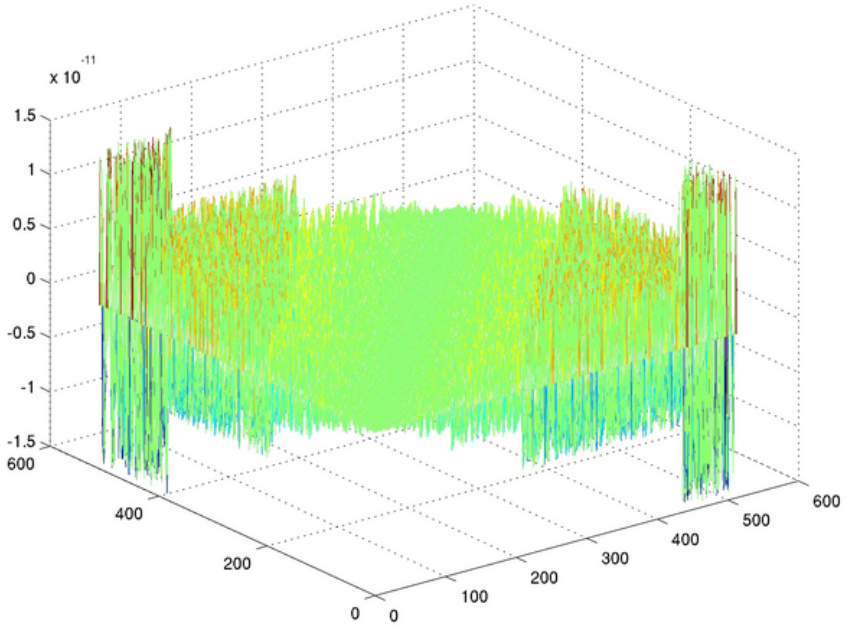
(c) (d)

Figure 4.3: 2D Phase unwrapping of one component of lena, lena_3_2. (a) Wrapped phase. (b) Unwrapped phase. (c) $\cos[\varphi_{\mathrm{LS}}(\mathbf{n})] - \cos[\varphi(\mathbf{n})]$. (d) Congruence term $b(\mathbf{n})$ in (4.12).

Quantitative illustrations of the MSE between the unwrapped phase before and after the phase congruency is enforced are shown in Table. 4.1 and 4.3. We

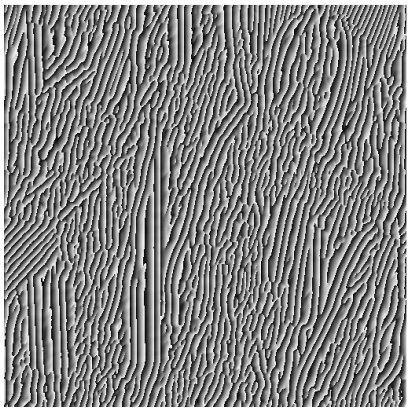


(a)

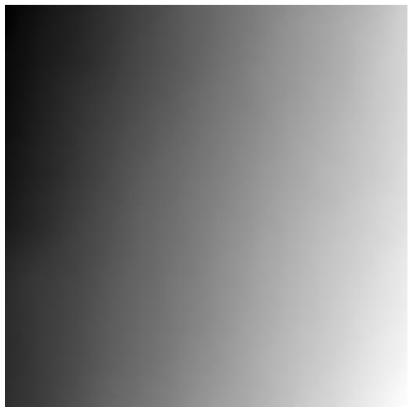


(b)

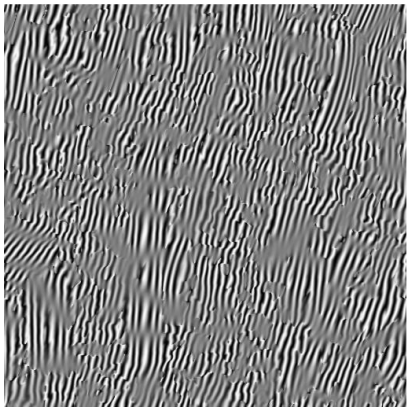
Figure 4.4: 3D error plot of the 2D Phase unwrapping for barbara_3_2. (a) $\cos[\varphi_{\mathrm{LS}}(\mathbf{n})] - \cos[\mathcal{W}\{\varphi(\mathbf{n})\}]$. (b) $\cos[\varphi(\mathbf{n})] - \cos[\mathcal{W}\{\varphi(\mathbf{n})\}]$.



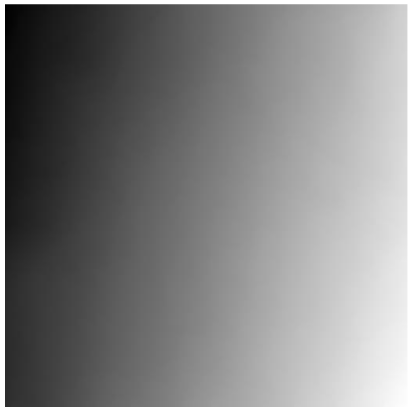
(a)



(b)

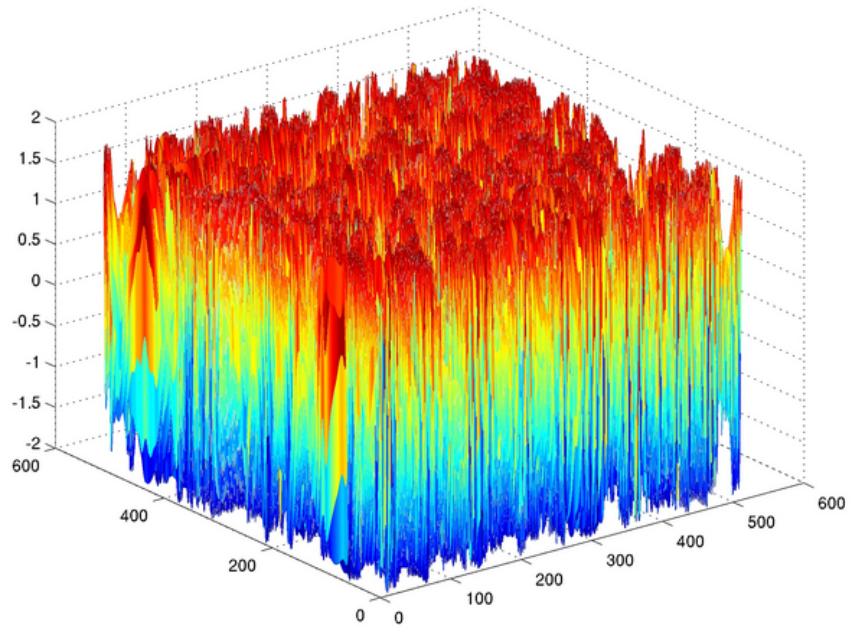


(c)

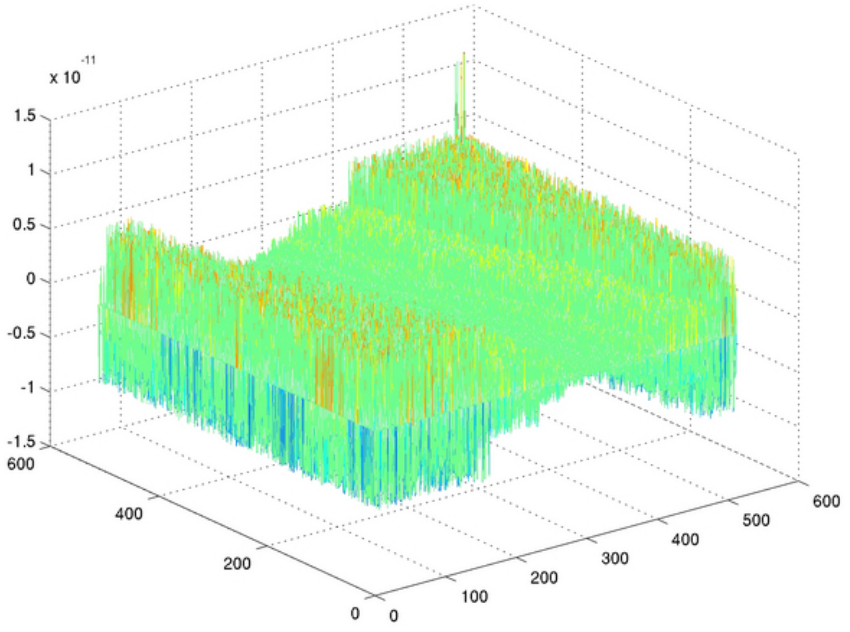


(d)

Figure 4.5: 2D Phase unwrapping of one component of barbara, babara_3_4. (a) Wrapped phase. (b) Unwrapped phase. (c) $\cos[\varphi_{\mathrm{LS}}(\mathbf{n})] - \cos[\varphi(\mathbf{n})]$. (d) Congruence term $b(\mathbf{n})$ in (4.12).



(a)



(b)

Figure 4.6: 3D error plot of the 2D Phase unwrapping for barbara_3_4. (a) $\cos[\varphi_{\mathrm{LS}}(\mathbf{n})] - \cos[\mathcal{W}\{\varphi(\mathbf{n})\}]$. (b) $\cos[\varphi(\mathbf{n})] - \cos[\mathcal{W}\{\varphi(\mathbf{n})\}]$.

can see that the MSE of 12 narrowband components effectively go zero after the phase congruency terms are added to the least-squares phase.

Table 4.1: Mean squared error (MSE) comparison between the true phase and the least-squares phase after and before phase congruency is added to the least-squares phase for the test image Barbara.

	after phase congruency	before phase congruency
barbara_2_2	1.393365×10^{-23}	0.979680
barbara_2_3	1.273848×10^{-23}	0.998202
barbara_2_4	1.308133×10^{-23}	1.001034
barbara_2_5	1.416999×10^{-23}	1.007397
barbara_3_2	3.394855×10^{-24}	0.977283
barbara_3_3	3.049092×10^{-24}	0.970878
barbara_3_4	3.016521×10^{-24}	0.977438
barbara_3_5	3.109758×10^{-24}	0.991340
barbara_4_2	7.302929×10^{-25}	0.973036
barbara_4_3	7.074440×10^{-25}	0.944815
barbara_4_4	7.274310×10^{-25}	1.065723
barbara_4_5	7.342723×10^{-25}	1.051791

4.4.3 Spline-based Perfect Reconstruction FM

Once the unwrapped phase $\varphi(\mathbf{n})$ is obtained, the FM is computed as the gradient of $\varphi(\mathbf{n})$. For a perfect reconstruction representation, the phase $\varphi(\mathbf{n})$ must be recovered from the computed FM without errors. Sivley and Havlicek [110] provided the first perfect reconstruction FM transform. They used the spline-based framework proposed by Unser, Aldroubi, and Eden [121, 122]. They fitted $\varphi(\mathbf{n})$ with a tensor product cubic spline model to create a continuous surface $\varphi_c(\mathbf{x})$. Specifically, they computed the cubic-spline coefficients from the discrete samples of $\varphi(\mathbf{n})$. The forward cubic spline transfer function is given by [122] by

$$[B^3(z)]^{-1} = \frac{6}{z + 4 + z^{-1}}. \tag{4.13}$$

Table 4.2: Mean squared error (MSE) comparison between the true phase and the least-squares phase after and before phase congruency is added to the least-squares phase for the test image Lena.

	after phase congruency	before phase congruency
lena_2_2	8.306846×10^{-26}	0.993641
lena_2_3	2.033533×10^{-25}	1.012747
lena_2_4	2.362168×10^{-25}	0.997782
lena_2_5	2.183022×10^{-25}	0.992178
lena_3_2	5.673893×10^{-26}	0.997217
lena_3_3	1.570788×10^{-25}	0.978311
lena_3_4	1.769817×10^{-25}	1.005338
lena_3_5	1.166994×10^{-25}	1.035526
lena_4_2	2.951476×10^{-26}	1.036047
lena_4_3	6.999355×10^{-26}	1.101886
lena_4_4	8.165159×10^{-26}	0.863878
lena_4_5	6.236149×10^{-26}	0.994436

Table 4.3: The mean squared error (MSE) comparison of the least-squares phase and the unwrapped phase with respect to the true phase for test image Lena.

The signal $\varphi(\mathbf{n})$ is recovered without errors by the indirect cubic spline transfer function

$$B^3(z) = \frac{z + 4 + z^{-1}}{6}. \tag{4.14}$$

For a 2D signal like $\varphi(\mathbf{n})$, the cubic spline filters $[B^3(z)]^{-1}$ and $B^3(z)$ are applied successively along rows and columns of the image.

Since the cubic spline interpolants were used in [109], the first order derivative of these splines is quadratic splines. In order to map the quadratic spline representation to the image intensity representation, one must apply the indirect quadratic spline with transfer function $C^2(z)$ is given by

$$C^2(z) = \frac{z + 1}{2}. \tag{4.15}$$

The first order derivative operator and the indirect quadratic spline $C^2(z)$ can be

grouped into one filter

$$D(z) = (1 - z^{-1})\frac{z+1}{2} = \frac{z - z^{-1}}{2}. \tag{4.16}$$

Let $g_r(\mathbf{n})$ be cubic spline representation of the phase function $\varphi(\mathbf{n})$ along the rows. Then $g_r(\mathbf{n})$ is obtained as

$$g_r(\mathbf{n}) = \varphi(\mathbf{n}) * [b^3(n)]^{-1}, \tag{4.17}$$

where $[b^3(n)]^{-1}$ is the time domain representation of $[B^3(z)]^{-1}$. Similarly, let $g_c(\mathbf{n})$ be the cubic spline representation along the column of the phase image $\varphi(\mathbf{n})$, $g_c(\mathbf{n}) = \varphi(\mathbf{n}) * [b^3(n)]^{-1}$. Let $U(\mathbf{n})$ and $V(\mathbf{n})$ be the horizontal and the vertical components of the FM vector. They are computed as

$$U(\mathbf{n}) = \frac{g_c(m, n+1) - g_c(m, n-1)}{2}, \tag{4.18}$$

$$V(\mathbf{n}) = \frac{g_r(m+1, n) - g_r(m-1, n)}{2}. \tag{4.19}$$

Sivley and Havlicek [110] showed that the original phase function $\varphi(\mathbf{n})$ can be perfectly reconstructed from the FM functions $U(\mathbf{n})$ and $V(\mathbf{n})$. However, the reconstruction algorithm required prior knowledge of four points of the original phase function. Particularly, they assumed that $\varphi(0,0)$, $\varphi(0,1)$, $\varphi(1,0)$, and $\varphi(1,1)$ are known. First, the algorithm computes $g_r(m,n)$ using (4.19) for the first two columns

$$g_r(m+1, n) = 2V(m, n) + g_r(m-1, n), \tag{4.20}$$

where $g_r(i,j) = \varphi(i,j)$ for $i, j \in [0,1]$. Let $g_c(i,j) = g_r(i,j)$, for $i \in [0, M-1]$ and $j \in [0,1]$. Then $g_c(\mathbf{n})$ is computed using (4.19) along the columns as

$$g_c(m, n+1) = 2U(m, n) + g_c(m, n-1), \tag{4.21}$$

where $g_c(i,j) = g_r(i,j)$, for $i \in [0, M-1]$ and $j \in [0,1]$. The original phase $\varphi(\mathbf{n})$ is then obtained by performing the indirect cubic spline transform

$$\varphi(\mathbf{n}) = g_r(\mathbf{n}) * b^3(n), \tag{4.22}$$

where $b^3(n)$ is the time domain representation of $B^3(z)$.

Notice that one could arrive at the same answer for $\varphi(\mathbf{n})$ if the reconstruction algorithm was started by integrating along the columns and then integrating along the rows. However, the algorithm still requires four points on the top left corner of the original phase function $\varphi(\mathbf{n})$ in that case.

4.4.4 Least-squares FM Perfect Reconstruction

While the spline-based algorithm in Section 4.4.3 provides perfect reconstruction of the phase from the FM, it is not suitable for FM signal processing applications. The first step of the reconstruction algorithm only operates on the first two columns. Therefore, any signal processing changes of $V(\mathbf{n})$ in columns other than these first two will not be reflected in the reconstructed phase. Similarly, changes in the first two columns of $V(\mathbf{n})$ will results in changes even though $V(\mathbf{n})$ is unmodified everywhere else. Second, the algorithm requires the knowledge of four points on the top left corner of the original phase function. A small change in any of these four points results in global changes to the reconstructed phase. Therefore, the algorithm does not give predictable results if the FM is altered by signal processing. This limitation prohibits the implementation of non-trivial filtering operations applying directly on the FM functions.

Nguyen, Campell, and Havlicek proposed a new FM perfect reconstruction algorithm in [88]. The FM reconstruction followed the same formulation as the

least-squares phase unwrapping in Section 4.3. Let $U(\mathbf{n})$ and $V(\mathbf{n})$ be the horizontal and vertical components of the FM function. They are computed from the unwrapped phase $\varphi(\mathbf{n})$ using the phase difference technique

$$U(m,n) = \varphi(m,n) - \varphi(m,n-1), \tag{4.23}$$

$$V(m,n) = \varphi(m,n) - \varphi(m-1,n), \tag{4.24}$$

where $\mathbf{n} = [m\ n]^T$. Recall that the discrete phase differences where also used in the least-squares optimization setup in (4.5).

In the FM reconstruction problem, I aimed to reconstruct the phase function $\varphi(\mathbf{n})$ from the FM functions $U(\mathbf{n})$ and $V(\mathbf{n})$. Let $\widetilde{\varphi}(\mathbf{n})$ be the reconstructed phase. Similar to the least-squares approach in Section 4.3, I set up the least-squares optimization problem as

$$\mathcal{E}[\widetilde{\varphi}(m,n)] = ||\widetilde{\varphi}_m(m,n) - V(m,n)||^2 + ||\widetilde{\varphi}_n(m,n) - U(m,n)||^2, \tag{4.25}$$

where $\widetilde{\varphi}_m(m,n)$ and $\widetilde{\varphi}_n(m,n)$ are the vertical and horizontal approximation of the derivatives. In particular, the FM components are computed using the phase differences, i.e., $\widetilde{\varphi}_m(m,n) = \widetilde{\varphi}(m,n) - \widetilde{\varphi}(m-1,n)$ and $\widetilde{\varphi}_n(m,n) = \widetilde{\varphi}(m,n) - \widetilde{\varphi}(m,n-1)$.

From (4.23)– (4.24), and (4.25), the solution of (4.25), if found, differs from the original phase $\varphi(\mathbf{n})$ by a global constant τ:

$$\widetilde{\varphi}(\mathbf{n}) = \varphi(\mathbf{n}) + \tau. \tag{4.26}$$

Therefore, the original phase $\varphi(\mathbf{n})$ can be perfectly reconstructed if τ is known a priori. I chose τ to be the top left corner pixel of the unwrapped phase, e.g., $\varphi(0,0)$. The MSE of the reconstructed phase $\tilde{\varphi}(\mathbf{n})$ with respect to the original

phase $\varphi(\mathbf{n})$ is given in Table 4.4 for 12 locally narrowband components of the well-known mandrill image. The algorithm is able to reconstruct the original phase from the FM functions $U(\mathbf{n})$ and $V(\mathbf{n})$ without errors.

Table 4.4: Mean squared error (MSE) comparison of the least-squares phase and the unwrapped phase with respect to the true phase for the test image mandrill.

	$(MN)^{-1}\|\|\tilde{\varphi}(\mathbf{n}) - \varphi(\mathbf{n})\|\|^2$
mandril_2_2	1.438606×10^{-14}
mandril_2_3	2.574802×10^{-14}
mandril_2_4	5.446122×10^{-14}
mandril_2_5	7.288013×10^{-14}
mandril_3_2	3.886939×10^{-15}
mandril_3_3	7.553478×10^{-15}
mandril_3_4	1.462437×10^{-14}
mandril_3_5	1.944426×10^{-14}
mandril_4_2	9.360980×10^{-16}
mandril_4_3	1.825935×10^{-15}
mandril_4_4	3.595069×10^{-15}
mandril_4_5	4.682343×10^{-15}

I illustrate FM reconstruction results in Fig. 4.7. The FM functions $U(\mathbf{n})$ and $V(\mathbf{n})$ are shown in Fig. 4.7(a) and Fig. 4.7(b), respectively. Fig. 4.7(c) depicts the least-squares reconstructed phase $\varphi_{\text{reconLS}}(\mathbf{n})$ obtained by solving (4.25). The final phase is shown in Fig. 4.7(d). The least-squares phase and the final phase differ by a constant τ. I show this difference as a mesh plot in Fig. 4.8(a). The error between the reconstructed phase and the original phase is also illustrated in Fig. 4.8(b).

The least-squares perfect reconstruction algorithm overcomes the limitations of the spline-based approach in Section 4.4.3. The algorithm requires a priori knowledge of one sample rather than four samples. In addition, the changes in $U(\mathbf{n})$ and $V(\mathbf{n})$ are reflected in the reconstructed phase which allows us to define

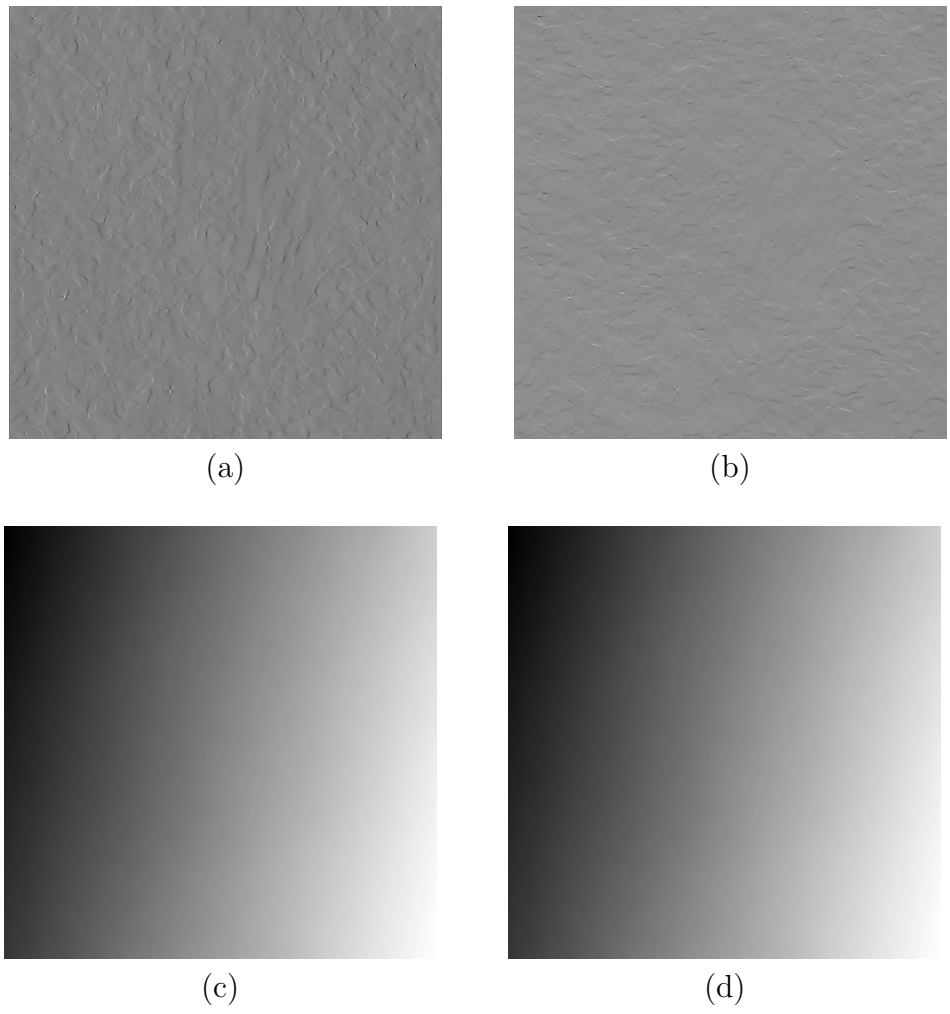
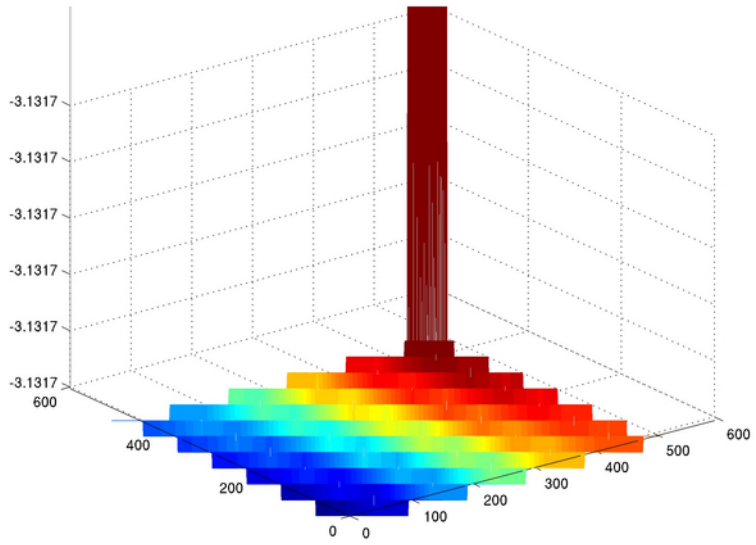
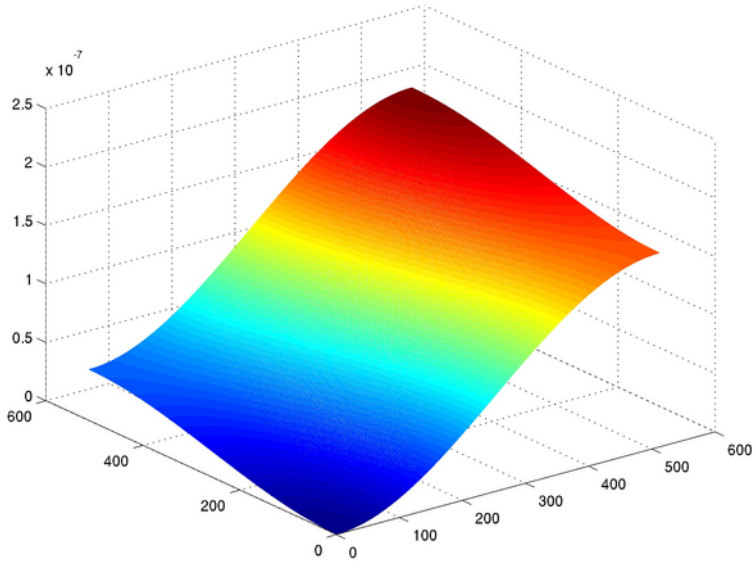


Figure 4.7: 2D phase reconstruction from FM functions. (a) $U(\mathbf{n})$. (b) $V(\mathbf{n})$. (c) Least-squares phase $\varphi_{\text{reconLS}}(\mathbf{n})$. (d) Reconstructed phase $\tilde{\varphi}(\mathbf{n})$.



(a)



(b)

Figure 4.8: 2D least-squares phase reconstruction. (a) Offset constant $\tau \approx -3.1317$ in (4.26). (b) $|\varphi(\mathbf{n}) - \tilde{\varphi}(\mathbf{n})| \in [0,\, 2.5 \times 10^{-7}]$.

filters acting on these FM components.

4.5 The Single Component PR AM-FM Transform

Once the FM perfect reconstruction algorithm was realized, I developed the perfect reconstruction (PR) model for the single component AM-FM transform. Given a real discrete input image $f(\mathbf{n})$, I constructed the imaginary image $q(\mathbf{n})$ using the pHT technique in Section 3.2.1. I then created the complex image as $z(\mathbf{n}) = f(\mathbf{n}) + jq(\mathbf{n})$. The AM and FM functions are computed as

$$a(\mathbf{n}) = \sqrt{f^2(\mathbf{n}) + q^2(\mathbf{n})}, \tag{4.27}$$

$$\nabla\varphi(\mathbf{n}) = \text{Im}\left[\frac{\nabla z(\mathbf{n})}{z(\mathbf{n})}\right]. \tag{4.28}$$

Since the derivative operator is not defined for discrete signals, $\nabla z(\mathbf{n})$ in (4.28) must be computed using an approximate discrete implementation. Sivley and Havlicek [108,110] used the spline-based approach to compute discrete derivatives. The spline-based approach permits an invertible FM transform. Farid and Simoncelli [33] designed separable kernels to perform differentiation for multidimensional discrete signals. These filters are of the finite impulse response (FIR) type. The accuracy of the approximation depends on the filter length. The filter coefficients are computed using a least-squares optimization technique. For practical applications, these derivative FIR filters have five or seven taps. In the proposed transform, a five tap FIR filter is used.

While the AM and FM functions have been computed numerically, we still can not reconstruct the original image $f(\mathbf{n})$ from these modulating functions. Because of the wrapped phase problem discussed in Section 4.1, one must perform phase unwrapping in order to obtain a perfect reconstruction AM-FM model. Let

$U(\mathbf{n})$ and $V(\mathbf{n})$ be the horizontal and vertical components of the FM function, i.e., $\nabla\varphi(\mathbf{n}) = [U(\mathbf{n})\ V(\mathbf{n})]^T$. These FM functions are then used as estimates for the least-squares phase unwrapping process in Section 4.3. Specifically, $U(\mathbf{n})$ plays the role of $\rho_n(m,n)$ and $V(\mathbf{n})$ plays the role of $\rho_m(m,n)$. The unwrapped phase $\varphi(\mathbf{n})$ is solved by (4.5). The FM functions are then computed from the unwrapped phase $\varphi(\mathbf{n})$ according to (4.23) and (4.24). The triplet $a(\mathbf{n})$, $U(\mathbf{n})$, $V(\mathbf{n})$ constitutes a perfect reconstruction AM-FM image representation for the image $f(\mathbf{n})$.

The original image $f(\mathbf{n})$ is reconstructed from the three components $a(\mathbf{n})$, $U(\mathbf{n})$, $V(\mathbf{n})$ without errors. I showed in Section 4.4.4 that the phase function $\varphi(\mathbf{n})$ can be perfectly reconstructed from the FM functions $U(\mathbf{n})$ and $V(\mathbf{n})$. While both the spline-based approach in Section 4.4.3 and the least-squares approach in Section 4.4.4 allow perfect reconstruction of the phase signal from the FM, I adopted the least-squares FM reconstruction approach because the least-squares reconstructed phase function is more robust to changes in the FM than that obtained from the spline-based approach. Once $\varphi(\mathbf{n})$ is recovered, the original image is given by

$$f(\mathbf{n}) = a(\mathbf{n})\cos[\varphi(\mathbf{n})]. \tag{4.29}$$

Algorithm 1 The analysis AM-FM transform

Let $f(\mathbf{n})$ be the discrete image.
Step 1: Construct the complex image $z(\mathbf{n})$:
$q(\mathbf{n}) = \mathcal{H}\{f(\mathbf{n})\}$, defined in Section 3.2.1.
$z(\mathbf{n}) = f(\mathbf{n}) + jq(\mathbf{n})$.
Step 2: Compute the AM using (4.27) and the FM using (4.28)
Step 3: Find the phase function $\varphi(\mathbf{n})$ by solving (4.25)
Step 4: Compute the FM functions using (4.23) and (4.24).

Steps for computing the analysis and synthesis AM-FM transform are given in Algorithm 1 and Algorithm 2, respectively.

Algorithm 2 The synthesis AM-FM transform

Require: $\varphi(0,0)$ known a priori.

Let $U(\mathbf{n})$ and $V(\mathbf{n})$ be the FM functions.

Step 1: Reconstruct the phase function $\varphi(\mathbf{n})$ by solving (4.25).

Step 2: Reconstruct the original image using (4.29).

I applied the perfect reconstruction AM-FM transform to two single component images: BentChirp and Chirp. The transform and reconstruction the BentChirp image are shown in Fig. 4.9, and those of the Chirp image are shown in Fig. 4.11. In addition, 3D plots of the computed AM and FM functions are depicted in Fig. 4.10 and Fig. 4.12. I quantified the reconstruction error by computing the mean squared error (MSE) and peak signal to noise ratio (PNSR)

$$\mathrm{MSE}(f,g) = \frac{\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}[f(m,n)-g(m,n)]^2}{MN}, \tag{4.30}$$

$$\mathrm{PSNR}(f,g) = 10\log_{10}\left(\frac{\max[f]}{\mathrm{MSE}(f,g)}\right), \tag{4.31}$$

where $g(m,n)$ is the reconstructed image. The results are given in Table 4.5. The low MSE (8.523966×10^{-18}) and high PSNR (170.71 dB) results confirm that the proposed algorithms provide a perfect reconstruction AM-FM transform.

Table 4.5: Reconstruction error of the one component AM-FM transform.

	MSE	PSNR (dB)
BentChirp	8.523966×10^{-18}	170.71
Chirp	1.912063×10^{-17}	167.18

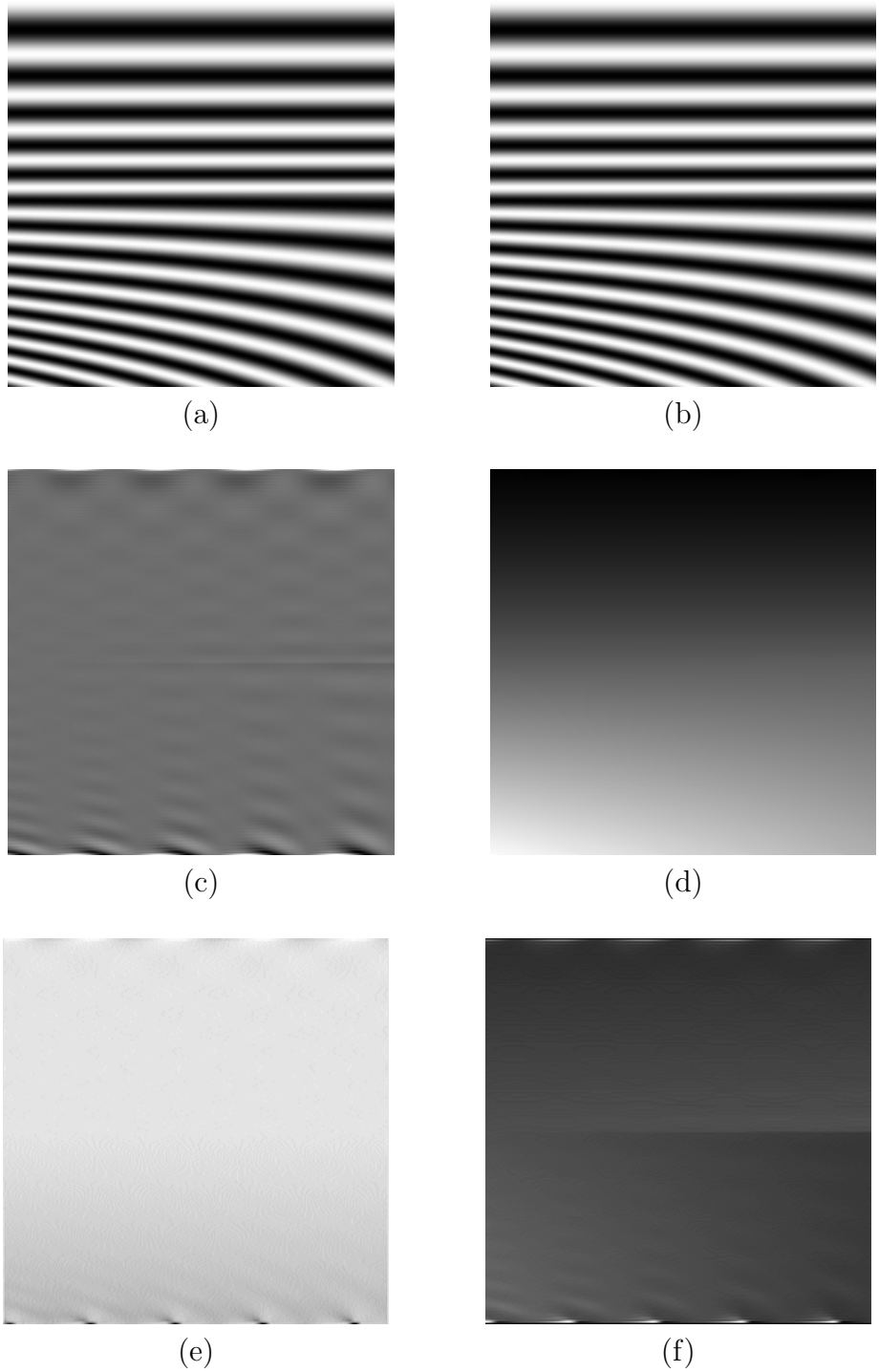
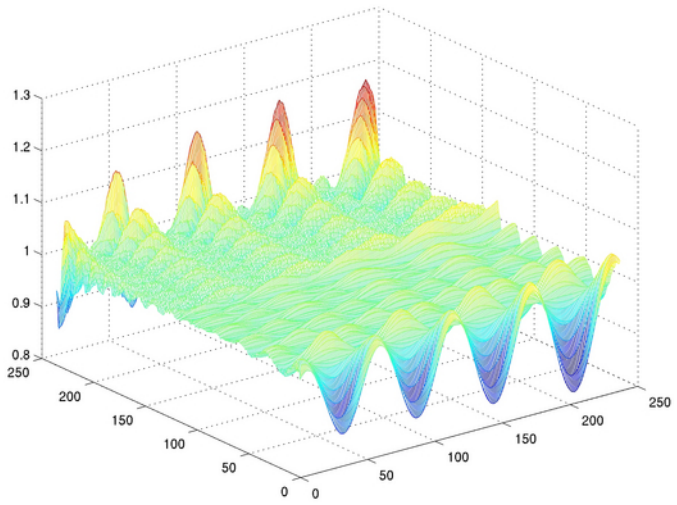
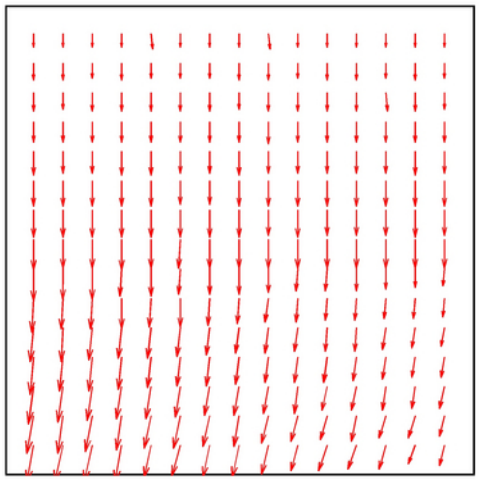


Figure 4.9: Single component perfect reconstruction AM-FM transform of the BentChirp image.. (a) Original bentChirp image. (b) Reconstructed bentChirp image. (c) Computed AM function. (d) Computed phase function $\varphi(\mathbf{n})$. (e) Horizontal component of $\varphi(\mathbf{n})$. (f) Vertical component of $\varphi(\mathbf{n})$.

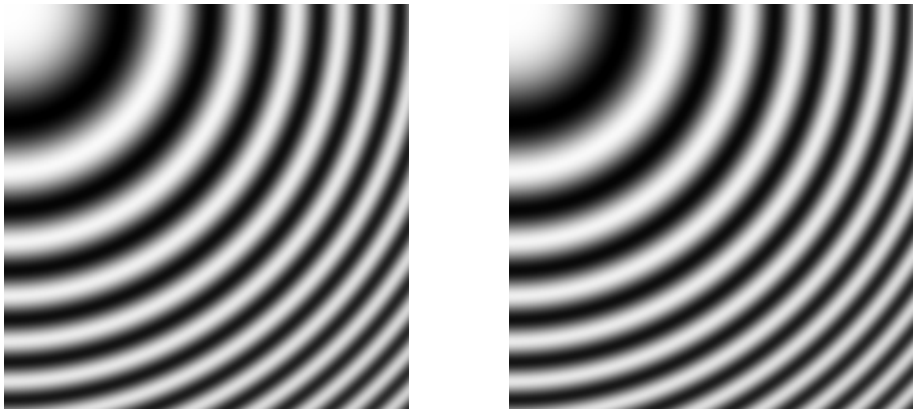


(a)



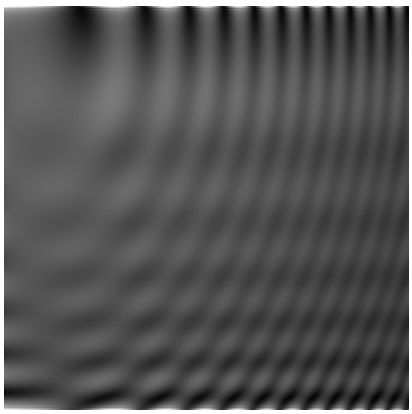
(b)

Figure 4.10: Single component perfect reconstruction AM-FM transform of the BentChirp image. (a) Computed AM. (b) Computed FM field.



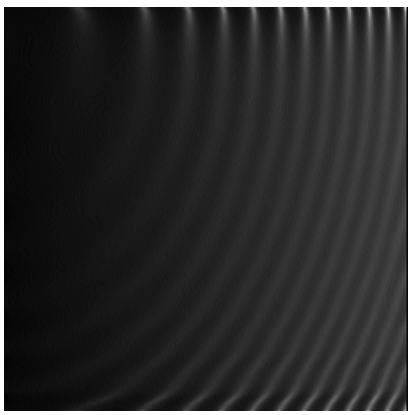
(a)

(b)

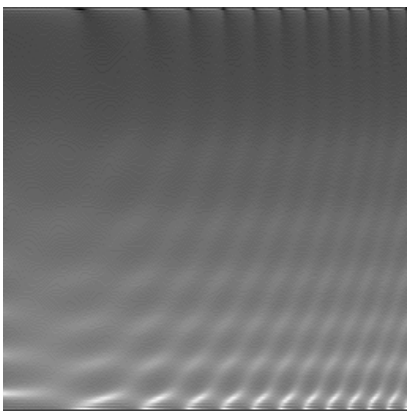


(c)

(d)

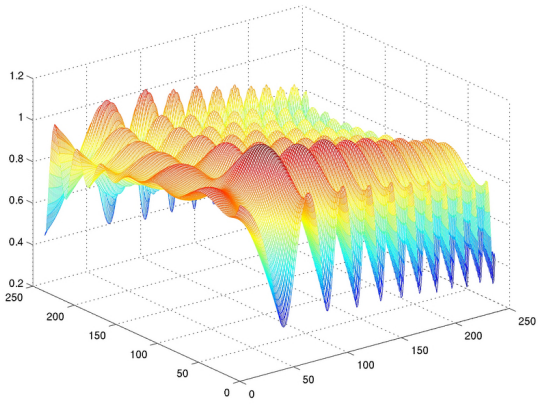


(c)

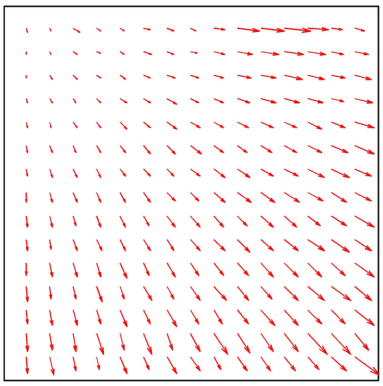


(f)

Figure 4.11: Single component perfect reconstruction AM-FM transform of the Chirp image. (a) Original chirp image. (b) Reconstructed chirp image. (c) Computed AM function. (d) Computed phase function $\varphi(\mathbf{n})$. (e) Horizontal component of $\varphi(\mathbf{n})$. (f) Vertical component of $\varphi(\mathbf{n})$.



(a)



(b)

Figure 4.12: Single component perfect reconstruction AM-FM transform of the Chirp image. (a) Computed AM. (b) Computed FM field.

4.6 Summary

I discussed the importance of 2D phase unwrapping in the context of AM-FM image synthesis applications. I reviewed four major approaches to perform phase unwrapping from the wrapped phase. I then discussed in detail the least-squares phase unwrapping approach which I then used to build the perfect reconstruction AM-FM transform. While both the spline-based approach and least-squares approach yield a perfect reconstruction FM transform, I argued for the least-squares approach because it is more robust to changes in the FM function than the spline-based approach. Finally, I introduced the AM-FM transform for single component images. I provided two algorithms that can compute the AM-FM representation of the image and then reconstruct the original image from the AM and FM functions without errors.

I argued that that the proposed AM-FM transform should be applied to single component images. Most practical images must be decomposed into multiple locally narrowband components before the proposed AM-FM transform can be applied. As a result, an image is represented as a finite sum of multiple single component AM-FM models. I develop the multi-component perfect reconstruction AM-FM transform in the next chapter.

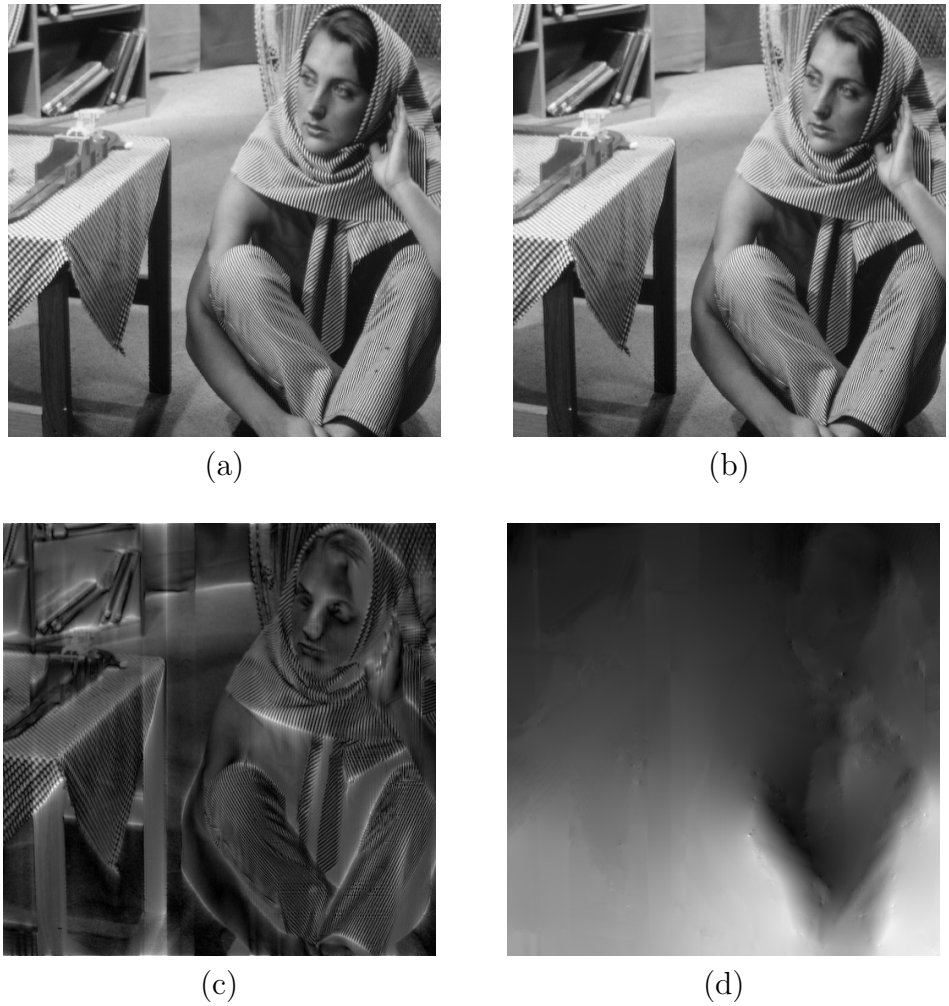
Chapter 5

Multi-component AM-FM Transform

5.1 Motivation for the Multi-component Approach

In Chapter 4, I discussed the single component AM-FM transform where an image is represented by one AM function and one FM function. The AM and FM functions are assumed to be smoothly varying functions. However, for practical images such as Barbara in Fig. 5.1(a), the single component AM-FM model is not an appropriate representation. Mathematically, the single component AM-FM representation of Barbara is still a perfect reconstruction model. Indeed, the MSE and PSNR of the reconstructed image are MSE = 5.952755×10^{-13} and PSNR = 164.44dB. Nevertheless, the computed AM and FM functions do not correspond with human visual perception of the image. The AM function of the single component prefect reconstruction AM-FM transform is given in Fig. 5.1(c); it contains texture information and oscillations. The FM is depicted with a needle diagram in Fig. 5.2. While the single component FM function captures some texture features, in many locations the needles in the FM field fail to indicate the correct texture orientations. In certain regions, the magnitude of the needles are also unstable due to the presence of phase discontinuities in the image. To be meaningful, an AM-FM representation must satisfy two constraints: perfect reconstruction and intuitive interpretation. To achieve these goals, for images like Barbara, a multi-component AM-FM model is required. Specifically, the image Barbara should be represented by K locally narrowband AM-FM components,

where K is a reasonably small integer.



(a) (b)

(c) (d)

Figure 5.1: Single component AM-FM representation of the barbara image. (a) Original barbara image. (b) Reconstructed barbara image. (c) Computed AM function. (d) Computed phase function.

Let $f(\mathbf{x})$ be a multi-component image. I model $f(\mathbf{x})$ with K AM-FM components according to

$$f(\mathbf{x}) = \sum_{k=1}^{K} f_k(\mathbf{x}) = \sum_{k=1}^{K} a_k(\mathbf{x}) \cos[\varphi_k(\mathbf{x})], \tag{5.1}$$

where $f_k(\mathbf{x}) = a_k(\mathbf{x}) \cos[\varphi_k(\mathbf{x})]$ is the single-component AM-FM model of image

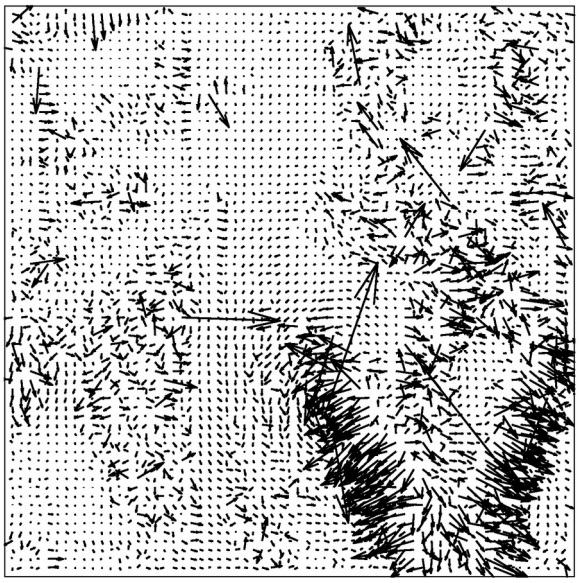


Figure 5.2: Computed FM function for single-component AM-FM transform of Barbara image.

component $f_k(\mathbf{x})$. In order to realize the representation (5.1), I first decompose the image $f(\mathbf{x})$ into multiple locally coherent components $f_k(\mathbf{x})$. The decomposition strategy is given in Section 5.2.

5.2 Perfect reconstruction filterbank

5.2.1 Arguments for the multi-scale multi-orientation filterbank

Biological evidence has been the driving force behind many successful visual processing algorithms. Hubel and Wiesel [55] provided a spatial mapping of the responses of cortical neurons. They found that these responses are sensitive light slits, edges, and bars of different orientations. Campbell and Robson [21] later suggested that the nervous system is also sensitive to selective ranges of spatial frequencies. They conjectured that there might be multiple spatial frequency channels involved in biological vision processing. Together with that of Hubel and Wiesel,

Campbell and Robson's work inspired the interpretation that our visual system functions similar to a spectrum analyzer [12], meaning that the visual passband is composed of multiple frequency and orientation selective channels.

Marcelja [82] provided a mathematical description of the responses of the simple cortical cells. He described the receptive fields in terms of 1D Gabor functions which are localized in both time and frequency. Daugman [28] extended Marcelja's work into 2D and proved that the 2D Gabor filters achieves the low-bound of the uncertainty principle. Jones and Palmer [61] confirmed the validity of the 2D Gabor filter model by comparing the Gabor responses with numerical data acquired from measurements in cat striate cortex.

The visual cortex model where the simple receptive field is modeled by 2D Gabor filters has been used in many important image processing algorithms. Bovik, Clark, and Geisler [17] used a 2D Gabor filterbank to perform image segmentation. Fleet and Jepson [35] computed the instantaneous frequency from the responses of a 2D Gabor filterbank to estimate optical flow in images. Manjunath and Ma [76] applied a Gabor filterbank to extract features for content-based image retrieval. Kovesi [66] computed the phase congruence from responses of Gabor filters. The phase congruency measure acts as a descriptor for image features such as corners and edges.

Besides the Gabor filterbank, multiscale transforms are also products of early human vision research. Marr and Hildreth [80] used the second derivative of a Gaussian to perform multi-scale edge detection. Burt and Adelson [20] decomposed an image into subbands using Gaussian filters. They then used different coding schemes for different subbands to achieve image compression gain. Koenderink [65] suggested that resolution can be a parameter describing images. Re-

cently, wavelet transform has been a popular multiresolution signal analysis tool for many practical application such as image compression and image analysis [73]. Freeman, Adelson, and Simoncelli introduced the steerable pyramid [38,107] where the image is decomposed into multiple orientations and frequency subbands similar to a Gabor filterbank.

To decompose an images into coherent and localized components, I adopt the multi-scale and multi-orientation approach. Specifically, I decompose the input image into multiple components of different scales and orientations using a modified version of the original steerable pyramid [89]. Implementation details of the modified steerable pyramid are given in Section 5.2.4.

5.2.2 Arguments for the Steerable Pyramid

In image analysis, texture orientation provides rich information about the object of interest. Steerable filters was originally proposed by Freeman and Adelson [38] to compute filtered images at arbitrary orientations from a small number of basis elements. Simoncelli, Freeman, and Adelson [107] built the steerable filters into a multi-scale transform to create the steerable pyramid. A fast implementation for the steerable pyramid was later introduced by Simoncelli and Freeman [106]. The steerable pyramid is a multi-resolution signal processing structure which enables a coarse to fine signal analysis. The input image is decomposed into subbands, each with frequency support lying in a finite partition of the original spectrum. The steerable pyramid has been used successfully in many computer vision applications such as image denoising [60], texture analysis [98], and image quality assessment [105].

The steerable pyramid provides both perfect reconstruction and orientation

selectivity. Therefore, it satisfies the requirements of the AM-FM decomposition. I now argue that among the prominent signal transformation and decomposition techniques the steerable pyramid is a suitable choice. Both the steerable pyramid and the Gabor filterbank possess important localization properties for image analysis. Both are multi-scale multi-orientation transforms. An important characteristic that makes them popular in image analysis is that these filters can create translation-invariant and rotation-invariant filterbanks. In addition, these filters allow joint time-frequency localization, even though the steerable filter are not optimal filter in the sense of achieving the lower bound of the uncertainty principal. However, the Gabor filters can not provide perfect reconstruction. This limitation prevents the use of Gabor filters in a perfect reconstruction representation. On the other hand, the steerable pyramid does provide perfect reconstruction. The steerable pyramid is also more effective than the traditional wavelet transform in the context of AM-FM applications. While the traditional wavelet transform allows perfect reconstruction and is complete, i.e., free of redundancy, it does not allow multi-orientation analysis. In addition, the transform is not translation-invariant and rotation-invariant, which are both critical when filtering operations are defined on the subbands. It is important to point out that there are other transforms that provide multi-scale and multi-orientation analysis such as the contourlet transform [29]. The contourlet was not considered because it is originally not translation-invariant.

5.2.3 The Original Steerable Pyramid

Steerable filters can extract important features of images such as texture orientation and edges from the oriented filters [38]. Steerability refers to the ability to

synthesize filters at arbitrary directions as a linear combinations of a small set of basis filters. The steerable filter is based on the multidimensional derivative operator [38]. The steerable pyramid decomposes an image according to the scheme shown in Fig. 5.3. I refer to this decomposition as the *original* steerable pyramid to distinguish it from the modified version that I will develop later in the chapter. The pyramid consists of two phases, namely the analysis phase and the synthesis

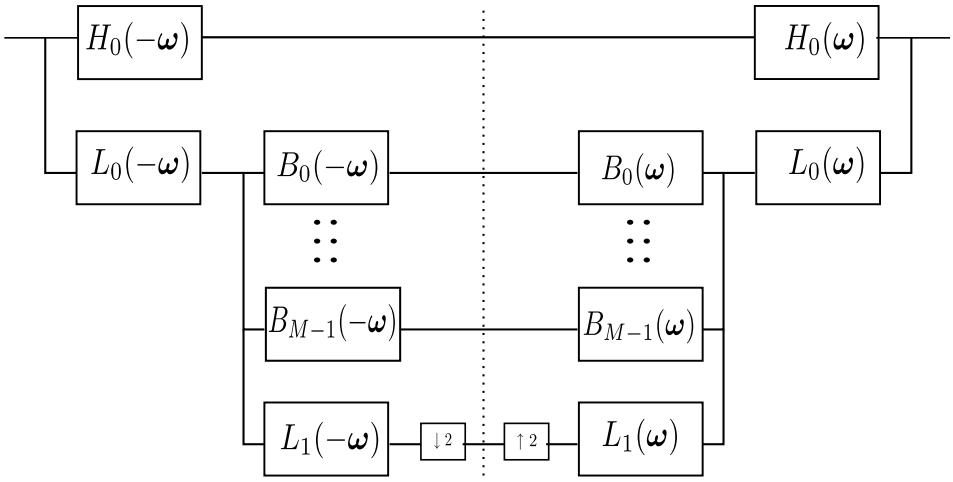


Figure 5.3: Decomposition scheme of the original steerable pyramid [106].

phase. The analysis filterbank is on the left of the vertical dotted line in Fig. 5.3, while the synthesis filterbank is shown on the right of the dotted line.

Let $\widehat{f}(\boldsymbol{\omega})$ be the Fourier spectrum of the input image. In the first decomposition level, the image is decomposed into one highpass component with frequency spectrum $\widehat{f}(\boldsymbol{\omega})H_0(-\boldsymbol{\omega})$, M bandpass components with frequency spectra $\widehat{f}(\boldsymbol{\omega})L_0(-\boldsymbol{\omega})B_1(-\boldsymbol{\omega}), \cdots, \widehat{f}(\boldsymbol{\omega})L_0(-\boldsymbol{\omega})B_M(-\boldsymbol{\omega})$, and one lowpass component with spectrum $\widehat{f}(\boldsymbol{\omega})L_0(-\boldsymbol{\omega})L_1(-\boldsymbol{\omega})$. For all subsequent levels, the decomposition starts over with a downsample image after the application of the lowpass filter $L_1(-\boldsymbol{\omega})$. In other words, the output of the previous lowpass decomposition level

is fed into the next level as the input signal. Hence, the steerable pyramid can be implemented with a recursion approach [106]. In the steerable pyramid decomposition, the highpass filter $H_0(-\boldsymbol{\omega})$ is applied only at the first decomposition level.

In order to yield a perfect reconstruction transform, the filters $H_0(\boldsymbol{\omega})$, $L_0(\boldsymbol{\omega})$, $L_1(\boldsymbol{\omega})$, and $B_k(\boldsymbol{\omega})$ have to meet the following three constraints [106].

1. Aliasing-free subbands:

$$L_1(\boldsymbol{\omega}) = 0, \text{for } |\boldsymbol{\omega}| > \pi/2 \tag{5.2}$$

2. Recursion constraint:

$$|L_1(\boldsymbol{\omega})|^2 + \sum_{k=0}^{M-1} |B_k(\boldsymbol{\omega})|^2 = 1 \tag{5.3}$$

3. Perfect reconstruction:

$$|H_0(\boldsymbol{\omega})|^2 + |L_0(\boldsymbol{\omega})|^2 \left[|L_1(\boldsymbol{\omega})|^2 + \sum_{k=0}^{M-1} |B_k(\boldsymbol{\omega})|^2 \right] = 1 \tag{5.4}$$

The frequency responses of these filters are given in [98]. The same filter shape was also suggested earlier by Castleman, Schulze, and Wu [22]. The lowpass filter $L_0(\boldsymbol{\omega})$ has a frequency spectrum in polar coordinates given by

$$L_0(r,\theta) = \begin{cases} 1 & \text{if } r \leq \frac{\pi}{2}, \\ \cos\left[\frac{\pi}{2}\log_2\left(\frac{2r}{\pi}\right)\right] & \text{if } \frac{\pi}{2} < r < \pi, \\ 0 & \text{if } r \geq \pi. \end{cases} \tag{5.5}$$

The frequency response of the highpass filter $H_0(\boldsymbol{\omega})$ is

$$H_0(r,\theta) = \begin{cases} 0 & \text{if } r \leq \frac{\pi}{2}, \\ \cos\left[\frac{\pi}{2}\log_2\left(\frac{r}{\pi}\right)\right] & \text{if } \frac{\pi}{2} < r < \pi, \\ 1 & \text{if } r \geq \pi. \end{cases} \tag{5.6}$$

Finally, the oriented filter $G_k(\boldsymbol{\omega})$ has a frequency response given in polar form by

$$G_k(r,\theta) = \begin{cases} \alpha_k \left[\cos\left(\theta - \frac{k\pi}{M}\right)\right]^{M-1} & \text{if } \left|\theta - \frac{k\pi}{M}\right| < \frac{\pi}{2}, \\ 0 & \text{otherwise,} \end{cases} \tag{5.7}$$

where $0 \leq k \leq M-1$ indicates orientation index and the constant α is defined by

$$\alpha = \frac{2^\ell \ell!}{\sqrt{N(2\ell)!}}. \tag{5.8}$$

I show in Appendix A that

$$\sum_{k=0}^{M-1} |G_k(\boldsymbol{\omega})|^2 = 1. \tag{5.9}$$

I now verify that the choices of $L_0(\boldsymbol{\omega})$ and $H_0(\boldsymbol{\omega})$ satisfy the three perfect reconstruction constraints.

1. Aliasing-free subbands:

   Let $L_1(r,\theta) = L_0(2r,\theta)$. The spectrum of $L_1(r,\theta)$ is

   $$L_1(r,\theta) = \begin{cases} 1 & \text{if } r \leq \frac{\pi}{4}, \\ \cos\left[\frac{\pi}{2}\log_2\left(\frac{2r}{\pi}\right)\right] & \text{if } \frac{\pi}{4} < r < \frac{\pi}{2}, \\ 0 & \text{if } r \geq \frac{\pi}{2}. \end{cases} \tag{5.10}$$

   It is clear that $L_1(r,\theta) = 0$ if $r > \pi/2$.

2. Recursion constraint:

   Let $H_1(r,\theta) = H_0(2r,\theta)$. Let $B_k(r,\theta) = H_1(r,\theta)G_k(r,\theta)$.

$$\begin{aligned} \sum_{k=0}^{M-1} |B_k(\boldsymbol{\omega})|^2 &= \sum_{k=0}^{M-1} |H_1(r,\theta)G_k(r,\theta)|^2 \\ &= |H_1(r,\theta)|^2 \sum_{k=0}^{M-1} |G_k(r,\theta)|^2 \\ &= |H_1(r,\theta)|^2. \end{aligned} \tag{5.11}$$

I aimed to show that $|L_1(r,\theta)|^2 + |H_1(r,\theta)|^2 = 1$, where $L_1(r,\theta)$ is given in (5.10) and $H_1(r,\theta)$ is given by

$$H_1(r,\theta) = \begin{cases} 0 & \text{if } r \leq \frac{\pi}{4}, \\ \cos\left[\frac{\pi}{2}\log_2\left(\frac{2r}{\pi}\right)\right] & \text{if } \frac{\pi}{4} < r < \frac{\pi}{2}, \\ 1 & \text{if } r \geq \frac{\pi}{2}. \end{cases} \tag{5.12}$$

From these definitions, for $r \leq \frac{\pi}{4}$ and $r \geq \frac{\pi}{2}$, we have $|L_1(r,\theta)|^2 + |H_1(r,\theta)|^2 = 1$. For $\frac{\pi}{4} < r < \frac{\pi}{2}$, let $S = |L_1(r,\theta)|^2 + |H_1(r,\theta)|^2$. Thus,

$$\begin{aligned} S &= \cos^2\left[\frac{\pi}{2}\log_2\left(\frac{4r}{\pi}\right)\right] + \cos^2\left[\frac{\pi}{2}\log_2\left(\frac{2r}{\pi}\right)\right] \\ &= \cos^2\left[\frac{\pi}{2}\log_2\left(\frac{2r}{\pi}\right) + \frac{\pi}{2}\log_2(2)\right] + \cos^2\left[\frac{\pi}{2}\log_2\left(\frac{2r}{\pi}\right)\right] \\ &= \sin^2\left[\frac{\pi}{2}\log_2\left(\frac{2r}{\pi}\right)\right] + \cos^2\left[\frac{\pi}{2}\log_2\left(\frac{2r}{\pi}\right)\right] \\ &= 1. \end{aligned} \tag{5.13}$$

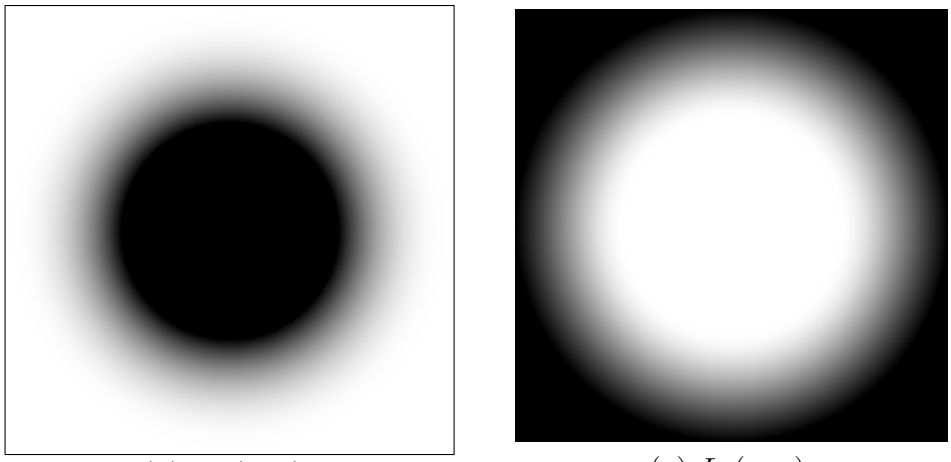
Therefore, $|L_1(r,\theta)|^2 + |H_1(r,\theta)|^2 = 1$.

3. Perfect reconstruction:

   The result follows directly from (5.13) that $|L_0(r,\theta)|^2 + |H_0(r,\theta)|^2 = 1$.

Since the chosen filters satisfy the three perfect reconstruction constraints, the steerable pyramid is a perfect reconstruction transform. I show the frequency responses of $L_0(-\boldsymbol{\omega})$ and $H_0(-\boldsymbol{\omega})$ in Fig. 5.4. The frequency magnitude varies from zero to one where one is represented with white and zero is represented with black.

While the original steerable pyramid allows perfect reconstruction, the highpass filter $H_0(-\boldsymbol{\omega})$ in Fig. 5.3 is not orientation selective. As a result, it is unsuitable for generating locally narrowband AM-FM image components. In



(b) $H_0(-\boldsymbol{\omega})$. (a) $L_0(-\boldsymbol{\omega})$.

Figure 5.4: Frequency responses of $H_0(-\boldsymbol{\omega})$ and $L_0(-\boldsymbol{\omega})$.

Section 5.2.4, I break this highpass component into M orientation selective sub-channels. In addition, I decompose the residual lowpass channel into M separate sub-channels of different orientations. With these two modifications, I develop a multi-component perfect reconstruction AM-FM image transform in Section 5.3. The modified steerable pyramid is a true multi-scale multi-orientation decomposition.

5.2.4 The Modified Steerable Pyramid

As mentioned in Section 5.2.3, the original steerable pyramid (OSP) retains one highpass component. Depending on the input image, the highpass component may have pixels that exhibit multiple orientations. The single component AM-FM model is insufficient to represent such component. Therefore, I further decomposed this component into M oriented components. The decomposition scheme for the highpass component is illustrated Fig. 5.5. Compared to the original decomposition in Fig. 5.3, the highpass component of the modified steerable pyramid (MSP) has an additional element denoted with dotted lines. The output of the

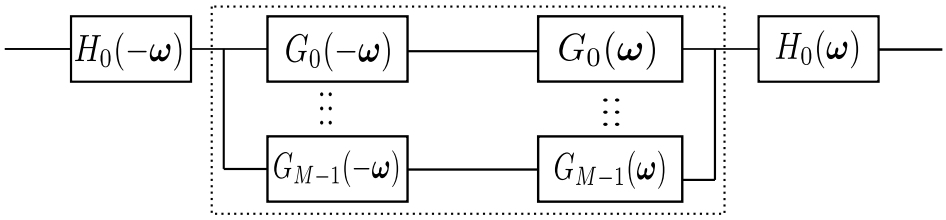


Figure 5.5: Highpass decomposition scheme of the MSP.

analysis highpass filter is filtered with M oriented filters G_k resulting M oriented components.

Let $\widehat{f}(\boldsymbol{\omega})H_0(-\boldsymbol{\omega})$ be the output of the highpass filter $H_0(-\boldsymbol{\omega})$. Each oriented component $\widehat{s}_k(\boldsymbol{\omega})$ is computed as

$$\widehat{s}_k(\boldsymbol{\omega}) = \widehat{f}(\boldsymbol{\omega})H_0(-\boldsymbol{\omega})G_k(-\boldsymbol{\omega}), \tag{5.14}$$

where $k \in [0, M-1]$. I show in Appendix A that the M component decomposition with filters $G_k, 0 \leq k \leq M-1$ is invertible. In other words, the highpass component $\widehat{f}(\boldsymbol{\omega})H_0(-\boldsymbol{\omega})$ can be reconstructed from the M oriented components without error, i.e.,

$$\begin{aligned}
\widehat{f}(\boldsymbol{\omega})H_0(-\boldsymbol{\omega}) &= \sum_{k=0}^{M-1} \widehat{f}(\boldsymbol{\omega})H_0(-\boldsymbol{\omega})G_k(-\boldsymbol{\omega})G_k(\boldsymbol{\omega}) \\
&= \widehat{f}(\boldsymbol{\omega})H_0(-\boldsymbol{\omega}) \sum_{k=0}^{M-1} G_k(-\boldsymbol{\omega})G_k(\boldsymbol{\omega}) \\
&= \widehat{f}(\boldsymbol{\omega})H_0(-\boldsymbol{\omega}) \underbrace{\sum_{k=0}^{M-1} |G_k(\boldsymbol{\omega})|^2}_{1} \\
&= \widehat{f}(\boldsymbol{\omega})H_0(-\boldsymbol{\omega}).
\end{aligned} \tag{5.15}$$

The residual lowpass component, which is obtained with the lowpass filter $L_1(-\boldsymbol{\omega})$ and the downsampling by a factor of two in Fig. 5.3, can also form M

oriented components. Such additional decomposition is necessary if the residual lowpass component contains significant image structures that a single component AM-FM model can not fully capture. To generate M oriented components from the lowpass residual, I applied the same decomposition scheme as for the highpass component to the lowpass residual component. I illustrate the orientation decomposition of the highpass component and of the lowpass component in Fig. 5.6 and 5.7. The frequency response of the original lowpass component is shown in Fig. 5.6(a), while its decomposition into eight oriented sub-channels are illustrated in Fig. 5.6(b)-(i). The frequency response of the original highpass channel is shown in Fig. 5.7(a) which is then decomposed into eight oriented sub-channels. These oriented sub-channels are depicted in Fig. 5.7(b)-(i).

With these two modifications, I depict the MSP decomposition scheme in Fig. 5.8. Compared to the OSP in Fig. 5.3, the MSP in Fig. 5.8 has two extra signal processing boxes indicated by dotted lines. In each of these boxes, I performed an additional decomposition step to obtain M oriented components. Together, the MSP has $2M - 2$ more filter channels than the OSP.

After decomposing the highpass channel and lowpass channel of the OSP into $2M$ orientation selective sub-channels, one can choose to compute the AM-FM representation either after the analysis step or after the synthesis step. Both approaches yield perfect reconstruction AM-FM representations of the original image. Specifically, one can compute the AM-FM representation from responses of analysis filters $G_k(-\boldsymbol{\omega})$ and $B_k(-\boldsymbol{\omega})$ and then use the synthesis filters $G_k(\boldsymbol{\omega})$ and $B_k(\boldsymbol{\omega})$ to reconstruct the original image. Alternatively, one can compute the AM-FM representation from the image components $f_k(\mathbf{x})$ individually by setting all but one input of the synthesis filterbank in Fig. 5.8 to zero. This produces

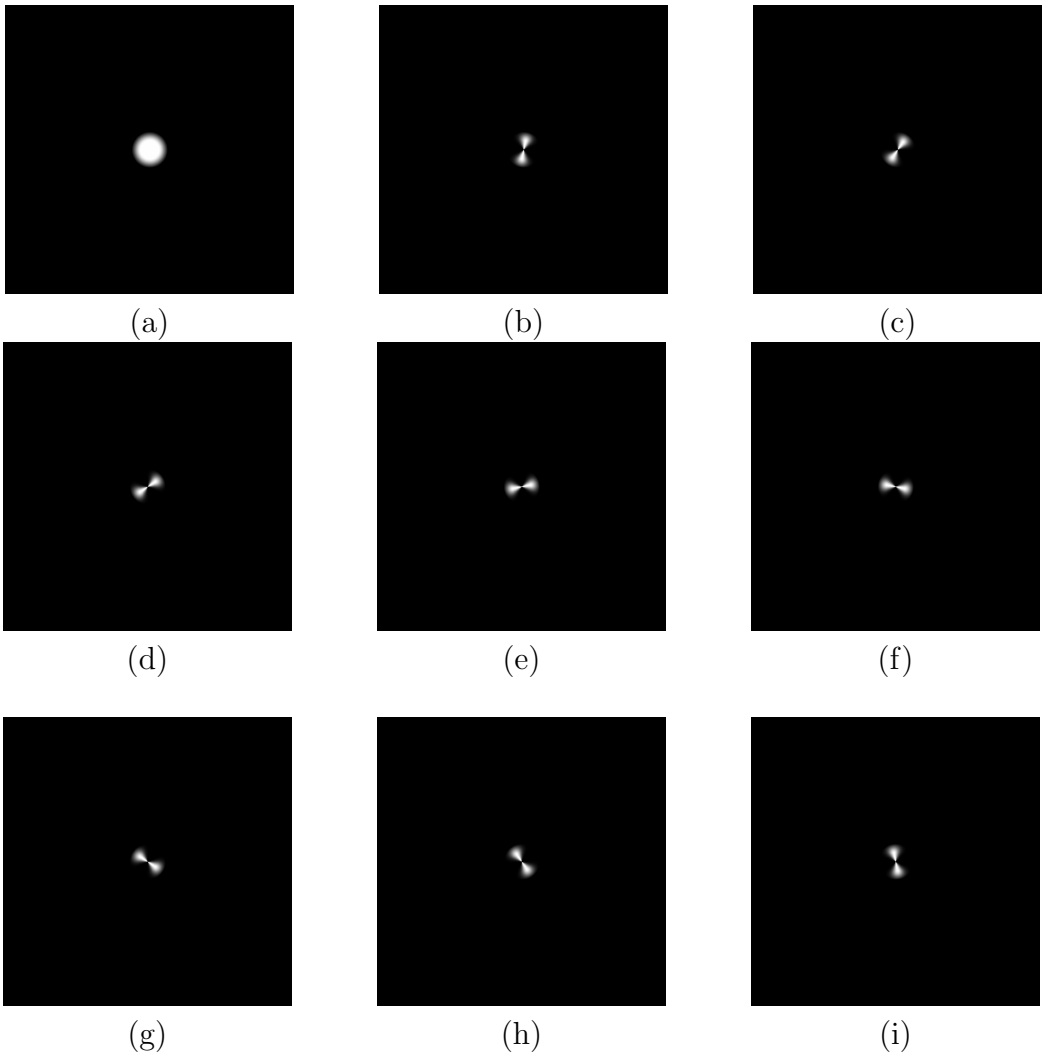


Figure 5.6: Decomposition of the lowpass channel into eight orientation selective sub-channels. (a) Original lowpass channel. (b) 1^{st} component. (c) 2^{nd} component. (d) 3^{rd} component. (e) 4^{th} component. (f) 5^{th} component. (g) 6^{th} component. (h) 7^{th} component. (i) 8^{th} component.

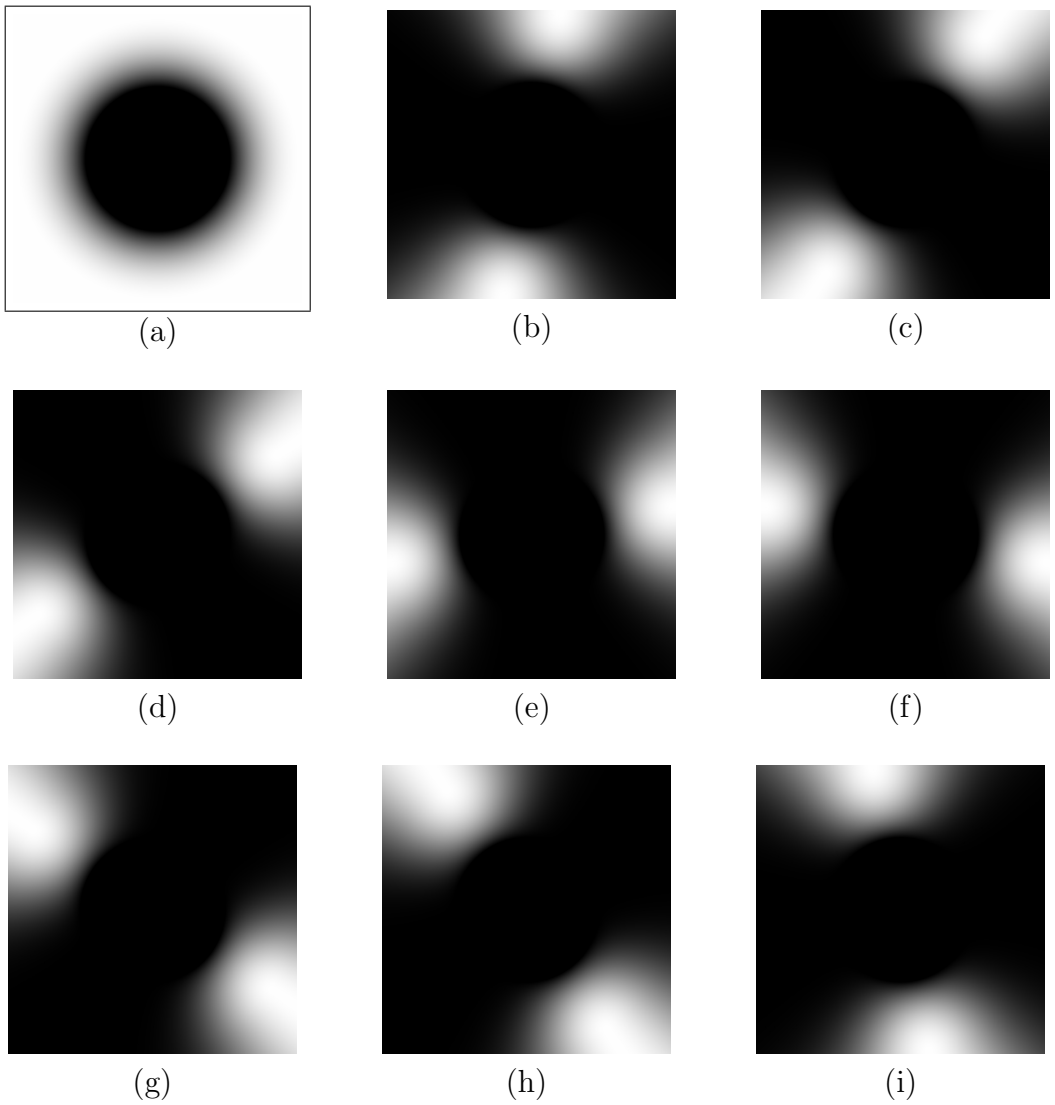


Figure 5.7: Decomposition of the highpass into eight orientation selective subchannels. (a) Original highpass channel. (b) 1^{st} component. (c) 2^{nd} component. (d) 3^{rd} component. (e) 4^{th} component. (f) 5^{th} component. (g) 6^{th} component. (h) 7^{th} component. (i) 8^{th} component.

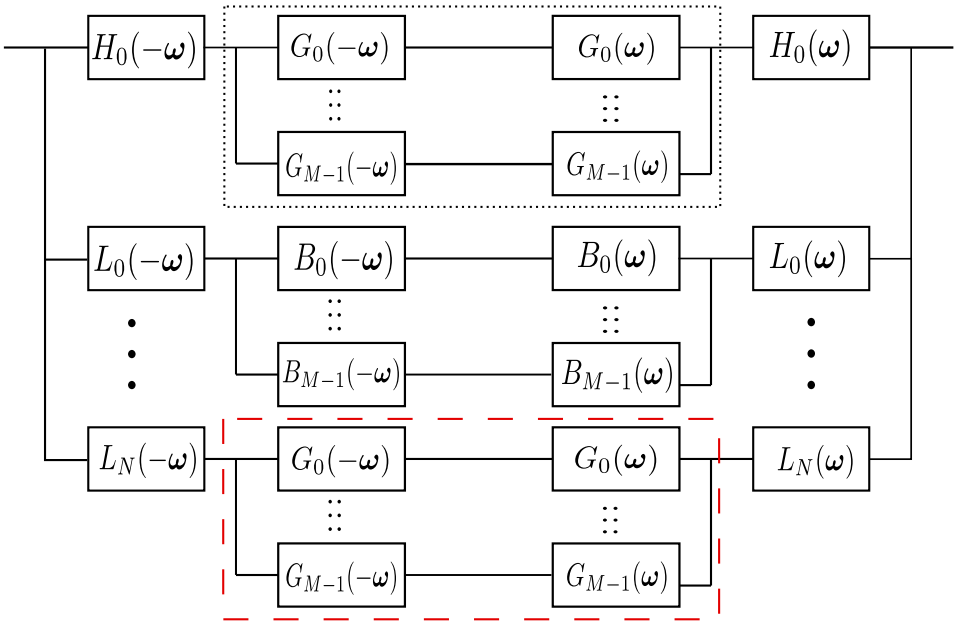


Figure 5.8: Decomposition scheme of the MSP.

the individual component component shown in Fig. 5.9, which may be added to obtain the reconstructed image. Here, I show that both approaches produce the same reconstructed image. Let $\tilde{f}_{0,0}$ be the reconstructed image obtained from the MSP in Fig. Fig. 5.8 by setting all but the top input to the synthesis filterbank on the right half of the figure equal to zero. Let $f_{0,0}$ be the top output of the MSP in Fig. 5.9. I show this equality in the frequency domain as

$$\begin{aligned}
\widehat{\tilde{f}_{0,0}}(\boldsymbol{\omega}) &= \widehat{f}(\boldsymbol{\omega})H_0(-\boldsymbol{\omega})G_0(-\boldsymbol{\omega})G_0(\boldsymbol{\omega})H_0(\boldsymbol{\omega}) \\
&= \widehat{f}(\boldsymbol{\omega})|H_0(\boldsymbol{\omega})|^2|G_0(\boldsymbol{\omega})|^2 \\
&= \widehat{f_{0,0}}(\boldsymbol{\omega}). \qquad (5.16)
\end{aligned}$$

Similar calculation shows that zeroing out all but the k^{th} input to the synthesis filterbank in Fig. 5.8 produces exactly the k^{th} output of the MSP in Fig. 5.9.

I prefer to perform the image decomposition according to the latter ap-

proach, i.e., performing AM-FM modeling on the image components after the synthesis filters. This approach saves computational resources because the analysis and synthesis filters of the steerable pyramid can be implemented together in one filtering process as shown in Fig. 5.9. More importantly, if signal processing

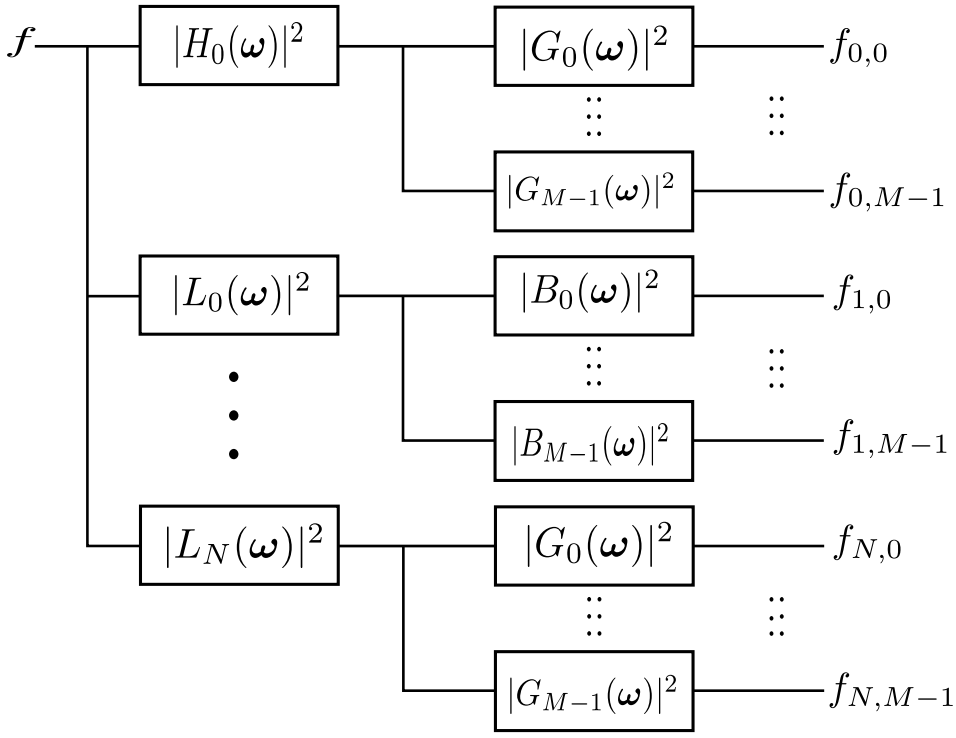


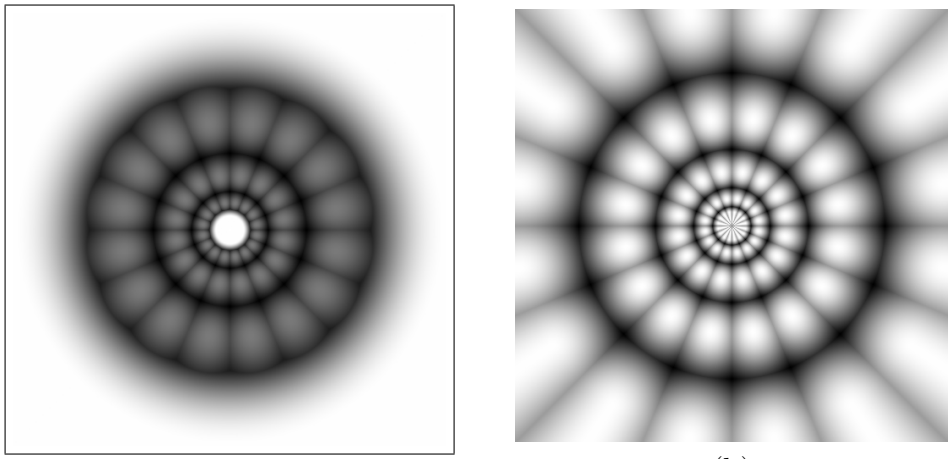
Figure 5.9: Alternative interpretation of the decomposition scheme of the MSP.

is to be performed on the AM and FM functions, this implies that they will be the modulating functions of the fully reconstructed image component, which relate well to human visual perception, as opposed to those of the unreconstructed image components occurring in the center of Fig. 5.8 which is less intuitive and significantly more difficult to interpret.

Let N be the number of decomposition levels (scales) and let M be the number of orientations in a level. The MSP decomposes an input image into $(N+$

1)M subcomponents. I show the frequency spectrum of the OSP and the MSP in Fig. 5.10, where the number of scales is $N = 5$ and the number of orientations $M = 8$. These numbers are chosen to be comparable with the widely used biologically motivated 43-channel Gabor filterbank in [49]. The range of intensities in these two spectral plots varies from zero to one where one is represented with white and zero is represented with black. In actuality, the frequency responses of adjacent channels overlap to some degree. For clarity, only magnitude of the channel with maximum response is shown at each point on the frequency plane in Fig. 5.10.

The difference between between the original filterbanks and the modified filterbank is obvious in the inner most and outer most rings of the filters in the figure. The inner most and outer most rings of the original filterbank in Fig. 5.10(a) are each decomposed into M orientation selective channels in the modified steerable pyramid of Fig. 5.10(b).



(a) (b)

Figure 5.10: Steerable Pyramid filterbanks. (a) OSP filterbank. (b) MSP filterbank.

5.3 The Multi-component PR AM-FM Transform

Before describing the multi-component AM-FM transform, I review the demodulation process by which the AM and FM functions are computed. I first construct the imaginary image $q(\mathbf{n})$ from the real input image using the partial Hilbert transform (pHT) described in Section 3.2.1. From these two signals, the complex image $z(\mathbf{n}) = f(\mathbf{n}) + jq(\mathbf{n})$ is created, where $q(\mathbf{n}) = \mathcal{H}_\theta[f(\mathbf{n})]$ and $\mathcal{H}_\theta$ is the pHT acting in the direction θ. The AM and FM functions may then be computed as [49]

$$a(\mathbf{n}) = |z(\mathbf{n})|, \tag{5.17}$$

$$\nabla\varphi(\mathbf{n}) = \mathrm{Im}\left[\frac{\nabla z(\mathbf{n})}{z(\mathbf{n})}\right]. \tag{5.18}$$

5.3.1 The Rotated Hilbert Transform

Because the pHT requires the direction of action to be specified, the output of the pHT is directional dependent; different angles θ can result in different imaginary image $q(\mathbf{n})$. Consequently, the computed AM and FM functions also depend on θ. The singularity in the pHT kernel can induce strong artifacts in the computed AM and FM functions along a line perpendicular to θ in the image. These are most often manifested as amplitude rippling artifacts and potentially as significant distortions of the FM. One of the important contributions of this dissertation is that I devised a novel scheme for rotating the pHT axis away from the significant spectral support of the image on a component by component basis in a multi-component AM-FM representation. Since the channels of the modified steerable pyramid used to isolate the image components $f_k(\mathbf{x})$ (or $f_k(\mathbf{n})$ in the discrete case) in (5.1) are highly orientation selective, an effective approach is to rotate the direction of action of the pHT, which used to compute the complex extension of each

image component, as far as possible away from the significant spectral content of the component. This can be achieved by setting θ in the pHT $\mathcal{H}_\theta$ orthogonal to the orientation of the center frequency vector of the steerable pyramid channel filter. This can be done on a component by component (e.g., channel by channel) basis to ensure that the arisen artifacts has little or no energy. Therefore, this new approach drastically reduces impact of the amplitude rippling artifacts and FM orientation distortions relative to a naïve approach of applying a fixed pHT (e.g., $\mathcal{H}_0$) to all the channels in the steerable pyramid. This rotation results in substantially improved correspondence between the computed modulations and human visual perception of the image structure. This improvement is an important advance that effectively overcomes one of the main obstacles that has precluded the development of practical perceptually motivated signal processing directly in terms of the AM and FM functions.

I illustrate the amplitude rippling artifacts, FM distortions, and improvements obtained with the rotated pHT in Fig. 5.11 and 5.12. Fig. 5.11(a) shows the original Barbara image. The real image component $f_{4,1}$ from the first oriented filter in level four is given in Fig. 5.11(b). The computed imaginary images obtained with the rotated pHT $\mathcal{H}_\theta$ and with the standard pHT $\mathcal{H}_0$ are depicted in Fig. 5.11(c) and (d). The spectral support of the complex images $z(\mathbf{n})$ obtained for $f_{4,1}$ using $\mathcal{H}_\theta$ and $\mathcal{H}_0$ is illustrated in Fig. 5.11(e) and (f), respectively. While the differences between the two imaginary images may appear subtle, the artifacts that are induced in the corresponding AM and FM functions are different in Fig. 5.12(c)-(f). The amplitude rippling artifacts induced by $\mathcal{H}_0$ are obvious, whereas they are entirely suppressed in the AM computed with the $\mathcal{H}_\theta$. In addition, the direction θ of the pHT also affects the computed FM. Observe that the

FM field computed with $\mathcal{H}_0$ in Fig. 5.12(f) shows needles pointing to wrong texture orientations. Similar examples of Mandrill image appear in Fig. 5.13 and 5.14.

5.3.2 The Multi-component AM-FM transform

Equipped with the PR filterbank discussed in Section 5.2.4 and the single component PR AM-FM model described in Sec. 4.5, I defined the lossless multi-component AM-FM transform (xAMFM). The xAMFM consists of the analysis (forward) transform and the synthesis (backward) transform.

Fig. 5.15 shows the block diagram of the analysis xAMFM. First, the input image $f(\mathbf{n})$ is decomposed into K components using the modified steerable pyramid described in Section 5.2.4. I then compute the AM function $a_k(\mathbf{n})$ and FM function $\nabla\varphi_k$ from each component f_k using the perfect reconstruction FM demodulation algorithm described in Section 4.5. Fig. 5.15 indicates that the whole demodulation process of K components can be done in in parallel. The overall complexity of the analysis xAMFM is $O(N\log(N))$, where N is the larger of the horizontal and vertical image dimensions.

The synthesis transform is shown in Fig. 5.16. I reconstruct the original real image component $f_k(\mathbf{n})$ from the AM and FM functions according to the synthesis algorithm described in Section 4.5. In particular, the phase function $\varphi(\mathbf{n})$ is reconstructed by integrating the FM functions. Finally, I obtain the original image as the linear sum of K component according to

$$f(\mathbf{n}) = \sum_{k=0}^{K} f_k(\mathbf{n}) = \sum_{k=0}^{K} a_k(\mathbf{n})\cos[\varphi(\mathbf{n})]. \tag{5.19}$$

Since each component $f_k(\mathbf{n})$ is perfectly reconstructed as described in Section 4.5, the reconstructed image from K components is identical to the original

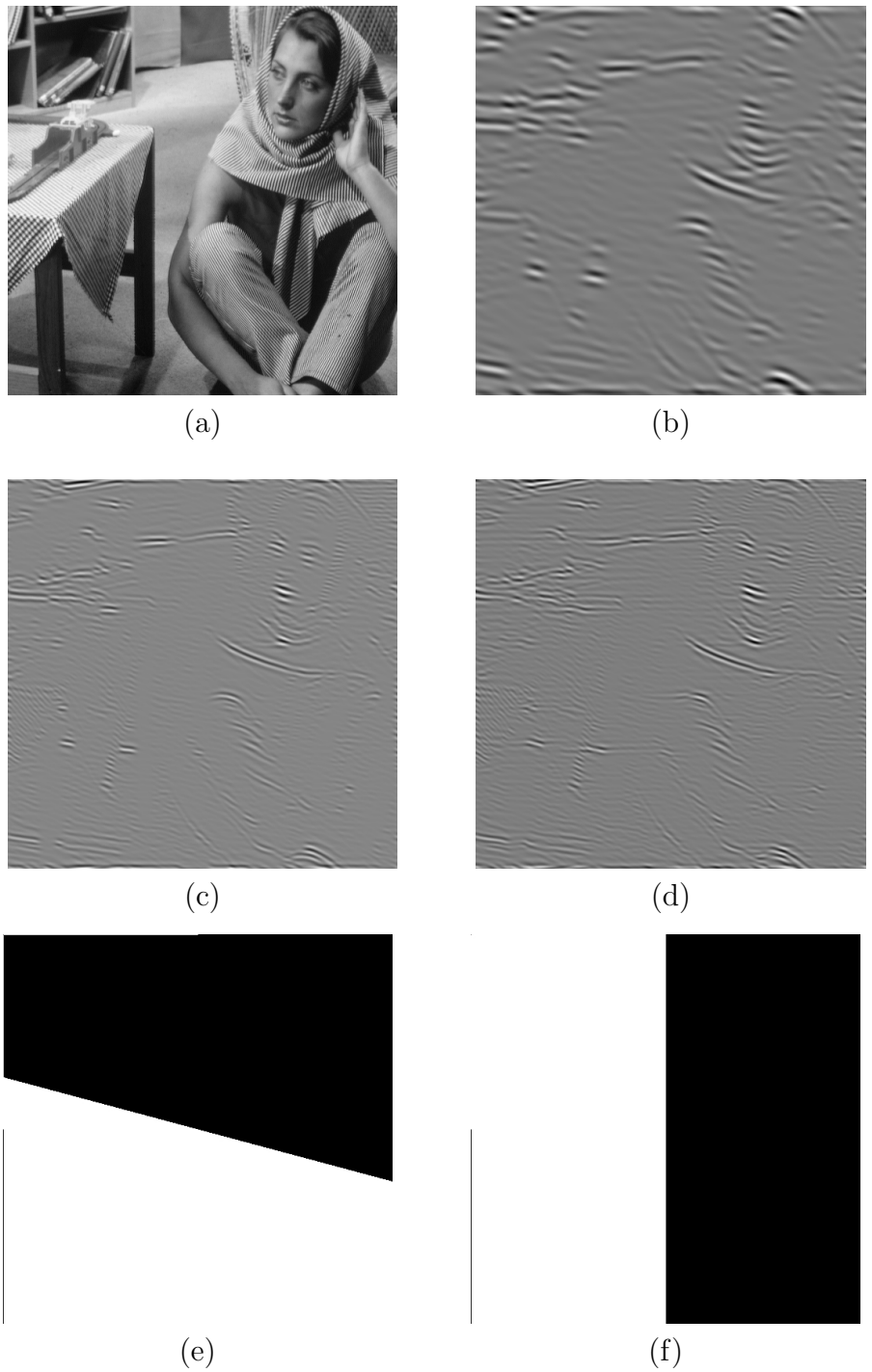


Figure 5.11: Barbara: imaginary image computed with rotated pHT and with $\mathcal{H}_0$. (a) Original Barbara image. (b) Component $f_{4,1}$. (c) $q_{4,1}$ with rotated pHT. (d) $q_{4,1}$ with $\mathcal{H}_0$. (e) Rotated pPHT. (f) pHT with $\theta = 0$.

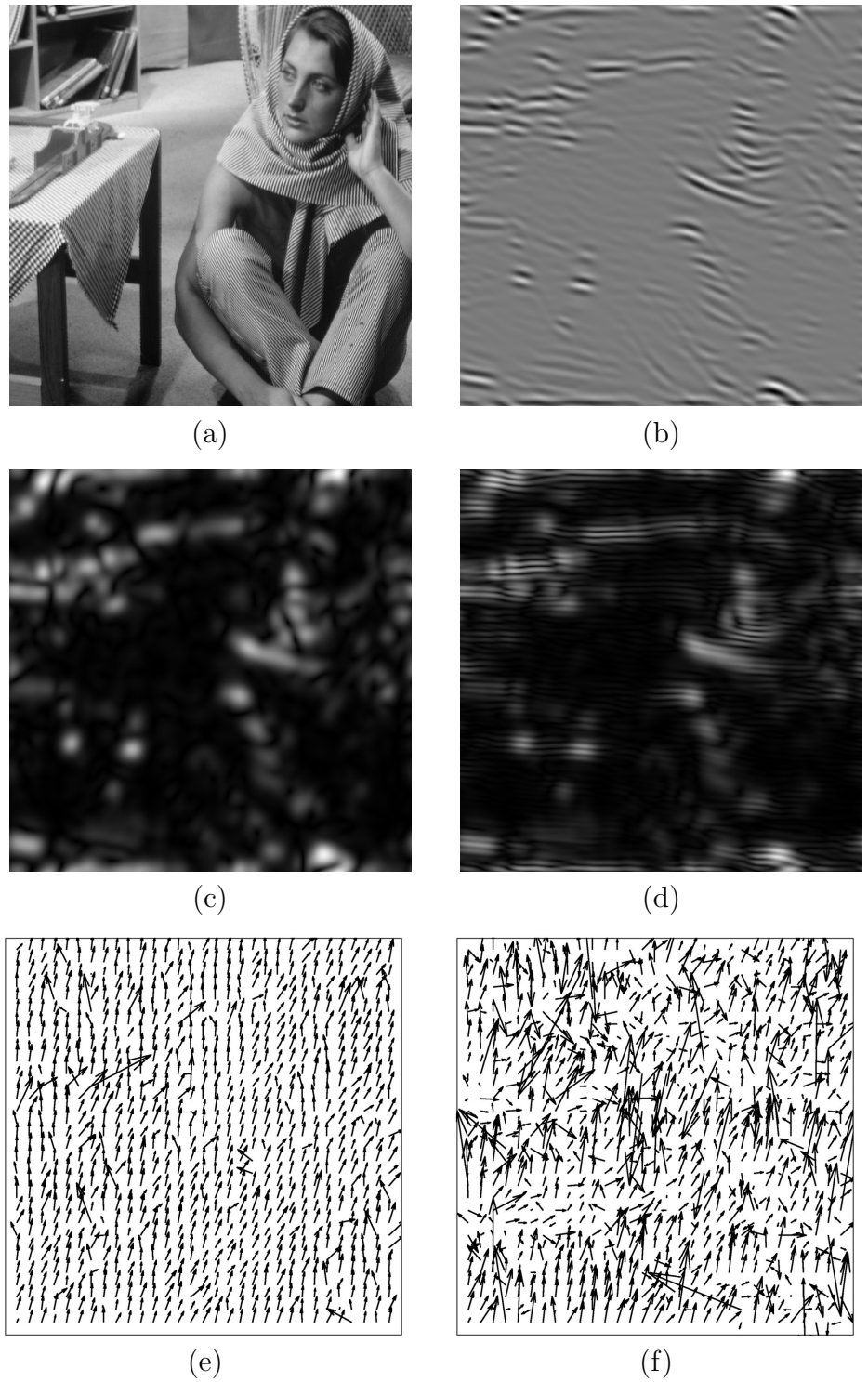


Figure 5.12: Barbara: computed AM-FM with rotated pHT and pHT with $\mathcal{H}_0$. (a) Original Barbara image. (b) Component $f_{4,1}$. (c) AM with rotated pHT. (d) AM with $\mathcal{H}_0$. (e) FM field of rotated pHT. (f) FM field with $\mathcal{H}_0$

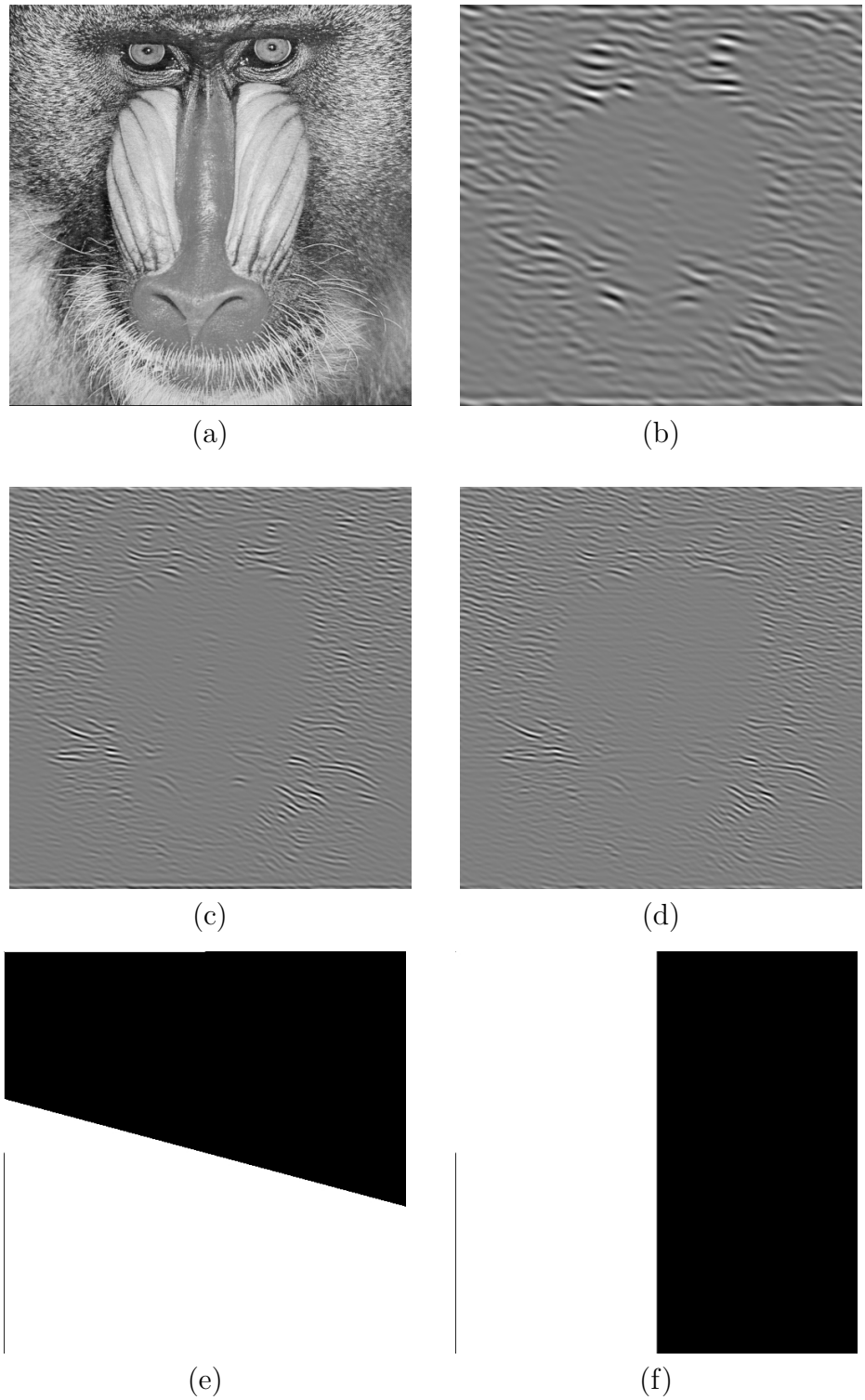


Figure 5.13: Mandrill: imaginary image computed with rotated pHT and with $\mathcal{H}_0$. (a) Original Mandrill image. (b) Component $f_{4,1}$. (c) $q_{4,1}$ with rotated pHT. (d) $q_{4,1}$ with $\mathcal{H}_0$. (e) Rotated pPHT. (f) pHT with $\theta = 0$.

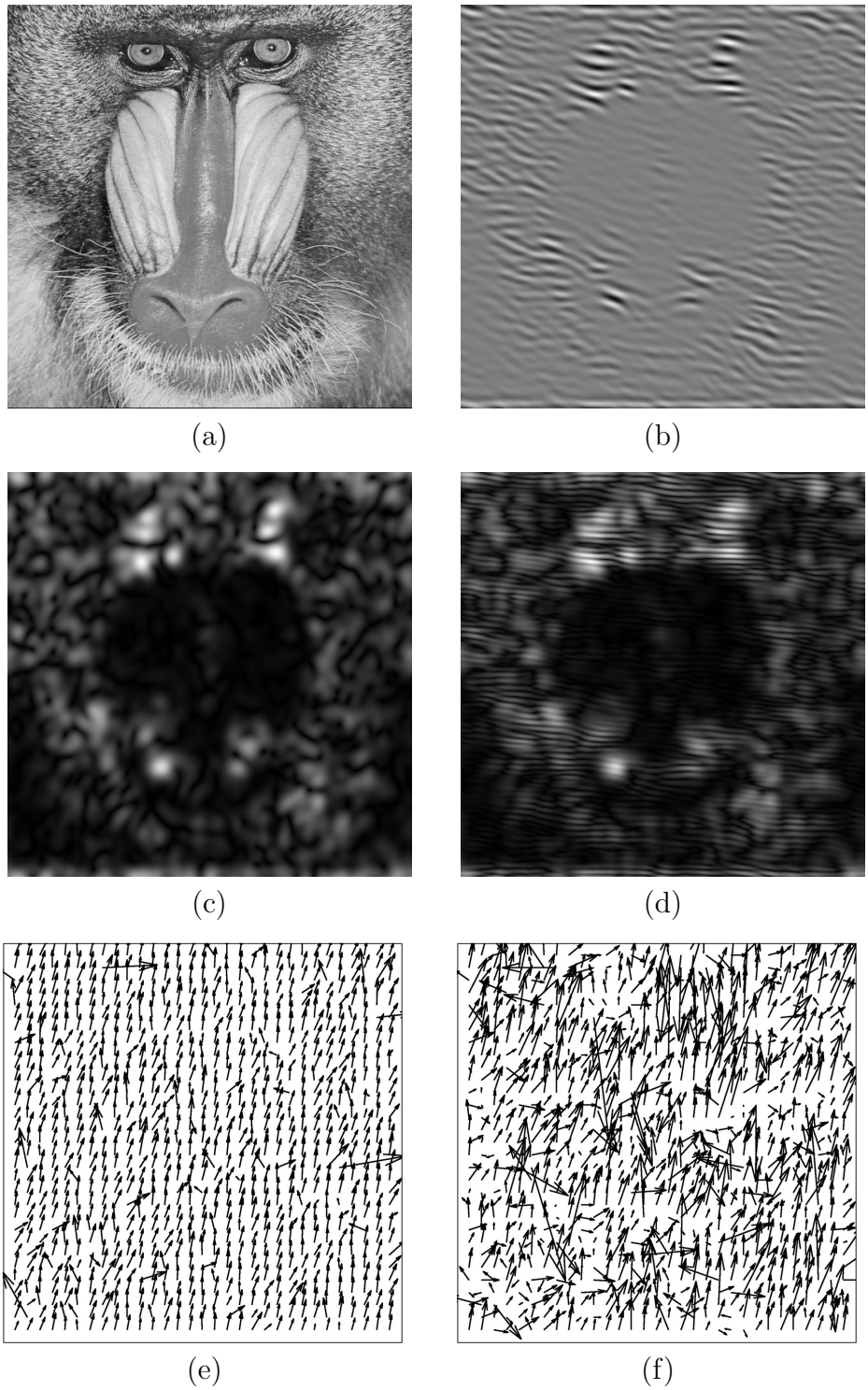


Figure 5.14: Mandrill: computed AM-FM with rotated pHT and $\mathcal{H}_0$. (a) Original Mandrill image. (b) Component $f_{4,1}$. (c) AM with rotated pHT. (d) AM with $\mathcal{H}_0$. (e) FM field of rotated PHT. (f) FM field with $\mathcal{H}_0$.

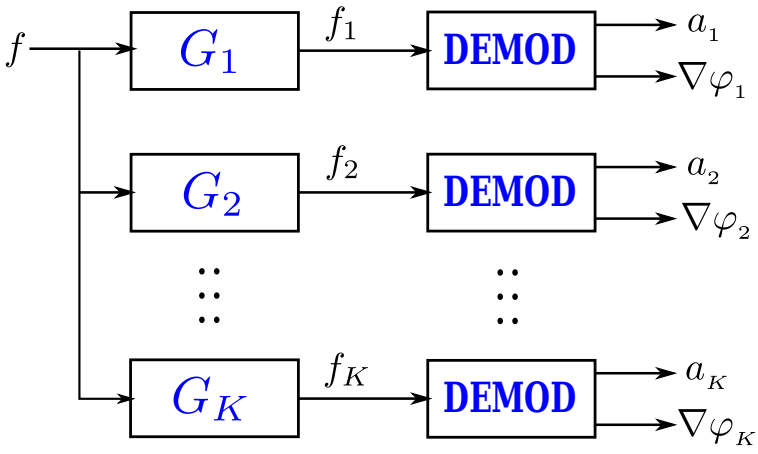


Figure 5.15: The analysis xAMFM.

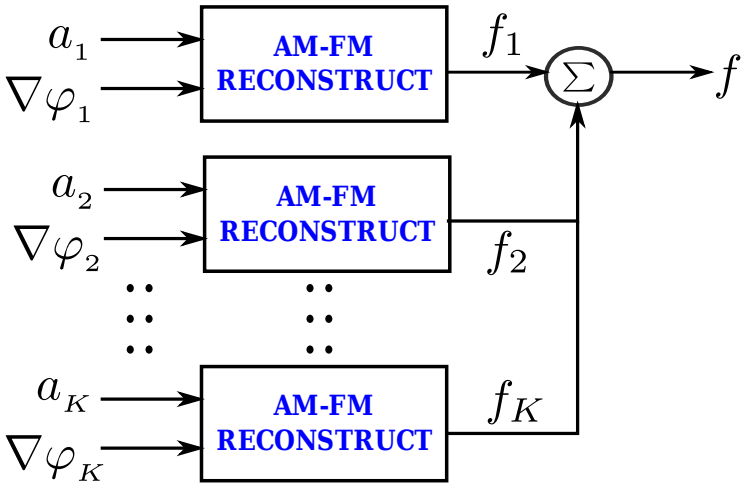


Figure 5.16: The synthesis xAMFM.

image. I measure error between the reconstruction error in terms of the mean squared error (MSE) and peak signal to noise ratio (PSNR). Definitions for these two measures are given in (4.30) and (4.31). The test images include standard grayscale test images and some color images from the Kodak image library [37]. The reconstruction errors statistics are shown in Table 5.1.

Table 5.1: Reconstruction error of the multi-component AM-FM transform.

	Dimension	PSNR (dB)	MSE
Lena	512x512	80.903801	4.874747×10^{-04}
Barbara	512x512	74.902187	1.957273×10^{-03}
Boat	512x512	78.522112	9.138380×10^{-04}
EinSlack	375x500	78.706215	8.085571×10^{-04}
Fingerprint	512x512	74.680474	2.213262×10^{-03}
Flintstones	512x512	73.169935	3.133912×10^{-03}
House	256x256	84.772617	1.903418×10^{-04}
kodim01	512x768	74.077160	2.543105×10^{-03}
kodim05	512x768	73.993926	2.592315×10^{-03}
kodim08	512x768	70.226141	6.172573×10^{-03}
kodim17	768x512	78.422051	9.351371×10^{-04}
kodim22	512x768	77.991403	1.032619×10^{-03}
kodim23	512x768	79.323793	7.598032×10^{-04}

5.4 Summary

In this chapter, I argued that the single component AM-FM model is not sufficient to represent most images of practical interest. A multi-component approach is required to fully capture important image features in a way that is natural, intuitive, and corresponds well to human visual perception. In addition, I claimed that a multi-scale multi-orientation filterbank such as the steerable pyramid is a suitable technique for decomposing an image into meaningful, locally narrowband components for the AM-FM representation.

To integrate the original steerable pyramid into the AM-FM transform, I modified the original design to create a true multi-scale and multi-orientation image decomposition. The modified filterbank is similar in many respects to the perceptually motivated Gabor filterbank, particularly with regards to its joint time-frequency localization properties.

I overcame one of the most important problem associated with the partial Hilbert transform (pHT). The computed AM and FM functions generally contain artifacts if the image component has frequency content perpendicular to the pHT filtering axis. I effectively solved this problem by rotating the filtering axis of the pHT on a component by component basis. Finally, I defined the multi-component AM-FM transform (xAMFM) and showed quantitatively that the xAMFM provides perfect reconstruction.

With the perfect reconstruction xAMFM, the door is open to begin development of practical processing operations in the modulation domain.

Chapter 6

AM-FM Image Processing

6.1 Motivation for Transform-Domain Filtering

Discrete image filters are usually designed to operate on the pixel grey levels intensities, which I refer to as pixel domain processing. There is a rich and established theory of linear shift-invariant (LSI) filters which may equivalently be implemented in the frequency domain by multiplication of the Fourier transforms. In the pixel domain, LSI filtering is implemented by the sliding window technique. The filter output at a pixel location (m, n) is obtained by computing the dot product of the filter kernel and the pixel values lying under the kernel support centering at (m, n). The filter kernel is translated to every pixel in the image in order to compute the full output image.

For any given signal processing task, an important question is whether filters operating in another domain yield better performance than a direct pixel domain implementation. The answer depends on the applications. There are applications where transform domain approaches offer significant advantages compared to the pixel domain. For example, higher compression rates can be obtained when the compression takes place in the transform domain such as those associated with the discrete cosine transform (DCT) [126] or wavelets [3]. In JPEG image compression, the image is first transformed to the DCT domain on a block by block basis where each block is encoded in terms of its quantized and entropy coded DCT

coefficients. In the DCT domain, high value coefficients are more important than low value ones. Therefore, compression gain is achieved by allocating fewer bits for small coefficients in the encoding process. In addition, the selection of a suitable transform affects the final result. For instance, image coding in the wavelet domain generally yields a higher compression rate than that of the DCT domain [111].

Wavelet image denoising is another example of transform-domain applications that achieves state of the art performance. First, the noisy image is transformed from the pixel domain to the wavelet domain. The wavelet coefficients are then thresholded because low coefficients tend to correlate with noise [30]. Advanced statistical techniques such as wavelet coefficient modeling [99] and transform-domain collaborative filtering [27] can improve the denoising performance.

The main advantage of transform domain approaches over pixel domain approaches is that transform domain techniques can capture important signal features that are not salient in the pixel domain. Bovik, Clark, and Geisler [13] used the AM-FM model to capture nonstationary image features such as texture orientations and texture granularity. They represented an image with a multicomponent AM-FM image models (5.1) [14, 49]. Given the image $f(\mathbf{x})$, a computed AM-FM model consists of estimates of the K AM functions $a_k(\mathbf{x})$ which provide a dense local characterization of the local texture contrast and the K FM functions $\nabla\varphi_k(\mathbf{x})$ which provide a dense characterization of the local texture orientation and pattern spacing.

Such models have been used with great success in a variety of image *analysis* applications, including texture segmentation, 3D shape from texture, texture-based stereopsis, fingerprint classification, content-based retrieval, and regeneration of occluded and damaged textures [52], as well as for infrared target tracking [51] and

in the analysis of 2D spectrograms of human speech signals [32]. Prior to the work described in this dissertation, however, they have been considerably less successful in applications requiring image *synthesis* in addition to analysis (to the best of my knowledge, reconstruction from a computed AM-FM model has been attempted previously only in [48, 49, 52, 72, 108, 109]). The reason is that some means must be devised for decomposing the image into a sum of components (5.1) that are *locally coherent* [14, 49] and for isolating these components from one another on a jointly localized basis in space and spatial frequency prior to demodulation. Because of their excellent joint localization properties, banks of Gabor or Gabor-like bandpass filters have been used almost universally for this purpose. Unfortunately, the properties that make these filters attractive, *viz.*, localization and smoothness of the impulse and frequency responses simultaneously, also imply that they cannot provide perfect reconstruction [52].

While the continuous Gabor filter is optimal in the sense of attaining the low bound on joint localization of the time-frequency uncertainty principle, it is not invertible because it fails to admit compact frequency support. Hence, it is not suitable for image synthesis applications. The steerable filters, despite being suboptimal relative to the Gabor filters with respect to joint localization, offers perfect reconstruction while still providing excellent joint time-frequency localization. These properties are crucial in AM-FM image synthesis applications. Therefore, I incorporated the steerable filters in the xAMFM as described in Chapter 5.

I illustrate the AM-FM image processing framework in Fig. 6.1. The input signal f is first decomposed into K component f_k, $1 \leq k \leq K$. Demodulation algorithm is performed on each component f_k to compute the AM function a_k and FM function $\nabla\varphi_k$. Upon completion of the demodulation step, the image f

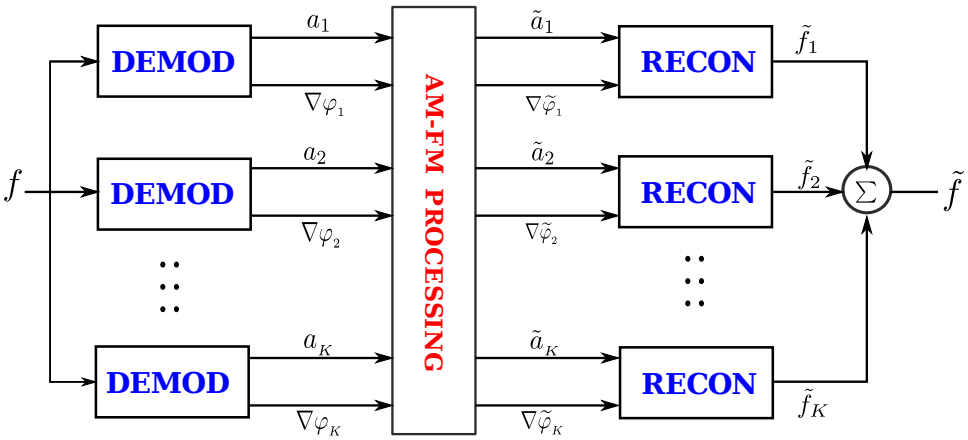


Figure 6.1: AMFM-based Image Filtering.

has been transformed into the modulation domain, i.e., it has been represented with K AM-FM components. I now define the AM-FM image processing block in Fig. 6.1 where the AM and FM functions are filtered with modulation domain filters, the design of which is the main subject of this chapter. Subsequent to AM-FM filtering, the processed image component $\tilde{f}_k$ is reconstructed from the filtered modulating functions $\tilde{a}_k$ and $\nabla\widetilde{\varphi}_k$. The overall output image is computed as a linear sum of the K reconstructed components $\tilde{f}_k$. For applications using the FM, the FM function is usually represented in polar form $\nabla\varphi_k(\mathbf{n}) = [r_k(\mathbf{n})\,\theta_k(\mathbf{n})]$, where $r_k(\mathbf{n}) = |\nabla\varphi_k(\mathbf{n})|$ is the frequency magnitude of the FM vector and $\theta_k(\mathbf{n}) = \arg\nabla\varphi_k(\mathbf{n})$ is the orientation of the FM vector.

I divide the discussion of AM-FM image filters into two major classes, namely the AM-based filters and the AM-FM filters. The AM-based filters interact primarily with the AM functions, while the AM-FM filters operate on both the AM and FM functions. I compare the effectiveness of the AM-based and AM-FM filters with traditional LSI filters in several illustrative image filtering tasks.

6.2 AM-based Filters

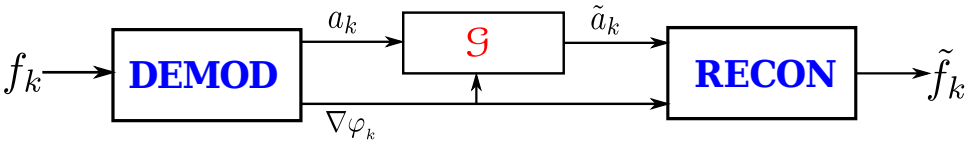


Figure 6.2: AM-based Image Filtering.

I define the AM-based filters to be operations where the AM functions are filtered. A block diagram of AM-based filtering for the component $f_k(\mathbf{n})$ is illustrated in Fig. 6.2. In the AM-based scheme, the FM functions are not modified. The FM is used as an auxiliary measurement to provide texture orientations and frequency content that may be used to determine what processing is applied to the AM functions and at what pixels. For each component $f_k(\mathbf{n})$, I apply the same filter $\mathcal{G}$ to the AM function $a_k(\mathbf{n})$. The output of this filtering process is denoted $\widetilde{a}_k(\mathbf{n})$. The phase $\varphi_k(\mathbf{n})$ is obtained by integrating the FM $\nabla\varphi_k(\mathbf{n})$. Notice that the frequency and phase are not affected by the filtering process because the FM functions were not modified. The processed component f_k is then reconstructed according to

$$\widetilde{f}_k(\mathbf{n}) = \widetilde{a}(\mathbf{n})\cos[\varphi_k(\mathbf{n})]. \tag{6.1}$$

6.2.1 Orientation Selective AM Filtering

Orientation Selective Attenuation of Structure in Synthetic Image

A synthetic radial chirp image is shown in Fig. 6.3(a). By construction, the AM is constant and the phase is quadratic along radials emanating from the center of the image. The signal processing goal is to attenuate nonstationary structure that is oriented at odd multiples of $\pi/4$. As a baseline for comparison, I implement an LSI notch filter with frequency response given in Fig. 6.3(b). The nonlinear AM

notch filter is given by

$$\widetilde{a}_k(\mathbf{x}) = \begin{cases} \frac{16}{\pi}\delta_k(\mathbf{x})a_k(\mathbf{x}), & \delta_k(\mathbf{x}) < \pi/16, \\ a_k(\mathbf{x}), & \text{otherwise}, \end{cases} \tag{6.2}$$

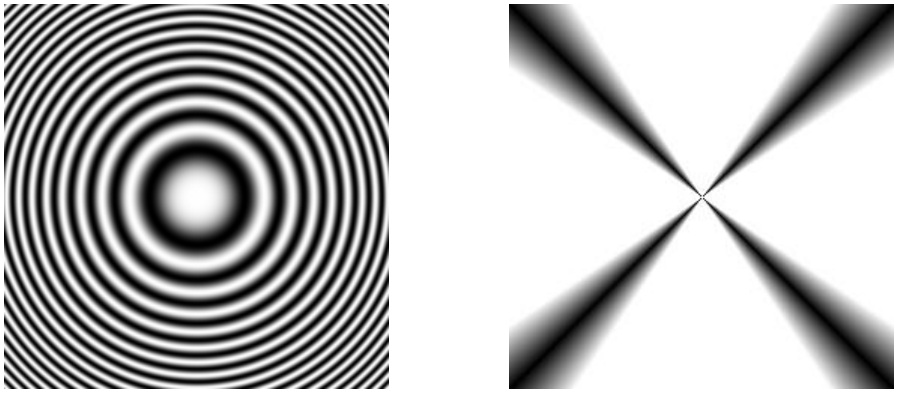
$$\nabla\widetilde{\varphi}_k(\mathbf{x}) = \nabla\varphi_k(\mathbf{x}), \tag{6.3}$$

where $\delta_k(\mathbf{x}) = |(|\theta_k(\mathbf{x})| - \pi/4)|$ is an amplitude scaling factor equal to the radian angular distance between $\theta_k(\mathbf{x})$ and $\pm\pi/4$.

The output of the LSI filter is shown in Fig. 6.3(c) and exhibits undesirable artifacts as expected: the LSI filter attenuates Fourier components on a spatially global scale, which achieves the desired result but also degrades the subtle constructive and destructive interference between Fourier components that creates the image structure at orientations other than odd multiples of $\pi/4$. This is demonstrated by Fig. 6.3(e), which gives the residual between the original image in Fig. 6.3(a) and the LSI processed image in Fig. 6.3(c). On the other hand, the AM-based filter is able to produce the desirable output which is shown in Fig. 6.3(d). The residual image in Fig. 6.3(e) indicates that texture components at odd multiples of $\pi/4$ are attenuated. Moreover, because the AM filter is capable of attenuating oriented structure on a spatially local basis, it delivers a result that is free of the undesirable artifacts seen in the LSI filter output. In contrast, the LSI filter tends to spread the filtering effects to a broader range of orientations and frequencies as seen in the residual image in Fig. 6.3(e).

Orientation Selective Attenuation of Structure in a Natural Image.

The natural wood grain texture image *Tree* given in Fig. 6.4(a) bears strong similarity to the synthetic chirp in Fig. 6.3(a). Here, I repeat the texture attenuation experiment on this image, but with a more aggressive notch of half-width $\pi/8$ in



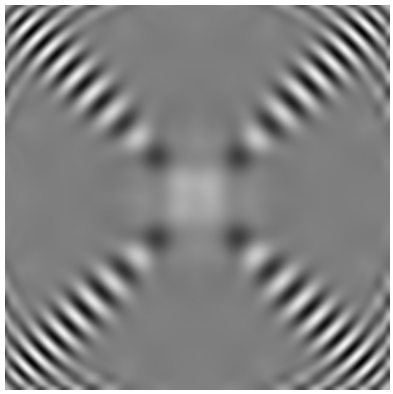
(a)

(b)

(c)

(d)

(e)



(f)

Figure 6.3: AM-based selective orientation attenuation. (a) Original chirp image. (b) Notch filter response. (c) LSI result. (d) AM-based result. (e) LSI residual. (f) AM-based residual.

order to attenuate a wider band of orientations. The result is given in Fig. 6.4(b) where it may again be seen that the perceptually-based signal processing goal has been achieved. The effect of this AM-based filtering is clearly visible in the left upper and lower quadrants of the image. Interesting subtle effects are also present. For example, careful examination of the center of the original image in Fig. 6.4(a), just to the right of the knot of the wood grain, reveals a small "hook" that is oriented along the main diagonal. Consistent with the processing goal, this hook is smoothly but totally obliterated in the result image of Fig. 6.4(b).

Spatially Selective Removal of Oriented Structure from a Natural Image.

The signal processing goal is to remove the bands from Lena's hat. The original image is given in Fig. 6.4(a). Let $\mathcal{X}$ denote the interior of the black rectangle shown superimposed on Fig. 6.4(b). AM-based filtering is applied to the components $f_k(\mathbf{n})$, but only for $\mathbf{n} \in \mathcal{X}$. With $\delta_k(\mathbf{n}) = |\theta_k(\mathbf{n}) - \pi/4|$, the AM-based filtering operation is given by

$$\widetilde{a}_k(\mathbf{n}) = \begin{cases} 0, & \delta_k(\mathbf{n}) < \pi/8 \text{ and } \mathbf{n} \in \mathcal{X}, \\ a_k(\mathbf{n}), & \text{otherwise} \end{cases} \tag{6.4}$$

and $\nabla\widetilde{\varphi}_k(\mathbf{n}) = \nabla\varphi_k(\mathbf{n})$. As shown in Fig. 6.4(c), the perceptually-based signal processing goal is achieved with a smooth, natural appearance. There are a few unwanted artifacts that result from the fact that we were quite imprecise in the specification the spatial region $\mathcal{X}$ desired for processing. For example, the central portions of the upper and lower edges of the hat brim were attenuated, as were certain orientations in the upper portion of the feather, and a slight shadow was induced on Lena's forehead. Interestingly, the shadow appearing on the upper portion of the hat is virtually unaffected. All of the unwanted artifacts could be

avoided by specifying the region $\mathcal{X}$ more precisely, but doing so would require increased effort in designing the filter.

(d) (e) (f)

Figure 6.4: AM-based texture removal. (a) Original Lena. (b) Operating window. (c) AM-based texture removal.

6.2.2 Frequency Selective Filtering

In addition to orientation selective processing, AM-based filters can also be used to perform spatially local amplification and attenuation based on magnitude frequency. To isolate a certain frequency range, I design the AM-based bandpass filter according to

$$\widetilde{a}_k(\mathbf{x}) = \begin{cases} a_k(\mathbf{x}), & 0.2 < r_k(\mathbf{x}) < 0.35, \\ 0, & \text{otherwise} \end{cases} \tag{6.5}$$

$$\nabla\widetilde{\varphi}_k(\mathbf{x}) = \nabla\varphi_k(\mathbf{x}), \tag{6.6}$$

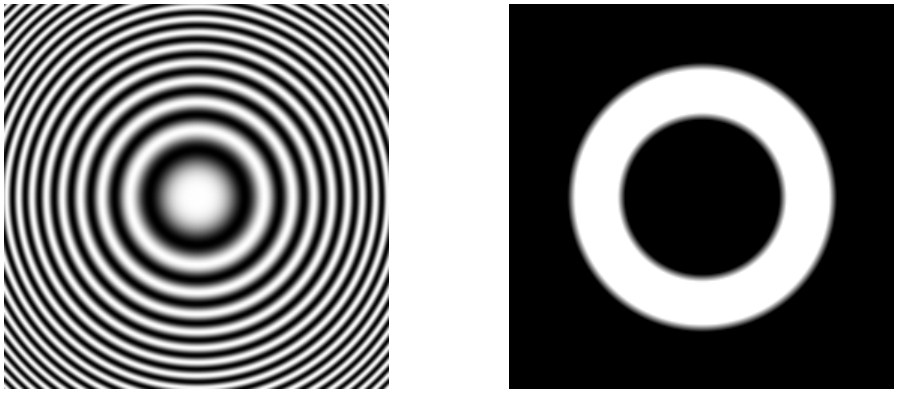
where $r_k(\mathbf{x})$ is given in units of cycles per pixel.

I apply the AM-based bandpass filter to the synthetic chirp image. The filtering result is given in Fig. 6.5. The image processing task is to isolate a circular ring from the original chirp image in Fig. 6.5(a). The frequency response of an appropriate LSI bandpass filter is given in Fig. 6.5(b). The cutoff frequencies were set to 0.2 and 0.35. The AM-based result is shown in Fig. 6.5(d). The AM-based

filter is able to isolate a ring of the original chirp image. I construct an equivalent LSI filter in the Fourier domain and depict its filtering result in Fig. 6.5(c). The LSI filter again spreads the filtering effects to a broader range of frequencies as seen in Fig. 6.5(c). The residuals of the LSI filter and the AM-based filters are shown in Fig. 6.5(e) and (f).

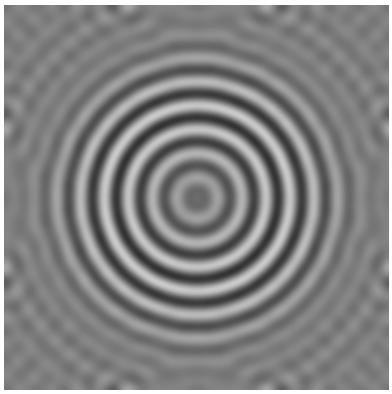
6.2.3 Selective Contrast Enhancement

Simplistic AM-based image enhancement is illustrated in Fig. 6.6. The original image is given in Fig. 6.6(a), while the image in Fig. 6.6(b) was obtained by applying a low-pass linear blur and adding Gaussian white noise. As a baseline comparison, the result of a naïve high-pass LSI filter approximating the pseudo-inverse is shown in Fig. 6.6(c). The design concept for the AM-based enhancement filter combines elements similar to both wavelet shrinkage and unsharp masking. The noise power is distributed widely throughout the steerable pyramid channels resulting in a relatively small contribution to the individual AM functions $a_k(\mathbf{n})$, whereas the coherent image structure tends to be jointly localized resulting in strong contributions to the AM functions, particularly in the vicinity of edges. I apply a simple threshold to the amplitude modulations computed from Fig. 6.6(a) to attenuate the noise, reconstruct, and then apply the same high-pass filter that is used in Fig. 6.6(c) to generate a high-pass mask. This mask is added back to the degraded image in Fig. 6.6(b) to obtain the enhanced/restored result shown in Fig. 6.6(d). We can see that the AM-based filter is able to attenuate the noise and enhance edges in the image. The pixel domain filter enhances the blurry image, but it also amplifies the noise.

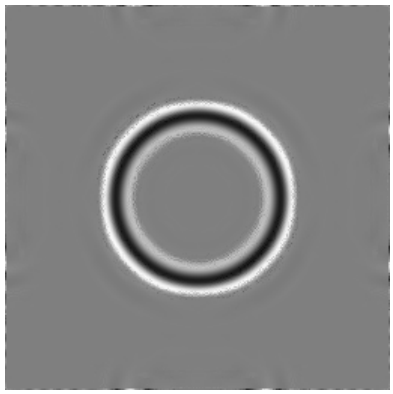


(a)

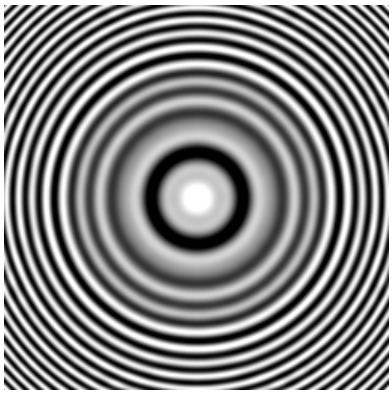
(b)



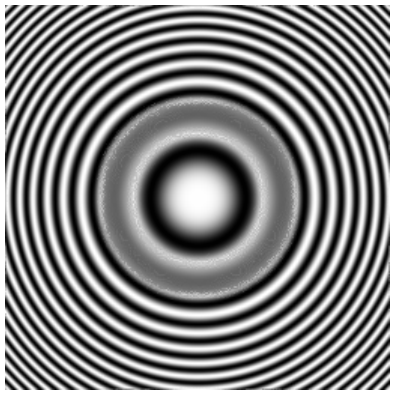
(c)



(d)



(e)



(f)

Figure 6.5: AM-based bandpass filter. (a) Original chirp image. (b) AM-based bandpass filter. (c) LSI result. (d) AM-based result. (e) LSI residual. (f) AM-based residual.

(a) (b)

(c) (d)

Figure 6.6: AM-based image enhancement by unsharp masking. (a) Original Lena. (b) Linear blur plus additive noise. (c) LSI highpass result. (d) AM-based enhancement result.

6.2.4 AM-based Image Fusion

Elementary examples of AM-based image fusion based on local contrast are given Fig. 6.7. The main idea is that sharp intensity edges imply high contrast and are associated with local AM values $a_k(\mathbf{x})$ that are relatively large, whereas defocused surfaces are associated with lower contrast and AM values that are relatively smaller than those in sharply focused or edge-rich regions. Thus, for a problem like the well-known pair of clock images shown in Fig. 6.7(a) and (b), a fused image showing both clocks in focus can be obtained as follows. I first compute a multi-component AM-FM model for each of the two input images shown in Fig. 6.7(a) and (b). For each pixel $\widetilde{f}_k(\mathbf{x})$ of AM-FM image component k in the fused image, we take $\widetilde{a}_k(\mathbf{x})$ and $\nabla\widetilde{\varphi}_k(\mathbf{x})$ directly from the input image for which the product $a_k(\mathbf{x})|\nabla\varphi_k(\mathbf{x})|^2$ is larger on a pixel-by-pixel basis. The fused image result is given in Fig. 6.7(c). An identical AM-based algorithm was used to obtain the result shown in Fig. 6.7(f) by fusing the CT image of Fig. 6.7(d) and the MR image of Fig. 6.7(e) (CT and MR images courtesy of `imagefusion.org` [58]).

Quantitative evaluations of the AM-based image fusion technique using three objective assessment metrics are shown in Tab. 6.1, Tab. 6.2, and Tab. 6.3. Here, I compare the performances of the AM-based image fusion against the well-known multi-scale image fusion techniques such as the Laplacian pyramid (LP) fusion and the wavelet fusion [69]. Best results are emphasized with bold-face numbers. Simulation results indicate that the AM-based image fusion technique generally does not perform as well as the Laplacian fusion technique and the wavelet fusion technique. The AM-based fusion only performs better than the other two techniques in a three cases. However, the current AM-based fusion scheme is pixel-wise and it does not exploit orientation information, which is an important feature

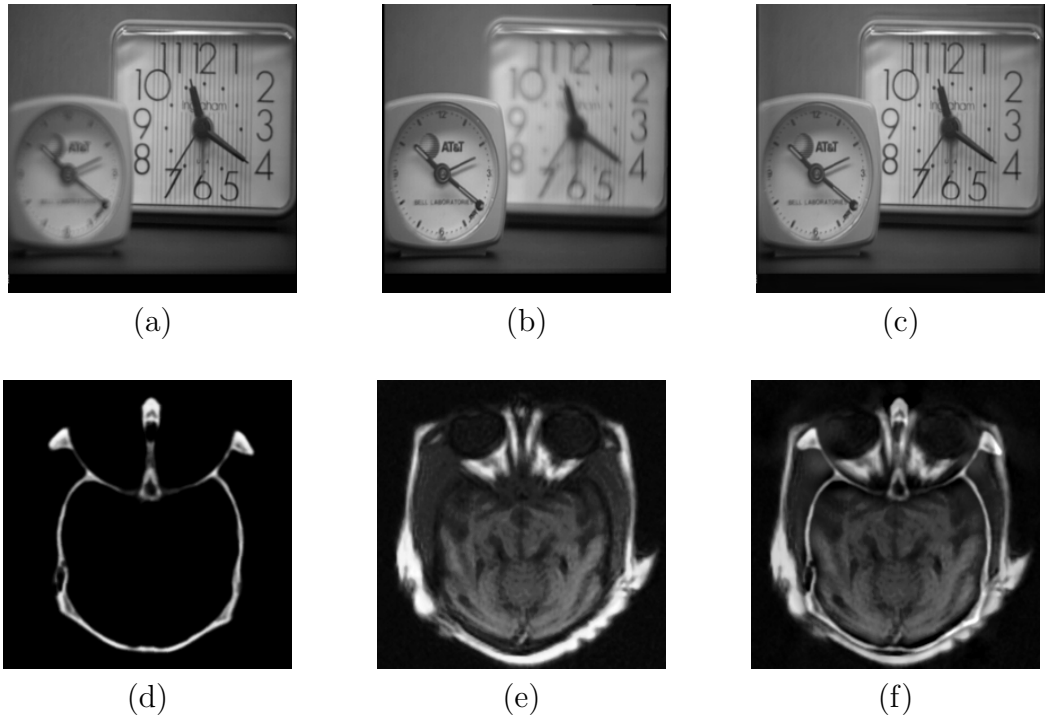


Figure 6.7: AM-based image fusion. (a) Clock A. (b) Clock B. (c) Fusion of (a) and (b). (d) CT image. (e) MR image. (f) Fusion of (d) and (e).

of the AM-FM image representation. My future work will be in the direction of incorporating orientation information into fusion algorithms.

Table 6.1: Objective performance of image fusion techniques measured by the mutual information metric [101].

	Wavelet	LP	AM-based
clock	0.577370	**0.630707**	0.610730
tiffany	0.669841	0.64423	**0.671626**
lena	**0.747307**	0.72831	0.704051
medical	**0.413435**	0.365575	0.305687
navigation	**0.242832**	0.25451	0.239581

Table 6.2: Objective performance of image fusion techniques using the objective pixel-level image fusion metric [128].

	Wavelet	LP	AM-based
clock	0.731238	**0.754339**	0.742081
tiffany	0.735087	**0.739462**	0.738636
lena	**0.741060**	0.739007	0.728825
medical	0.692751	**0.789651**	0.716454
navigation	0.613523	**0.697676**	0.640881

Table 6.3: Objective performance of image fusion techniques measured by the SSIM metric [127].

	Wavelet	LP	AM-based
clock	0.490198	**0.509316**	0.502113
tiffany	0.531843	**0.535119**	0.534410
lena	0.548370	**0.549710**	0.544959
medical	**0.301499**	0.279124	0.261661
navigation	0.198293	0.295320	**0.417149**

6.3 AM-FM Image Filtering

The AM-FM filters perform signal processing by operating on the computed AM $a(\mathbf{n})$ and FM functions $\nabla\varphi(\mathbf{n})$. A schematic diagram for the FM-based filtering

is illustrated in Fig. 6.8. In this filtering scheme, the FM function $\nabla\varphi_k(\mathbf{n})$ is modified by the filter $\mathcal{G}$, producing $\nabla\widetilde{\varphi}_k(\mathbf{n})$. The modified phase image, $\widetilde{\varphi}_k(\mathbf{n})$, is reconstructed from the filtered FM using the least squares integration approach discussed in Chapter 4. The AM may also be changed in general. The output image is then obtained as

$$\widetilde{f}_k(\mathbf{n}) = \widetilde{a}_k(\mathbf{n})\cos[\widetilde{\varphi}_k(\mathbf{n})]. \tag{6.7}$$

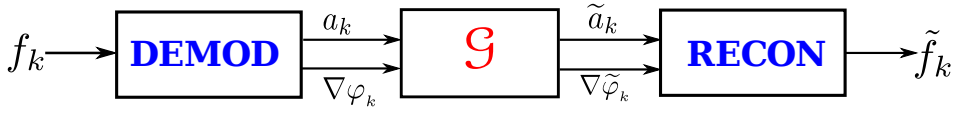


Figure 6.8: FM-based Image Filtering.

Let $\psi_k(\mathbf{n})$ be the reconstructed least-squares phase function. In Sec. 4.4.4, the least-squares phase $\psi_k(\mathbf{n})$ can be different from the true $\varphi(\mathbf{n})$. The reason is that $\psi_k(\mathbf{n})$ is obtained from a least-squares optimization approach. In other words, $\psi_k(\mathbf{n})$ is generally inconsistent with the principle phase values $\arccos[f_k(\mathbf{n})/a_k(\mathbf{n})]$ obtained from the original image components at some pixels. Sivley and Havlicek [109] defined the function $\rho_k(\mathbf{n})$ according to

$$\varphi_k(\mathbf{n}) = \psi_k(\mathbf{n}) + \rho_k(\mathbf{n}), \tag{6.8}$$

where the phase congruence term $\rho_k(\mathbf{n})$ is chosen to enforce agreement between the principle values of $\psi_k(\mathbf{n})$ and the principle values of $\varphi_k(\mathbf{n})$.

The phase model (6.8) is generally satisfactory for performing image synthesis on the discrete lattice $\mathbb{Z}^2$ – even after AM-based image filtering or elementary FM-based processing have been applied. However, the phase congruence term

$\rho_k(\mathbf{n})$ generally fails to be smooth and, consequently, the presence of $\rho_k(\mathbf{n})$ in (6.8) tends to introduce undesirable artifacts in the reconstructed image when sophisticated FM filtering or geometric transformations requiring interpolation to a new spatial sampling lattice are applied. These artifacts arise because the integrated phase $\varphi_k(\mathbf{n})$ contains jumps that are introduced to the phase by the phase congruence term $\rho_k(\mathbf{n})$. These discontinuities subsequently generate artifacts in the filtered output [88].

I ameliorate the phase congruence problem by removing $\rho_k(\mathbf{n})$ out of each term in (6.8) to define new generalized AM functions $A_{1k}(\mathbf{n})$ and $A_{2k}(\mathbf{n})$ through

$$\begin{aligned} f_k(\mathbf{n}) &= a_k(\mathbf{n})\cos[\varphi_k(\mathbf{n})] = a_k(\mathbf{n})\cos[\psi_k(\mathbf{n}) + \rho_k(\mathbf{n})] \\ &= \underbrace{a_k(\mathbf{n})\cos[\rho_k(\mathbf{n})]}_{A_{1k}(\mathbf{n})}\cos[\psi_k(\mathbf{n})] - \underbrace{a_k(\mathbf{n})\sin[\rho_k(\mathbf{n})]}_{A_{2k}(\mathbf{n})}\sin[\psi_k(\mathbf{n})] \\ &\equiv A_{1k}(\mathbf{n})\cos[\psi_k(\mathbf{n})] + A_{2k}(\mathbf{n})\sin[\psi_k(\mathbf{n})]. \end{aligned} \tag{6.9}$$

In order to define modulation domain signal processing operations capable of delivering filtered images that are free from undesirable phase reconstruction artifacts, AM-only processing should be applied to the (non-generalized) amplitude modulation functions $a_k(\mathbf{n})$. However, for joint AM-FM filtering, the generalized AM functions $A_{1k}(\mathbf{n})$ and $A_{2k}(\mathbf{n})$ should be processed. FM processing should be applied only to the generalized FM functions $\nabla\psi_k(\mathbf{n})$ and *not* to $\nabla\varphi_k(\mathbf{n})$. Subsequent to such processing, the generalized AM and FM functions can be interpolated to synthesize image samples on a modified sampling lattice as required to implement geometric image transformations. I collaborated with Adrian Campell, Johnathan Williams, and Murad Özaydin in developing this AM-FM filtering technique [88, 92]. This AM-FM filtering is a a new and largely unexplored area. For

the remainder of this chapter I will restrict my attention to only on class of AM-FM filters which I have investigated, the AM-FM geometric image transformations.

6.3.1 Translation

The discrete image $f(\mathbf{n})$ is modeled as $f(\mathbf{n}) = \sum_{k=1}^{N} a_k(\mathbf{n}) \cos[\varphi_k(\mathbf{n})]$. The filtered image translated by the displacement vector $\mathbf{u} = (u_0, v_0)$ is given by

$$f(\mathbf{n} - \mathbf{u}) = \sum_{k=1}^{N} a_k(\mathbf{n} - \mathbf{u}) \cos[\varphi_k(\mathbf{n} - \mathbf{u})], \tag{6.10}$$

where $\mathbf{u} \in \mathbb{R}^2$. Therefore, the modulation domain translation operation can be achieved by translating $a_k(\mathbf{n})$ and $\varphi_k(\mathbf{n})$ by $\mathbf{u}$, which generally yields samples that fail to lie on the discrete sampling lattice of the original image $f(\mathbf{n})$. Consequently, I apply bicubic interpolation to the resulting generalized AM and FM functions to synthesize new image samples on the translated sampling lattice.

An illustration of the image shifting is shown in Fig. 6.9. The original image Barbara is shown in Fig. 6.9(a). The image processing goal is to shift the original image by $\mathbf{u} = (14.01, 27.01)$ where $(0, 0)$ is the top left of the image. Fig. 6.9(b) is result of the shift operation in the pixel domain. The AM-FM result is given in Fig. 6.9(c). It is very difficult to spot any differences between the two approaches visually. Similar experiments are given for the fingerprint image in Fig. 6.9(d) with $\mathbf{u} = (34.70, 50.30)$. The outputs of the pixel domain approach and the AM-FM approach are shown in Fig. 6.9(e) and (f). The boat example has $\mathbf{u} = (24.50, 37.70)$.

6.3.2 Scaling

Here, scaling refers to magnification or "zoom". The modulation domain scaling operation is designed to admit positive real scaling factors. For example, if the

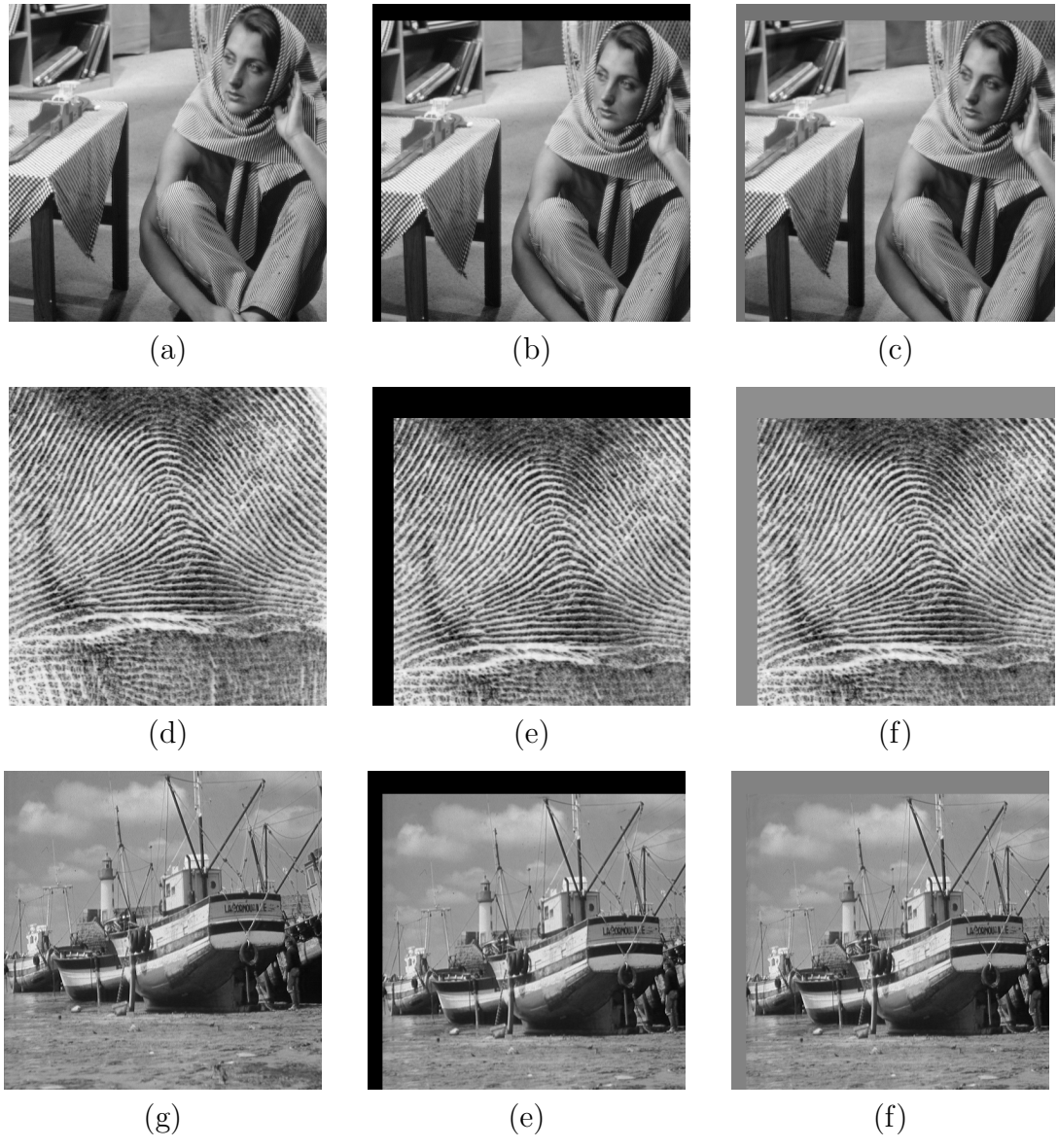


Figure 6.9: AM-FM image shift. (a) Original Barbara image. (b) Spatial shift. (c) AM-FM shift. (d) Original fingerprint image. (e) Spatial shift. (f) AM-FM shift. (g) Original boat image. (h) Spatial shift. (i) AM-FM shift.

image is enlarged by a factor of two, the magnitudes of the FM frequency vectors in the enlarged image must be decreased by two. Consequently, to produce consistent modulation domain image scaling operations, I design an AM-FM filtering scheme to accommodate changes of the FM function $\psi_k(\mathbf{n})$. The modulation domain scaling operation is depicted in Fig. 6.10. The AM signals $A_{ik}(\mathbf{n})$ are first up/down sample by a predefined positive real factor L and then are interpolated by either bilinear or bicubic interpolants. The FM signal $\nabla\psi_k(\mathbf{n})$ is also up/down sampled by the predefined factor L. This modified gradient signal is then compensated by $1/L$ in order to preserve the texture structure spacing and orientations. The modified gradient is then integrated to find the modified phase function $\widetilde{\psi}_k(\mathbf{n})$. As the processed gradient field $\nabla\widetilde{\psi}_k(\mathbf{n})$ may not be conservative, the modified phase is computed by performing the least squares phase unwrapping method proposed by Ghiglia and Romero [42]. The scaled output signal is then given as a summation of the filtered components $\widetilde{f}_k(\mathbf{n})$.

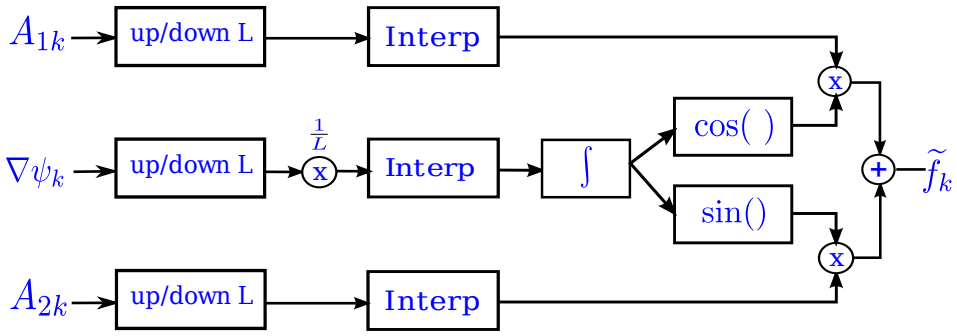


Figure 6.10: FM-based Image Scaling.

An illustration of AM-FM image magnification by a factor of two is shown in Fig. 6.11. The original image Barbara, Lena, and boat are shown in the first column in Fig. 6.11(a),(d),(g). The second column shows the spatial domain image

magnification using the standard bicubic interpolation technique. The last column depicts the AM-FM image magnification results.

While it is difficult to see the difference between the spatial domain technique and the proposed technique visually, a quantitative evaluation of these two approaches is given in Table. 6.4. The bicubic interpolation is a classical image magnification method [62]. I restrict the comparison to the upsampling operation because of the lack of ground truth data for the rotation and translation operations. I measure the performance of the two techniques in terms of the peak signal to noise ratio measure (PSNR) (4.31) and the perceptually motivated structural similarity index (SSIM) index [127]. For the upsampling operation, the proposed algorithm outperforms the classical bicubic interpolation in the PSNR and in the SSIM.

Table 6.4: Comparison of the upsampling operation.

	PSNR (dB)		SSIM	
	Bicubic	AM-FM	Bicubic	AM-FM
Boat	33.488	**33.697**	0.765	**0.799**
Barbara	32.091	**32.179**	0.716	**0.728**
Lena	35.018	**35.214**	0.852	**0.862**
Fingerprint	30.362	**30.488**	0.864	**0.869**

6.3.3 Image Rotation

A classical image rotation involves a rotation of the image grid and an interpolation scheme. In the modulation domain, a rotation on the image grid will also rotate the orientation of the FM function $\nabla\psi_k(\mathbf{n})$. In order to preserve the visually

Figure 6.11: AM-FM image 2x zoom. (a) Original barbara image. (b) Spatial zoom. (c) AM-FM zoom. (d) Original Lena image. (e) Spatial zoom. (f) AM-FM zoom. (g) Original boat image. (h) Spatial zoom. (i) AM-FM zoom.

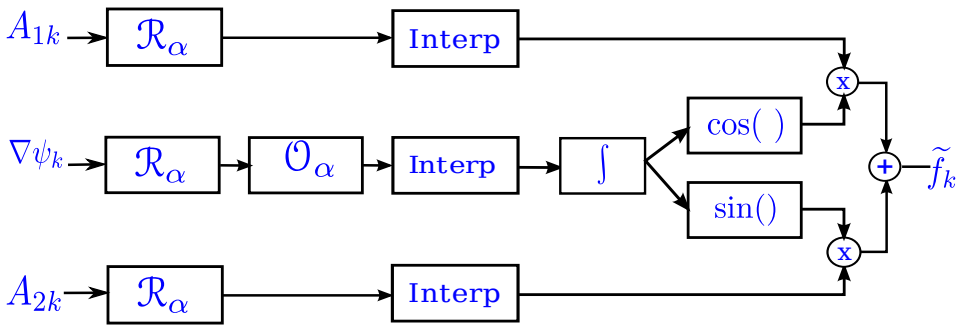


Figure 6.12: FM-based Image Rotation.

important texture structure and orientation, I implement the orientation change of the gradient field by multiplying with a rotation matrix $\mathcal{O}_\alpha$.

The modulation domain image rotation operation is depicted in Fig. 6.12. The rotation operator $\mathcal{R}_\alpha$ is first applied to the AM signals $A_{ik}(\mathbf{n})$. The rotated AM signals are then interpolated to find values lying on the pixel lattice. The rotation operator $\mathcal{R}_\alpha$ is also applied to the gradient field $\nabla\psi_k(\mathbf{n})$ and then multiplied with the rotation matrix $\mathcal{O}_\alpha$. Therefore, the rotation operation for the FM signal $\nabla\psi_k(\mathbf{n})$ is defined as $\mathcal{R}_\alpha\mathcal{O}_\alpha\nabla\psi_k(\mathbf{n})$. Similar to Section 6.3.2, the modified FM function $\nabla\widetilde{\psi}_k$ is then integrated to find the modified phase $\widetilde{\psi}_k$ by solving for the least squares solution of the phase unwrapping problem. Finally, the rotated output image is computed as a linear sum of the rotated components $\widetilde{f}_k(\mathbf{n})$.

In addition, for the rotation of the FM field, I showed that the counter rotation operator $\mathcal{O}_\alpha$ commutes with the lattice rotation operator $\mathcal{R}_\alpha$:

Theorem: $\mathcal{R}_\alpha\mathcal{O}_\alpha\nabla\psi_k(\mathbf{n}) = \mathcal{O}_\alpha\mathcal{R}_\alpha\nabla\psi_k(\mathbf{n})$.

Proof.

$$
\begin{aligned}
\mathcal{R}_\alpha \mathcal{O}_\alpha \nabla\psi(\mathbf{n}) &= \begin{bmatrix} \mathcal{R}_\alpha \mathcal{O}_\alpha U(\mathbf{n}) \\ \mathcal{R}_\alpha \mathcal{O}_\alpha V(\mathbf{n}) \end{bmatrix} \\
&= \mathcal{R}_\alpha \begin{bmatrix} \cos(\alpha)U(\mathbf{n}) + \sin(\alpha)V(\mathbf{n}) \\ -\sin(\alpha)U(\mathbf{n}) + \cos(\alpha)V(\mathbf{n}) \end{bmatrix} \\
&= \begin{bmatrix} \mathcal{R}_\alpha\Big(\cos(\alpha)U(\mathbf{n}) + \sin(\alpha)V(\mathbf{n}) \Big) \\ \mathcal{R}_\alpha\Big(-\sin(\alpha)U(\mathbf{n}) + \cos(\alpha)V(\mathbf{n}) \Big) \end{bmatrix} \\
&= \begin{bmatrix} \cos(\alpha)\mathcal{R}_\alpha U(\mathbf{n}) + \sin(\alpha)\mathcal{R}_\alpha V(\mathbf{n}) \\ -\sin(\alpha)\mathcal{R}_\alpha V(\mathbf{n}) + \sin(\alpha)\mathcal{R}_\alpha V(\mathbf{n}) \end{bmatrix} \\
&= \begin{bmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{bmatrix} \begin{bmatrix} \mathcal{R}_\alpha U(\mathbf{n}) \\ \mathcal{R}_\alpha V(\mathbf{n}) \end{bmatrix} \\
&= \mathcal{O}_\alpha \mathcal{R}_\alpha \nabla\psi(\mathbf{n})
\end{aligned}
$$

□

I illustrate the AM-FM image rotation in Fig. 6.13. Test images are Barbara in Fig. 6.13(a), boat in Fig. 6.13(d), and Lena in Fig. 6.13(g). The original Lena is rotated clockwise 27°. The pixel domain rotation of Lena is shown in Fig. 6.13(b) and the AM-FM rotation result is depicted in Fig. 6.13(c). Visually, the two results look the same. Similar to the scaling, quantitative evaluation is not given because the ground truth values are not available. Rotation results of the boat image and the Lena image are shown in Fig. 6.13(e),(f) and Fig. 6.13(h),(i), respectively.

6.4 Summary

For the first time, I introduced a systematic, high fidelity practical framework for AM-FM image processing. The AM-FM filters offer intuitive designs with fewer parameters and specifications compared to the traditional LSI filters. I provided two classes of AM-FM image filters, the AM-based and AM-FM filters.

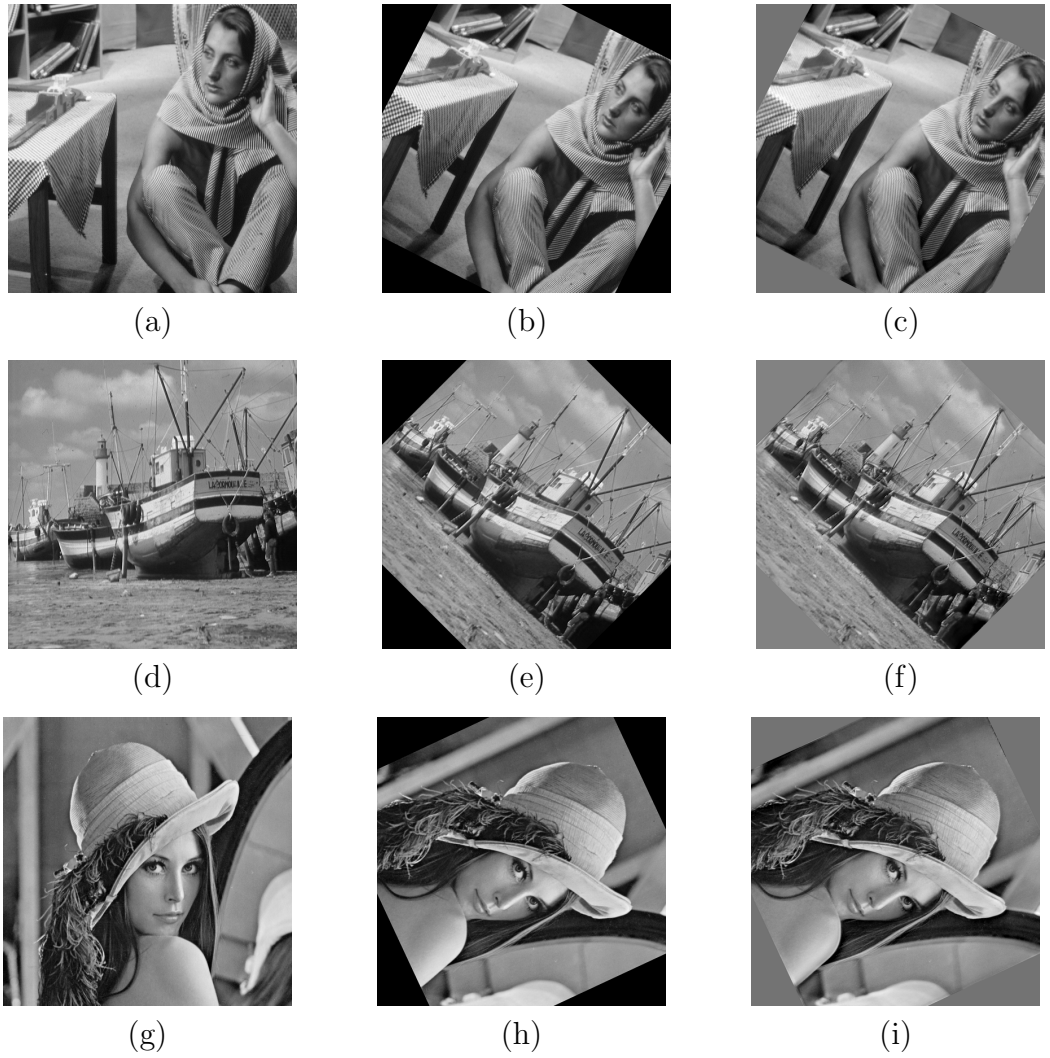


Figure 6.13: AM-FM image rotation. (a) Original Barbara. (b) Spatial domain rotation 27°. (c) AM-FM rotation 27°. (d) Original boat. (e) Spatial domain rotation 45°. (f) AM-FM rotation 45°. (d) Original Lena. (e) Spatial domain rotation 65°. (f) AM-FM rotation 65°.

I designed AM-based filters for selective orientation filtering, selective frequency filtering, image enhancement, and image fusion. In addition, I designed the AM-FM geometric image transformation including translation, rotation, and scaling. The experimental results indicate that the AM-based can achieve filtering results that are difficult or impossible to achieve with LSI filters, while the AM-FM image transformations deliver result with fidelity comparable to the pixel domain. For the image magnification operation, the AM-FM filter outperforms the LSI filter.

The proposed AM-FM filters are important as they establish the foundation for future research in high fidelity FM image processing research. The design of filters in the AM-FM framework is intuitive and requires fewer parameters than traditional filter design methods. In addition, the framework produces results that correspond well with human visual perception.

Chapter 7

Extended Results of the xAMFM

In this chapter, I described three extensions of the xAMFM for image analysis applications. These extensions are discussed in a separate chapter because they do not fit in the modulation domain image filtering framework. They are analysis applications and do not require the AM-FM synthesis transform. Therefore, I discussed each extension separately. In particular, I first motivate the problem, then discuss the use of the xAMFM for that problem, and finally evaluate the effectiveness of the proposed approach with experiments. Readers will see some equations and concepts discussed in previous chapters.

7.1 Coherent Texture Decomposition

Decomposing a complicated signal into perceptually meaningful components is an important problem that has received increasing attention recently [4, 74, 113, 125]. Well-known examples include the cocktail party speech separation problem and image restoration from multiple sub-image sources. Here, I am interested in the image decomposition problem where a texture image is broken into multiple *visually meaningful* components, e.g., simple and locally coherent constituents. Unfortunately, such decomposition is an ill-posed inverse problem [4, 113]. Starck, et al. [113] illustrated that a K component image of N pixels will require $N \times K$ unknowns to be solved. Therefore, prior knowledge of the signal components should

be incorporated, e.g., signal statistics, image models, and sparsity.

Early approaches used multiresolution techniques to describe the image as a sum of localized parts. Daugman [28] proposed a human visual system (HVS) inspired Gabor filterbank to decompose an image into smooth and localized components. The computed components are band-pass and orientation selective. Alternatively, Simoncelli and Freeman [38] introduced the steerable pyramid for image analysis. The steerable pyramid decomposes an image into multiple scales and multiple orientations. In the past two decades, wavelets have been used extensively in denoising and compression applications. Wavelets offer good time-frequency localization and a compact representation [73]. While these techniques are essential in many image processing applications, they lack specific models for coherent texture components. Consequently, the decomposed components obtained by these analysis techniques frequently fail to correspond well with human visual perception of the image.

Recently, Meyer [84] pioneered a nonlinear partial differential equation (PDE) approach to image decomposition. In this formulation, an image is broken into a sum of two parts: a cartoon part and a texture part. The cartoon describes a homogeneous region with sharp boundaries, and is modeled by a bounded variational function. The texture part is modeled by certain energy norms. Both of these components are computed simultaneously by a total variation minimization framework [4, 125]. Similar to Meyer's cartoon and texture decomposition idea, Starck, et al. [113] formulated an image decomposition using a combination of basis pursuit denoising and total-variation regularization. They used two optimized and sparse dictionaries, one for the cartoon and one for the texture, to extract image components. Even though the two component image decomposition

model delivers meaningful results, this approach does not generate locally coherent components.

In contrast to the two component image model approach, cartoon and texture, this dissertation has developed techniques to decompose images into sums of multiple locally narrow-band components. Each component is represented with smoothly varying amplitude modulation (AM) and frequency modulation (FM) functions, thereby explicitly computing AM-FM image models in the modulation domain. A Kalman filtering framework was developed in [48] to track texture multi-components spatially across the channels of a Gabor filterbank and extract them. While this approach did not prove sufficiently robust to enable reliable analysis of general images, it should be noted that extended Kalman filtering was applied successfully in [72] to track multicomponent amplitude and frequency modulations temporally in human speech. The spatially adaptive Kalman filters of [48] were replaced by a static global decomposition into components based on the filterbank structure in [49], which led to a robust and readily computable multicomponent image model. However, such decomposition precluded the possibility of perfect reconstruction and produced components that were less strongly connected to human visual perception. Evangelopoulos and Maragos [31] also used the modulation domain model for image decomposition, but their approach was limited to the two component decomposition model, cartoon and texture.

In this section, I propose a novel iterative algorithm for decomposing a texture image into homogeneous textural patches that are locally coherent and visually meaningful. I introduce a new quantitative modulation domain coherency measure. The components are iteratively extracted by a greedy algorithm that is similar to matching pursuit [74]. The experimental results show that the extracted

components are locally coherent and agree well with human perception.

7.1.1 Modulation Domain Texture Decomposition

Let $f(\mathbf{x})$ be an image defined on $\mathbb{R}^2$. The modulation domain image model for $f(\mathbf{x})$ is given by

$$f(\mathbf{x}) = \sum_{k=1}^{K} f_k(\mathbf{x}) = \sum_{k=1}^{K} a(\mathbf{x}) \cos[\varphi(\mathbf{x})]. \tag{7.1}$$

Let $m, n \in \mathbb{N}$ and let $\mathbf{I}(m,n)$ contain the samples of the continuous image $f(\mathbf{x})$ in (7.1). Let $\mathbf{I}_k(m,n)$ contain the samples of component $t_k(\mathbf{x})$ in (7.1). Then $\mathbf{I}_k(m,n)$ admits a modulation domain representation $\Gamma_k = [A_k(m,n)\ R_k(m,n)\ \theta_k(m,n)]$ which may be computed using the spline-based demodulation framework given in [109] or in Chapter. 4. $R_k(m,n)$ and $\arg\nabla\varphi_k(m,n)$ are the polar representation of the FM function, and they are computed as $R_k(m,n) = |\nabla\varphi_k(m,n)|$ and $\theta_k(m,n) = \arg\nabla\varphi_k(m,n)$. The overall K component image $\mathbf{I}(m,n)$ is described in the modulation domain by the multicomponent representation $\mathbf{\Gamma} = [\Gamma_1\ \Gamma_2\ \ldots\ \Gamma_k]^T$ obtained by concatenating the representations of the individual components. The vector $\mathbf{\Gamma}$ is then used as a dictionary for matching orientations as the dominant texture components are extracted.

Let A_k, R_k, and θ_k be the be the AM, FM magnitude, and FM orientation functions computed for component $I_k(m,n)$. I computed AM-weighted FM functions

$$\Theta(m,n) = \sum_{k=1}^{K} A_k(m,n)\theta_k(m,n), \tag{7.2}$$

where $\Theta(m,n)$ is the dominant orientation. I illustrate the dominant orientation estimation process in Fig. 7.1. The dominant FM field of the woven brass image is shown in Fig. 7.1(b). From this dominant FM field, a dominant orientation

is estimated as in Fig. 7.1(c). The second dominant orientation is depicted in Fig. 7.1(d).

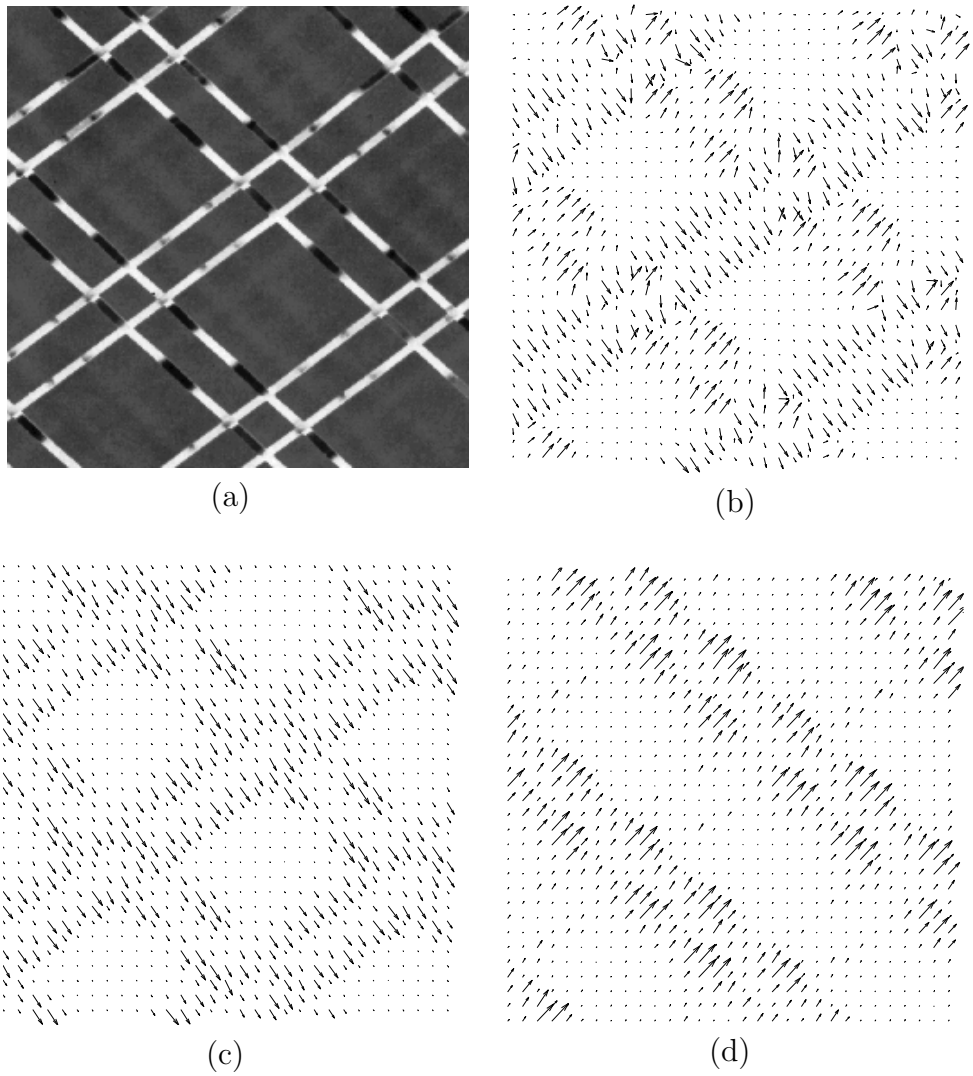


Figure 7.1: Dominant orientation estimation of the woven brass image. (a) Woven brass image. (b) Dominant FM field. (c) 1^{st} dominant orientation. (d) 2^{nd} dominant orientation.

The dominant orientation (7.2) is matched against the overcomplete dictionary $\mathbf{\Gamma}$ to assign a weight to the AM and FM functions in each Γ_k on a pixel-by-pixel

basis. The dominant texture component is then defined by the linear combination

$$C_d(m,n) = \sum_{k=1}^{K} \alpha_k(m,n) A_k(m,n) \cos[\varphi_k(m,n)], \tag{7.3}$$

where the coherency measure $\alpha_k(m,n)$ is defined by

$$\alpha_k(m,n) = \begin{cases} 1, & \delta_k(m,n) \leq \frac{\pi}{12}, \\ 1 - \{\frac{3}{\pi}[\delta_k(m,n) - \frac{\pi}{12}]\}^2, & \frac{\pi}{12} < \delta_k(m,n) \leq \frac{\pi}{6}, \\ 0, & \text{otherwise} \end{cases} \tag{7.4}$$

and $\delta_k(m,n) = |\theta_k(m,n) - \Theta(m,n)|$. Intuitively, the coherency measure α_k works in local spatial neighborhoods to group together and blend components $\mathbf{I}_k(m,n)$ having FM orientations that are *close* to the dominant orientation (7.2) at each pixel. The blended sum (7.3) then constitutes a coherent, textural component that is extracted from the image and the process is repeated iteratively until all coherent dominant orientations have been extracted from the image. The threshold bandwidth $\pi/6$ in (7.4) was chosen for agreement with the eight orientations present per level in the adapted steerable pyramid.

Algorithm 3 AM-FM Texture Decomposition Algorithm

$\mathbf{L} \leftarrow$ original image $\mathbf{I}(m,n)$
domOrien $\leftarrow \emptyset$
while $1 = 1$ **do**
1. Compute feature vector $\mathbf{\Gamma}$ for image $\mathbf{L}$
   using demodulation algorithm in Chapter. 4
2. Estimate the dominant orientation $\Theta(m,n)$ using (7.2).
   if ($\Theta \in$ **domOrien**) **break**;
3. Extract texture component C_d using (7.3) and (7.4).
4. $L \leftarrow L - C_d$
5. **domOrien** $\leftarrow$ [**domOrien** Θ]

end while

The pseudo code for AM-FM texture decomposition is given in Algorithm 7.1.1. The proposed algorithm can be interpreted as a frequency-based feature extraction technique. It is, however, different from other traditional techniques such as

Fourier and orientation-selective Gabor filtering. Our algorithm is capable of finding the dominant texture orientation of a component and extract it by matching its orientation against an overcomplete feature dictionary. Other frequency-based techniques rely on an energy-based analysis step to estimate texture orientations, and subsequently require one to design filters with appropriate bandwidths and orientations for the feature extraction process. Both of these estimation steps are sensitive to errors if the image structures and orientations are not known a priori.

7.1.2 Results and Discussion

I evaluated the algorithm described in Algorithm 7.1.1 against a variety of Brodatz textures. For each extracted component, we computed the modulation domain feature dictionary $\mathbf{\Gamma}$. The dominant texture orientation Θ was estimated as in (7.2). The extracted component $C_d(m, n)$ was then found by matching the dominant orientation $\Theta(m, n)$ against the dictionary $\mathbf{\Gamma}$ as described in (7.3) and (7.4). The results are illustrated in Fig. 7.2 and Fig. 7.3. Each column depicts the decomposition result of a test image. In each column, the first row contains original images. The second and third row show the first and second dominant components and the last row shows the residual after extraction of the first two dominant components. Images in the second, third, and fourth rows of each column are contrast stretched together and directly comparable in terms of gray scales.

The original burlap image is shown in Fig. 7.2(a). It can be interpreted as a two component image with horizontal and vertical stripes. The algorithm produces two dominant components shown in Fig. 7.2(b)-(c). These components are locally coherent and coincide with human perception. The lack of organized texture in the residual image in Fig. 7.2(d) demonstrates that the horizontal and

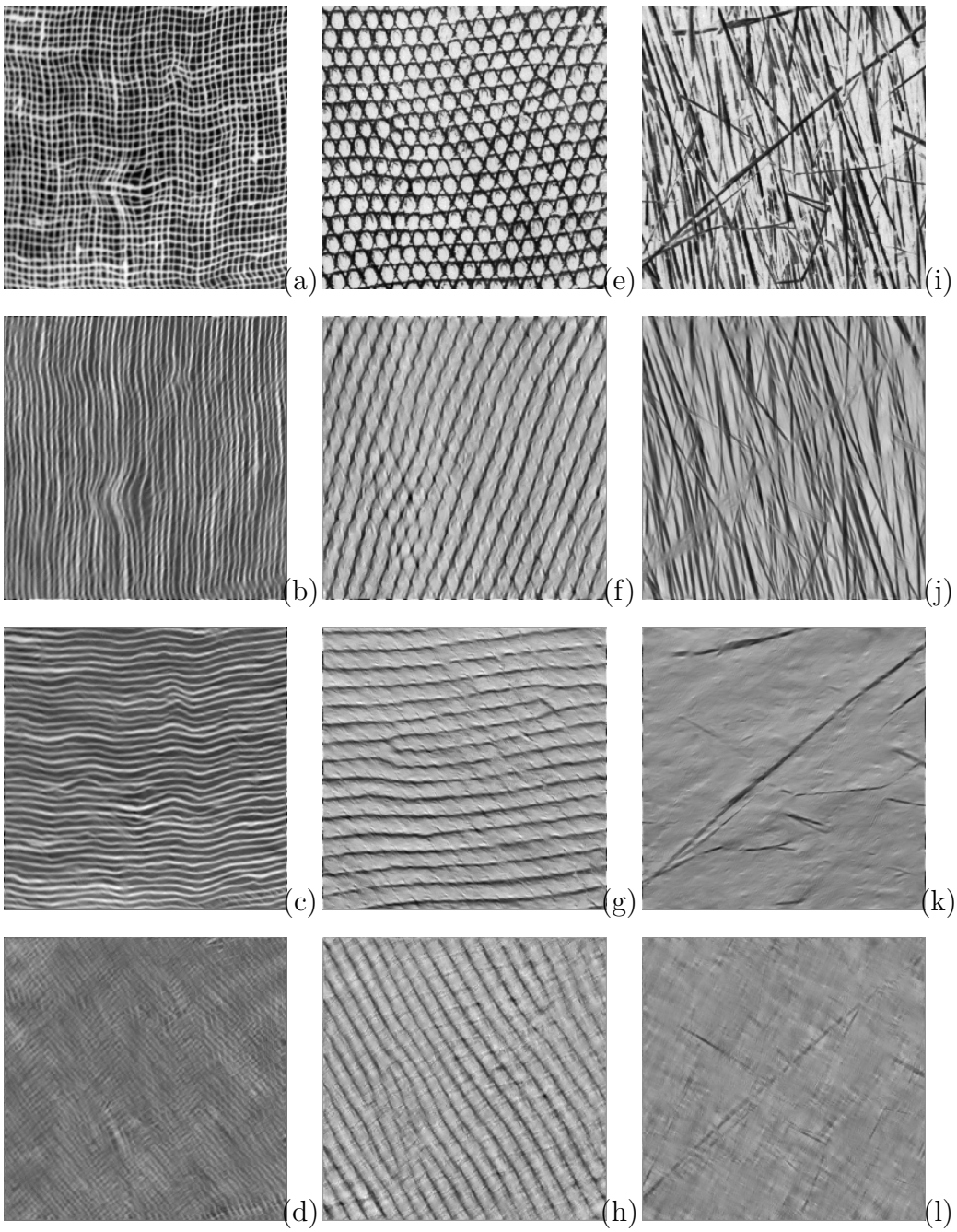


Figure 7.2: Examples. (a) Original burlap image. (b) First component of burlap. (c) Second component of burlap. (d) Residual of burlap. (e) Original reptile skin image. (f) First component of reptile skin. (g) Second component of reptile skin. (h) Residual of reptile skin. (i) Original straw image. (j) First component of straw. (k) Second component of straw. (l) Residual of straw.

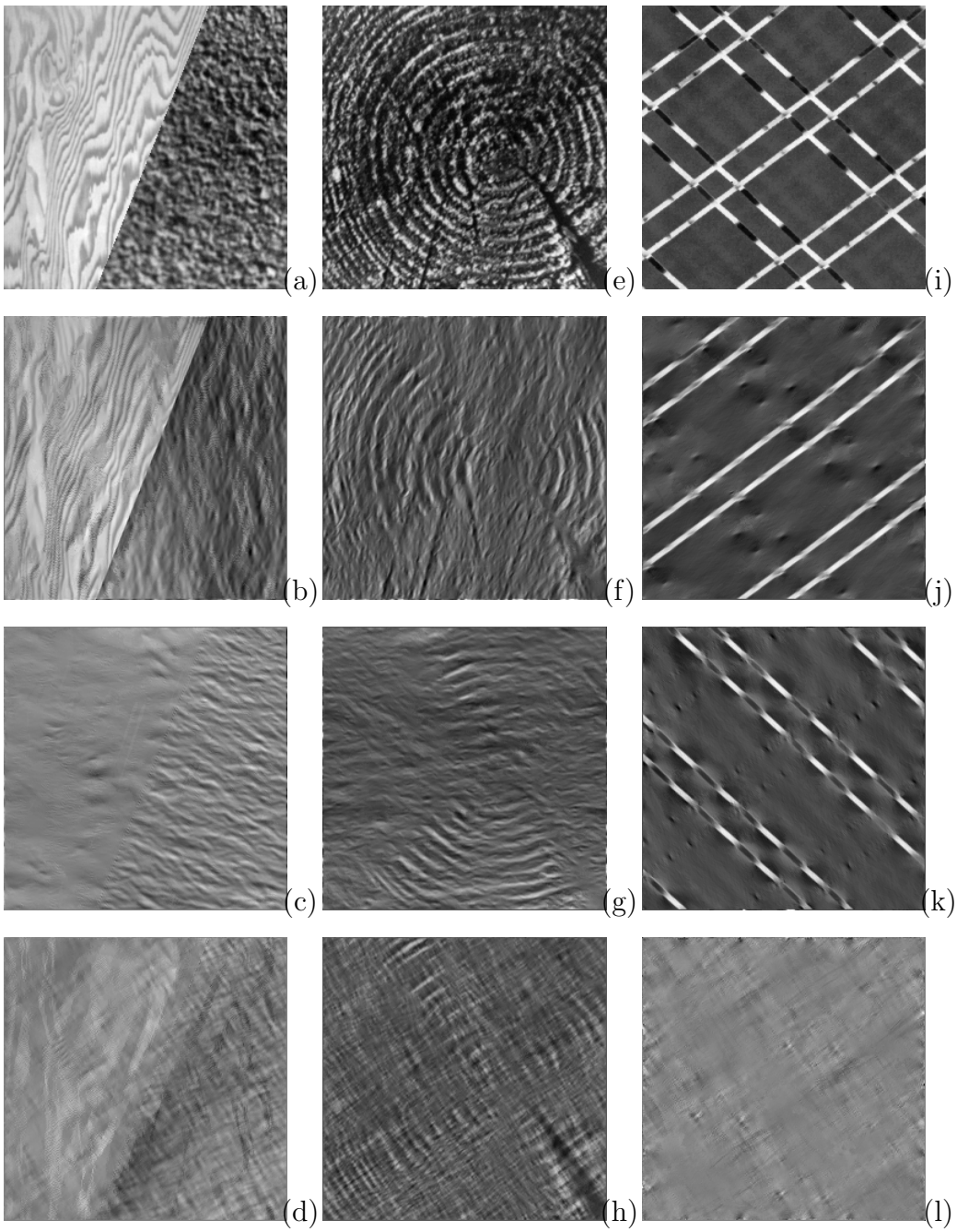


Figure 7.3: Examples. (a) Original wood/paper image. (b) First component of wood/paper. (c) Second component of wood/paper. (d) Residual of wood/paper. (e) Original tree image. (f) First component of tree. (g) Second component of tree. (h) Residual of tree. (i) Original cloth image. (j) First component of cloth. (k) Second component of cloth. (l) Residual of cloth.

vertical stripes of the original image are successfully extracted.

A more complex decomposition example of reptile skin texture is shown in Fig. 7.2(e). The two dominant texture components are shown in Fig. 7.2(f)-(g). The algorithm is capable of decomposing the original image into three locally coherent image components. It is interesting to observe that these components are not easily identified from the original image. Similar perceptually motivated results can be seen for the straw image in Fig. 7.2(i). Straws having vertical orientations are successfully grouped into the first coherent component in Fig. 7.2(j). The second extracted component in Fig. 7.2(k) depicts remaining textures that are orthogonal to the first.

Fig. 7.3(a) shows a composite image which contains a wood grain texture on the left and a uniform grainy texture on the right. The decomposed texture components are shown in Fig. 7.3(b)-(c). The first component in Fig. 7.3(b) is able to capture the wood grain texture part of the original image, while the second component in Fig. 7.3(c) extracts the remaining texture. This example suggests that better decomposition results can be achieved if the original image is segmented into homogeneous regions prior to applying this algorithm.

The proposed algorithm, however, is not effective for circularly symmetric texture images like the tree image in Fig. 7.3(e). Although the extracted components in Fig. 7.3(f) and Fig. 7.3(g) are locally coherent, they do not agree with human perception which interprets the image with one circular component. Such limitation can be explained by the range restriction imposed on the angular bandwidth threshold of the coherency measure in (7.4). The angular bandwidth was set to $\pi/6$, which is not wide enough to handle circular textural patterns.

7.1.3 Summary

In this section, I proposed an iterative texture analysis algorithm capable of extracting locally coherent and meaningful components from textural images. I introduced a new quantitative coherency measure in the modulation domain for image components. The effectiveness of the proposed algorithm is demonstrated with a variety of well-known Brodatz textures. The decomposed image components are visually motivated and their interpretations coincide with human perception.

The proposed algorithm, however, is limited to texture images with components having limited orientation bandwidth. The future work will be focused on fine-tuning the coherency measure in (7.4) to accommodate circularly symmetric textural patterns and apply the algorithm to nature images.

7.2 Cartoon + Texture Decomposition

In Section 7.1, the goal is to decompose an image into coherent components. The coherent components must be locally smooth varying. Here, I consider the problem of decomposing an image into a structural component and textural component. The structural portion, which is referred to as *cartoon*, carries broad information about an image and is usually piece-wise smooth curves. The texture component, subsequently referred to as *texture*, describes oscillating patterns of image textures and noise [18, 113]. For example, in a striped T-shirt, the *cartoon* consists of lines in the borders of the shirt and the *texture* are the stripes within the shirt.

A successful texture-cartoon decomposition can lead to improvements for subsequent image processing operations such as compression, edge detection, and image inpainting. For instance, higher overall compression gain can be obtained

by decomposing the image into different types of signals and designing optimized encoders to compress these type of signals separately [83]. The cartoon-texture decomposition can eliminate extraneous edges that can partially due to noise or unimportant objects [18, 113]. Such decomposition can retain important edges in image denoising [2, 102] or improve inpainting results with better texture replication [7].

The texture-cartoon decomposition is, however, an ill-posed problem. As texture and cartoon are loosely defined, image features can be considered as texture in one scale, but they can be cartoon at another scale. For example, when we look at a tree at a far distance, leaves can be classified as textures. However, one can consider these leaves as cartoon at a closer viewing distance. In addition, human judgement can also play an important role in deciding whether an object is part of texture or not.

Most works in the texture-cartoon decomposition are in the partial differential equation (PDE) setting. The texture and cartoon are modeled to lie in different functional spaces. The solution is found by solving a convex regularized optimization problem [4, 84, 102, 125]. The quality of texture and cartoon decomposition depends on signal models used to describe them and the regularization parameter. Despite approaches to find suitable values for the regularization parameter [4, 104], the cartoon edges often bleed into the texture components.

Meyer, Averbuch, and Coifman [83] proposed an image compression scheme where an image is decomposed into multi-layered components such as texture and cartoon. The authors used a suitable basis for each each layer of signal in order to increase compression gain. Stark, Elad, and Donoho drew ideas from [83] and [125] to create a hybrid approach that used total variation regularization and

basis matching. They designed two dictionaries, each of which contains basis functions that are *tuned* for either cartoon or texture. The texture and cartoon were subsequently extracted by projecting the image onto these basis functions.

Apart from the PDE and basis representation approaches, Buades et al. [18] proposed a nonlinear texture-cartoon decomposition. They observed that the total variation of texture and cartoon features behave differently before and after a lowpass filtering. A weight assignment scheme were then used to classify texture and cartoon features. While the algorithm produced good texture and cartoon separation, the results depended on the bandwidth parameter of the lowpass filter.

In this section, inspired by the work of Buades et al. [18], I proposed an automatic nonlinear texture-cartoon decomposition algorithm. In particular, I measured the ratio of gradient magnitude across modulation domain components and used this ratio to determine the component where the change between cartoon and texture of a pixel is most likely to occur. Once the component is determined, we used a hard threshold strategy to classify texture and cartoon pixels to obtain a weight matrix. The texture component is then obtained by multiplying the original image with the weight matrix. The cartoon is the difference between the original image and the texture component.

7.2.1 Background

Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a continuous image. Let $u : \mathbb{R}^2 \to \mathbb{R}$ be the cartoon component. Let $v : \mathbb{R}^2 \to \mathbb{R}$ be the texture component. The cartoon-texture decomposition aims to extract u and v as

$$f = u + v. \tag{7.5}$$

Total variation regularization approaches

Most of the texture-cartoon decomposition approaches are formulated in the partial differential equation setting, we will give a short description of the underlying models. Readers can refer to [4, 102, 125] for deeper analysis.

Rudin, Osher, and Fatemi [102] solved (7.5) in the context of a denoising problem. They assumed that the cartoon u belongs to a class of bounded variation (BV) functions and the texture v is a finite energy function. Both u and v are solved simultaneously in the convex minimization setup

$$\underset{u\in \mathrm{BV}, v\in \mathrm{L}^2}{\arg\min} \left(\int |Du| + \lambda ||v||^2_{\mathrm{L}^2} \right), \tag{7.6}$$

where λ is a positive tuning parameter and the integral of Du measures the total variation of signal u. The computed texture v, however, contains cartoon edges. Aliney [2] proposed an L^1 model for the texture component v in (7.6) to capture the salt and pepper noise property of corrupted signals. Meyer [84] provided an alternative model for the texture component in (7.6). Instead of being L^1 or L^2, v belongs to functions in a Banach space featured by a G-norm model which allows features to have high oscillation but can still retain low energy norm. Many successful texture-cartoon decomposition algorithms have been derived from the Meyer formulation, *e.g.*, [4, 125].

The Linearized Meyer Model

Buades et al. [18] observed that a linearized version of the original Meyer model is indeed the classical highpass-lowpass filtering problem. Let K_σ be a lowpass filter; the texture-cartoon decomposition problem can be viewed as a problem of

designing a suitable lowpass filter K_σ to capture u and v as

$$\begin{aligned} u &= K_\sigma * f \\ v &= f - K_\sigma * f, \end{aligned} \tag{7.7}$$

where σ is the scale parameter that determines the filter bandwidth and $*$ denotes the convolution operator.

Intuitively, the bandwidth parameter σ controls the amount of high frequency features that will be retained. Therefore, this model can not separate the texture and cartoon features when their frequencies are overlapped.

Nonlinear texture-cartoon classification

Buades et al. [18] observed that the local total variation (TV) of texture features and cartoon features behave differently when filtered by a lowpass filter K_σ. The ratio of local TV before and after the lowpass filter K_σ is applied tends to be lower in the texture region than that in the cartoon region. Based on this observation, the authors used a nonlinear mapping similar to soft-thresholding to classify pixels into the two categories.

Even though the decomposition algorithm does not compute solutions that converge to those of the TV regularization approaches [18], it produces good quality texture-cartoon separation with a non-iterative implementation. The solutions of this method, however, depend on the selection of the bandwidth σ of the lowpass filter K_σ. Without a properly tuned σ, the solutions can change drastically.

7.2.2 Texture-Cartoon Decomposition

I represented the image f as a sum of K non-stationary amplitude modulation (AM) functions and frequency modulation (FM) functions

$$f = \sum_{k=1}^{K} f_k = \sum_{k=1}^{K} a_k \cos(\varphi_k), \tag{7.8}$$

where $a_k : \mathbb{R}^2 \to \mathbb{R}^+$ is the AM function and $\varphi_k : \mathbb{R}^2 \to \mathbb{R}$ is the phase modulation function [49]. Both a_k and φ_k are assumed to be locally smooth. The FM functions are given by the gradient of φ_k, *i.e.*, $\nabla\varphi_k = [\varphi_{kx}\, \varphi_{ky}]^T$, where the second subscript denotes partial differentiation. The discrete AM and FM functions are computed using the demodulation algorithm in [91, 108]. I arranged the K AM-FM components in ascending order based on the magnitude of the FM vector in (7.8), *i.e.*, f_1 carries low-frequency components and f_K contains high-frequency components.

The key ingredient the cartoon-texture separation in [18] as well as in this paper lies in the computation of image gradient. For 1D AM-FM signal representation, the derivative of component f_k is obtained as

$$f_k' = a_k' \cos(\varphi_k) - \varphi_k' a_k \sin(\varphi_k). \tag{7.9}$$

I performed an approximation to (7.9) to make it more robust to noise. Since the AM function a_k is locally smooth, we estimated the 1D derivative of f_k in (7.9) as

$$f_k' \approx -\varphi_k' a_k \sin(\varphi_k). \tag{7.10}$$

Extended the 1D derivative in (7.10) to 2D, we computed a metric T_ℓ to quantify the gradient magnitude of the first ℓ AM-FM components according to

$$T_\ell \approx \sqrt{\left(\sum_{k=1}^{\ell} \varphi_{kx} a_k \sin(\varphi_k)\right)^2 + \left(\sum_{k=1}^{\ell} \varphi_{ky} a_k \sin(\varphi_k)\right)^2} \tag{7.11}$$

In (7.11), T_1 is the approximated gradient magnitude of the lowest frequency component while T_K is the approximated gradient magnitude of the image f.

Similar to Buades et al. [18], I defined the gradient magnitude ratio at every pixel as

$$D_\ell = \frac{T_K - T_\ell}{T_K}, \tag{7.12}$$

where $1 \leq \ell \leq K$. At a pixel (m, n) in the image grid, $D_\ell(m, n)$ measures the relative difference between the gradient magnitude of the whole image and the gradient magnitude of the first ℓ components. D_ℓ is maximum when $\ell = 1$ and decreases monotonically towards 0 as ℓ increases.

I defined f_β to be the AM-FM component where the change between texture and cartoon is likely to happen at each pixel. I first created a mask M such that $M(m, n) = \alpha$ if $D_\alpha > 0.25$ and $M(m, n) = 0$ otherwise. The index β is then estimated according to

$$\beta = \text{median}(M). \tag{7.13}$$

Finally, I applied a hard threshold strategy to create a weight matrix w where a weight of one means texture and a weight of zero means cartoon

$$w = \begin{cases} 1, & \text{if } T_\beta \geq 0.25, \\ 0, & \text{if } T_\beta < 0.25. \end{cases} \tag{7.14}$$

The cartoon and texture are then computed as

$$\begin{aligned} u &= w \cdot f, \\ v &= f - u. \end{aligned} \tag{7.15}$$

7.2.3 Simulation Results

I tested the proposed algorithm on the Kodak image dataset and standard test images. The results are shown in Fig. 7.4. For each test image, the texture-

cartoon decomposition results are demonstrated by row. The original image is in the left column, the cartoon component u is in the middle column, and the texture component u is on the right. Fig. 7.4(b) and Fig. 7.4(c) show the cartoon and the texture component of the image *kodim08*.In these figures, overall structure of the image is retained in the cartoon, while the fine textures in the roofs and windows are extracted into the texture. Fig. 7.4(h) and Fig. 7.4(i) depict the cartoon and the texture component of the *fingerprint* image. The algorithm is able to extract most of repeating curves in the original image and put into the texture component. The cartoon contains mostly low-frequency residual. Fig. 7.4(k) and Fig. 7.4(l) illustrates the cartoon and the texture component of the *Babara* image. Oscillating patterns in her pant, shirt, and in the table are successfully extracted to the texture. The edges in her hands and table are still kept in the overall structure of the image.

7.2.4 Summary

I proposed an automatic nonlinear texture-cartoon decomposition based on the frequency behavior of texture and cartoon across different scales. I measured the ratio of gradient magnitude across modulation domain components and use this ratio to classify the texture and cartoon pixels. The simulation results demonstrated that the proposed algorithm is able to extract texture and cartoon components from images efficiently. While this work followed a similar path as Buades et al. [18], the results do not depend on the lowpass filter bandwidth which is critical to the separation process. Currently, I set the threshold parameter in the hard threshold process empirically to 0.25. I am investigating machine learning techniques to overcome this limitation.

Figure 7.4: Texture Cartoon Decomposition Examples. (a) Original kodim08 from Kodak. (b) Original kodim05 from Kodak. (c) Cartoon component of (a). (d) Cartoon component of (b). (e) Texture component of (a). (f) Texture component of (b).

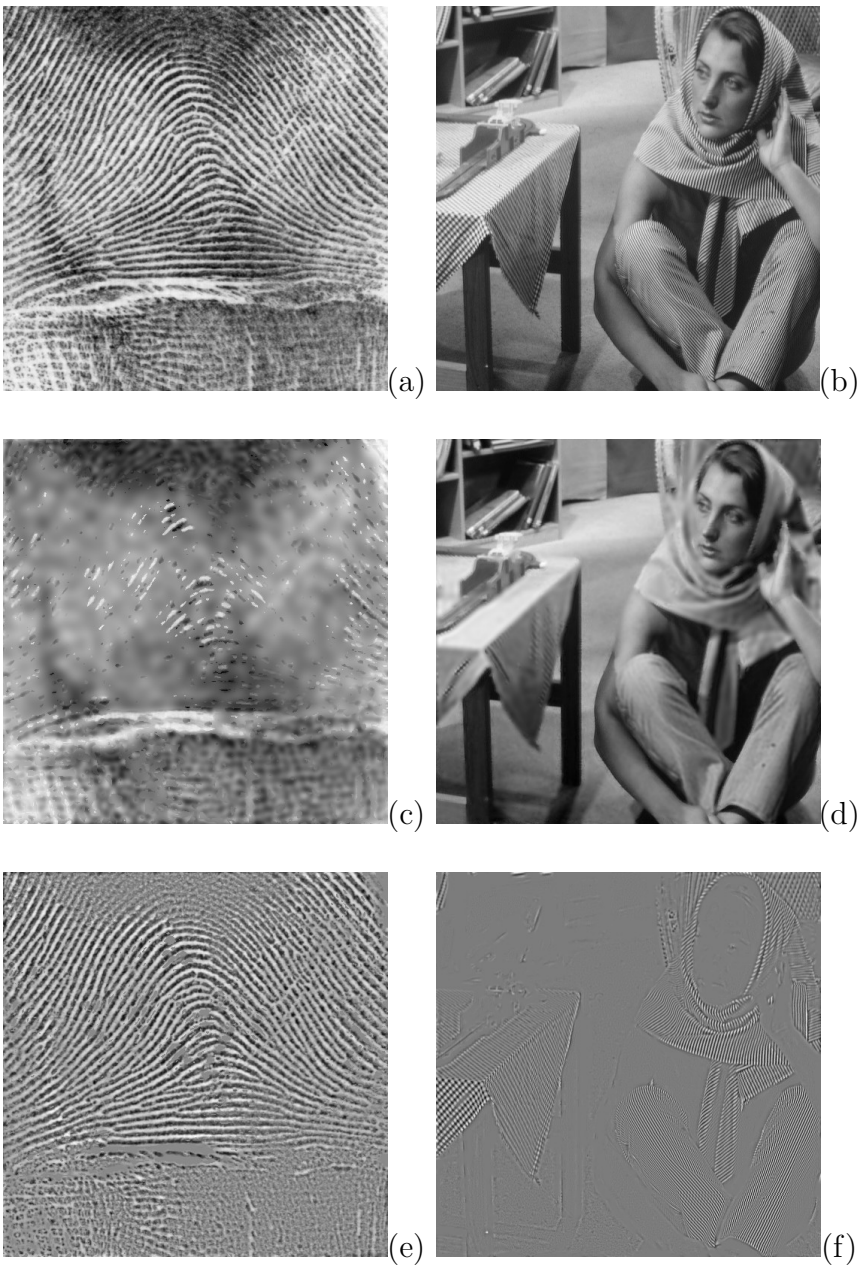


Figure 7.5: Texture Cartoon Decomposition Examples. (a) Original *fingerprint*. (b) Original *Barbara*. (c) Cartoon component of (a). (d) Cartoon component of (b). (e) Texture component of (a). (f) Texture component of (b).

7.3 Relationship with the Monogenic Signal

Fourier analysis is an important tool for analyzing and representing a stationary signal as a sum of pure sinusoids having constant amplitudes and constant frequencies. However, in practice many important signals are *nonstationary* in the sense that the amplitude envelope and frequency content are time varying (or space varying). Obvious examples include seismic survey data, radar returns, a variety of communication signals, and many biomedical signals [9]. The AM-FM model

$$f(\mathbf{x}) = a(\mathbf{x}) \cos[\varphi(\mathbf{x})] \tag{7.16}$$

has been widely used to represent such signals, where, for the 2D case we are concerned with in this paper, $a(\mathbf{x}) : \mathbb{R}^2 \rightarrow \mathbb{R}^+$ is a slowly varying non-negative amplitude modulation (AM) function and $\varphi(\mathbf{x}) : \mathbb{R}^2 \rightarrow \mathbb{R}$ is the phase modulation function. The frequency modulation (FM) function $\nabla\varphi(\mathbf{x})$ carries a rich local description of the surface pattern orientation and granularity.

The AM-FM model (7.16) is ill-posed in the sense that infinitely many pairs of $a(\mathbf{x})$ and $\varphi(\mathbf{x})$ exist which satisfy the equality (7.16). In his seminal 1D paper [41], Gabor used the Hilbert transform (HT) to disambiguate the AM-FM modeling problem by constructing a complex signal extension called the *analytic signal*. In arbitrary dimensions, any given complex extension associates unique AM and FM functions with a real signal which may be obtained directly by taking the magnitude of the complex signal and by differentiating the argument of the complex signal. More recently, Huang et al. developed the data adaptive Hilbert-Huang transform (HHT) technique to iteratively compute multicomponent AM-FM models. At each iteration, the HHT uses the empirical mode decomposition method to extract signal components and subsequently compute AM and FM

functions by applying the HT.

While Gabor's analytic signal is well-defined in 1D, extension to higher dimensions is nontrivial. The main reason is that there is no straightforward means of extending the Hilbert transform into multiple dimensions. Peyrin, Zhu, and Goutte [97] defined a 2D complex signal by performing a 1D Hilbert transform along a specified direction; such transforms are frequently referred to as *partial Hilbert transform* (pHT). Havlicek, Havlicek, and Bovik proposed an *adjusted* multiplier [53] to enforce harmonic correspondence [124] of the multidimensional pHT and used it to define a multidimensional *analytic image* [49]. Other related developments include Hahn's single orthant complex signal [45] and the hypercomplex signal of Bülow and Sommer [19], both of which are important but of limited interest for our purposes here because they do not provide a complete representation of all possible signal orientations. Felsberg and Sommer [34] introduced the multidimensional *monogenic* signal where the 1D HT is replaced by an nD Riesz transform. Independently, Larkin et al. [67] used the same signal model to study fingerprints.

While other approaches exist for computing AM-FM models without an explicit complex extension such as the Teager-Kaiser energy operator [77] and the quasi-local approximation [43, 87], the explicit complex extension approaches including the pHT model and the monogenic signal have remained highly popular. These two models have been successfully applied in many practical applications such as motion estimation [5], target tracking [90], fingerprint modeling [67], and texture analysis [96] just to name a few. Both compute an explicit complex extension for the real signal by adding an imaginary part that is equal to the pHT for the partial Hilbert approach and equal to the Riesz transform for the monogenic

signal. The main advantage of the monogenic signal lies in the *isotropic* kernel of the Riesz transform, which tends to avoid undesirable rippling in the AM function; such amplitude ripples are characteristic of the AM functions typically obtained via the pHT. As will be described in more detail below, the main advantage of the pHT approach is that it enables the FM functions to be obtained directly without the need for an auxiliary orientation estimation procedure as is required with the monogenic signal approach.

In this section, I focus on the pHT and monogenic signal approaches. I propose a new alternative algorithm for computing the monogenic signal FM functions that avoids the need for an auxiliary orientation estimation procedure. In addition, I demonstrate that in situations where a multi-scale multi-orientation signal decomposition is required, the pHT approach can deliver equivalent modulation functions to those obtained with the monogenic signal while maintaining a simpler representation. I argue that the pHT method and monogenic signal are both viable approaches. For signals that admit orientations (*e.g.*, spectral support) orthogonal to the direction of action of the pHT, the pHT approach will always suffer from undesirable amplitude rippling that is not present in the monogenic signal. However, pHT based models are always more efficient than the monogenic signal in the sense of requiring only one imaginary component as opposed to two. For signals that do not admit significant spectral support orthogonal to the pHT direction of action, both approaches typically deliver equivalent but slightly different AM-FM interpretations of the signal.

7.3.1 A new algorithm for computing the monogenic FM

As discussed in the complex signal extension approaches in Section 3.2.1, the pHT computes the imaginary part of the complex signal as $a_{\mathbf{e}}(\mathbf{x})\sin[\varphi_{\mathbf{e}}(\mathbf{x})]$ using (3.2), whereas the monogenic approach, without the local orientation $\phi(\mathbf{x})$, actually computes the imaginary part as $a_{\mathrm{mo}}(\mathbf{x})|\sin[\varphi_{\mathrm{mo}}(\mathbf{x})]|$. Therefore, computation of the monogenic FM function involves estimation of the local orientation $\phi(\mathbf{x})$ to deduce the correct sign of the imaginary component. For instance, Larkin, Bone, and Oldfield [67] and Unser, Sage, and Ville [123] adopted the classical tensor structure orientation estimation [64].

I observe that the local orientation $\phi(\mathbf{x})$ in (3.27) is not required for the computation of the FM functions. Taking the derivative of the model (3.1) in both horizontal and vertical directions, we obtain a relationship between the derivatives of the real signal $f(\mathbf{x})$ and the FM function $\nabla\varphi_{\mathrm{mo}(\mathbf{x})} = [\varphi_x(\mathbf{x})\ \varphi_y(\mathbf{x})]^T$ according to

$$\begin{aligned} f_x(\mathbf{x}) &= a_x(\mathbf{x})\cos[\varphi_{\mathrm{mo}}(\mathbf{x})] - \varphi_x(\mathbf{x})a_{\mathrm{mo}}(\mathbf{x})\sin[\varphi_{\mathrm{mo}}(\mathbf{x})], \\ f_y(\mathbf{x}) &= a_y(\mathbf{x})\cos(\varphi_{\mathrm{mo}}(\mathbf{x})) - \varphi_y(\mathbf{x})a_{\mathrm{mo}}(\mathbf{x})\sin[\varphi_{\mathrm{mo}}(\mathbf{x})], \end{aligned}$$

where $\nabla a_{\mathrm{mo}}(\mathbf{x}) = [a_x(\mathbf{x})\ a_y(\mathbf{x})]^T$ and $\nabla f(\mathbf{x}) = [f_x(\mathbf{x})\ f_y(\mathbf{x})]^T$. I then rearranged these derivatives to compute the magnitude of the monogenic FM functions

$$\begin{aligned} |\varphi_x(\mathbf{x})| &= \frac{|a_x(\mathbf{x})\cos[\varphi_{\mathrm{mo}}(\mathbf{x})] - f_x(\mathbf{x})|}{|a_{\mathrm{mo}}(\mathbf{x})\sin[\varphi_{\mathrm{mo}}(\mathbf{x})]|}, \\ |\varphi_y(\mathbf{x})| &= \frac{|a_y(\mathbf{x})\cos[\varphi_{\mathrm{mo}}(\mathbf{x})] - f_y(\mathbf{x})|}{|a_{\mathrm{mo}}(\mathbf{x})\sin[\varphi_{\mathrm{mo}}(\mathbf{x})]|}, \end{aligned} \tag{7.17}$$

where the denominator is obtained from (3.27) as

$$|a_{\mathrm{mo}}(\mathbf{x})\sin[\varphi_{\mathrm{mo}}(\mathbf{x})]| = \sqrt{f_1^2(\mathbf{x}) + f_2^2(\mathbf{x})}. \tag{7.18}$$

I calculated the orientation of the FM vector as

$$\theta(\mathbf{x}) = \arctan\left(\frac{a_y(\mathbf{x})\cos[\varphi_{\text{mo}}(\mathbf{x})] - f_y(\mathbf{x})}{a_x(\mathbf{x})\cos[\varphi_{\text{mo}}(\mathbf{x})] - f_x(\mathbf{x})}\right), \tag{7.19}$$

where $-\pi/2 \leq \theta(\mathbf{x}) \leq \pi/2$. Notice that the range restriction of $\theta(\mathbf{x})$ is the result of the phase ambiguity of the model (7.16) where both $-\varphi_{\text{mo}}(\mathbf{x})$ and $+\varphi_{\text{mo}}(\mathbf{x})$ are valid representations because $\cos[\varphi_{\text{mo}}(\mathbf{x})]$ is an even function. From (7.17) and (7.19), the FM functions are obtained according to

$$\varphi_x(\mathbf{x}) = \sqrt{|\varphi_x(\mathbf{x})|^2 + |\varphi_x(\mathbf{x})|^2}\cos[\theta(\mathbf{x})], \tag{7.20}$$

$$\varphi_y(\mathbf{x}) = \sqrt{|\varphi_y(\mathbf{x})|^2 + |\varphi_y(\mathbf{x})|^2}\sin[\theta(\mathbf{x})]. \tag{7.21}$$

I compare the computed FM function of the proposed method and the computed FM of the original monogenic signal. The simulation results are given in Fig. 7.6. The chirp image is shown in Fig. 7.6(a). The magnitude of the FM computed from the original monogenic signal model and from the proposed method are depicted in Fig. 7.6(c) and (d), respectively. The absolute difference between the two magnitude FM functions is given in Fig. 7.6(b). The artifact of the computed FM magnitude is clearly visible and dominates the image in Fig. 7.6(c). These high magnitude values occur at locations where the computed phase is wrapped - the phase contains jumps that are integer multiples of π. These discontinuities do not exist in FM magnitude computed by the proposed method. I also showed the FM field plots of the original monogenic signal in Fig. 7.6(e) and of the proposed method in Fig. 7.6(f).

7.3.2 Relationship between Monogenic and partial Hilbert approaches

In practical applications, a multipartite signal may admit multiple orientations at a given pixel. Therefore, it is desirable to decompose the signal into individual

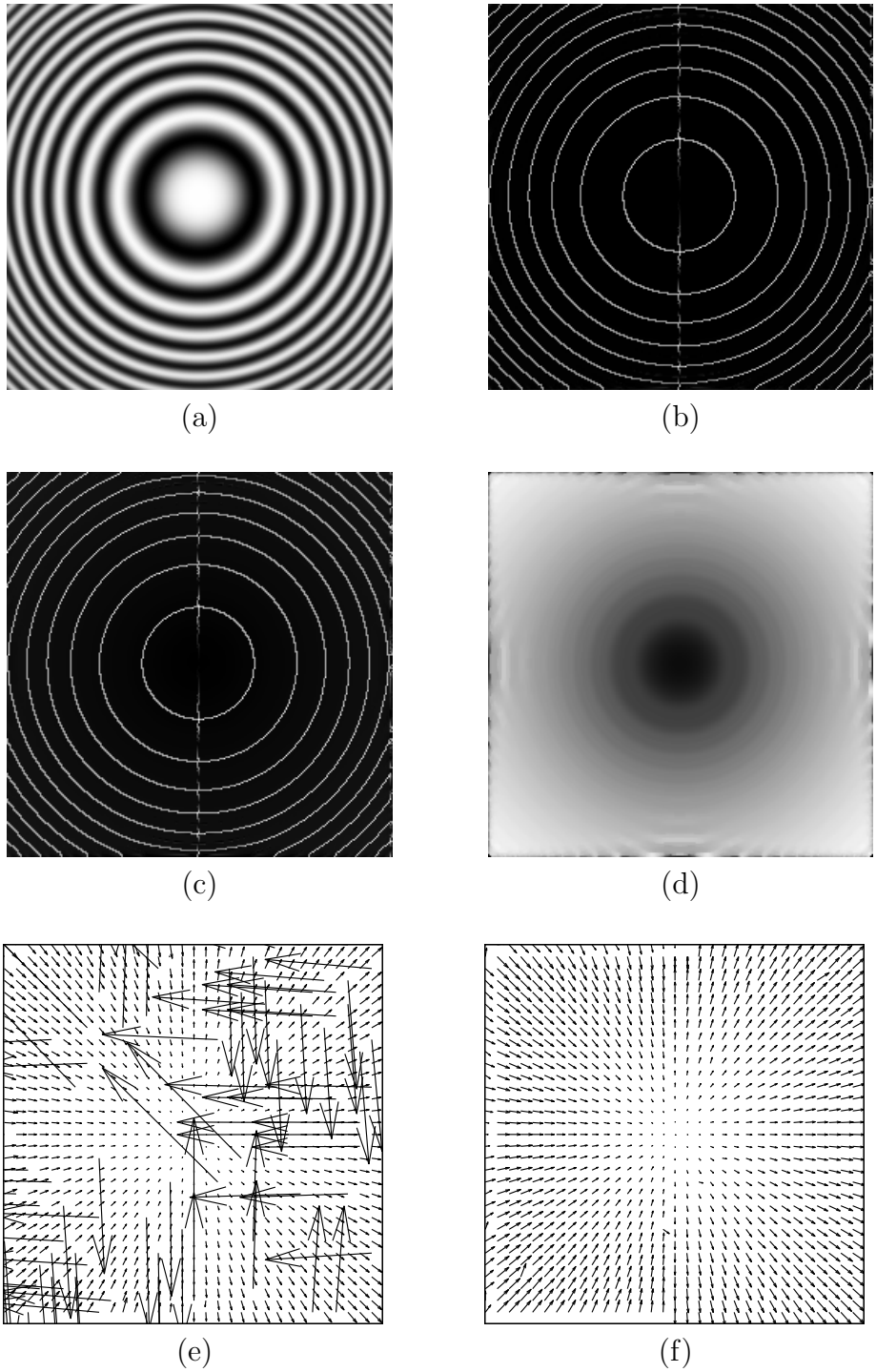


Figure 7.6: Regular FM v.s. modified FM computed using the monogenic signal. (a) Original chirp image. (b) $|(c) - (d)|$. (c) Monogenic $|\varphi_k(\mathbf{n})|$. (d) Modified Monogenic $|\varphi_k(\mathbf{n})|$. (e) Regular monogenic FM field. (f) Modified monogenic FM field.

components of different orientations so that the assumptions inherent in (7.16) are valid. For instance, the monogenic model has been used with multi-scale multi-orientation representations such as wavelets [94,123] and the steerable pyramid [54] to analyze local signal features. Given an input signal specified by a single orientation, both the pHT and the monogenic approach produce the same frequency magnitude for the imaginary components $q_{\mathbf{e}}(\mathbf{x})$ and $q_{\mathrm{mo}}(\mathbf{x})$.

Assume that $s(\mathbf{x})$ is the output from one channel in a multi-scale multi-orientation realization of the steerable pyramid filterbank [38] modified as described in [89]. In particular, let θ_0 be the orientation of filter center frequency; the spectrum of $s(\mathbf{x})$ can be written in polar form as

$$\widehat{s}(r,\theta) = \widehat{\xi}(r,\theta)[\cos(\theta-\theta_0)]^{2\ell}, \tag{7.22}$$

where $\widehat{\xi}(r,\theta)$ is the Fourier spectrum of the original image $f(\mathbf{x})$ at a given scale and ℓ denotes the number of orientations per scale.

For a real signal $s(\mathbf{x})$, the Riesz transform of $s(\mathbf{x})$ produces two components $s_1(\mathbf{x})$ and $s_2(\mathbf{x})$. Similar to [67, 123], we represent these two components by a complex signal $p(\mathbf{x}) = s_1(\mathbf{x}) + js_2(\mathbf{x})$. Note that the complex signal $p(\mathbf{x})$ plays the role of the imaginary image in the context of the complex signal extension approach. It may be shown that the Riesz transform has unity gain:

$$\begin{aligned} ||\widehat{p}(\boldsymbol{\omega})||^2 &= ||\widehat{s}(\boldsymbol{\omega})||^2 \left[\frac{\omega_x^2}{||\boldsymbol{\omega}||^2} + \frac{\omega_y^2}{||\boldsymbol{\omega}||^2}\right] \\ &= ||\widehat{s}(\boldsymbol{\omega})||^2. \end{aligned} \tag{7.23}$$

To demonstrate that the pHT also has unity gain for the specific type of signal in (7.22), let $g(\mathbf{x}) = \mathcal{H}_{\theta_0}\{s(\mathbf{x})\}$, let $\mathbf{e} = [\cos(\theta_0)\ \sin(\theta_0)]^T$ be the unit vector

with angle θ_0, and let $\kappa = [\cos(\theta) \ \sin(\theta)]^T$ be the polar representation of the rectangular frequency $\boldsymbol{\omega} = [\omega_x \ \omega_y]^T$. The pHT of $g(\mathbf{x})$ can be written as

$$\begin{aligned} \widehat{g}(r,\theta) &= -j\text{sgn}(\kappa^T\mathbf{e})\widehat{s}(r,\theta) \\ &= -j\text{sgn}[\cos(\theta-\theta_0)]\widehat{s}(r,\theta). \end{aligned} \tag{7.24}$$

The frequency magnitude of $g(\mathbf{x})$ is then given by

$$||\widehat{g}(r,\theta)||^2 = \begin{cases} ||\widehat{s}(r,\theta)||^2, & \text{if } \cos(\theta-\theta_0) \neq 0, \\ 0, & \text{if } \cos(\theta-\theta_0) = 0. \end{cases} \tag{7.25}$$

According to (7.22), $\widehat{s}(r,\theta)$ vanishes when $\cos(\theta-\theta_0) = 0$. Hence, $||\widehat{g}(\boldsymbol{\omega})||^2 = ||\widehat{s}(\boldsymbol{\omega})||^2 = ||\widehat{p}(\boldsymbol{\omega})||^2$.

While the frequency magnitude of the imaginary signals in the two approaches are the same, the computed AM functions are not necessarily equal.

Fig. 7.7 shows the AM computed by the two approaches. The original *barbara* image is given in Fig. 7.7(a). A spectral depiction of the modified steerable pyramid is given in Fig. 7.7(b). Fig. 7.7(c) shows one component of the multipartite image obtained as a steerable pyramid filterbank channel response. The absolute difference of the computed AM from the two approaches is given in Fig. 7.7(d), where brighter pixels denote a greater difference.

Figs. 7.7(e) and (f) depict the computed AM functions of the monogenic signal and the pHT approach respectively. In this example, the mean difference between the two AM functions is 0.022037 with reference to the range of AM functions [0, 12.3618]. While the pHT approach does not compute the exact AM and FM functions as those obtained by the monogenic signal, it provides an equivalent AM-FM representation.

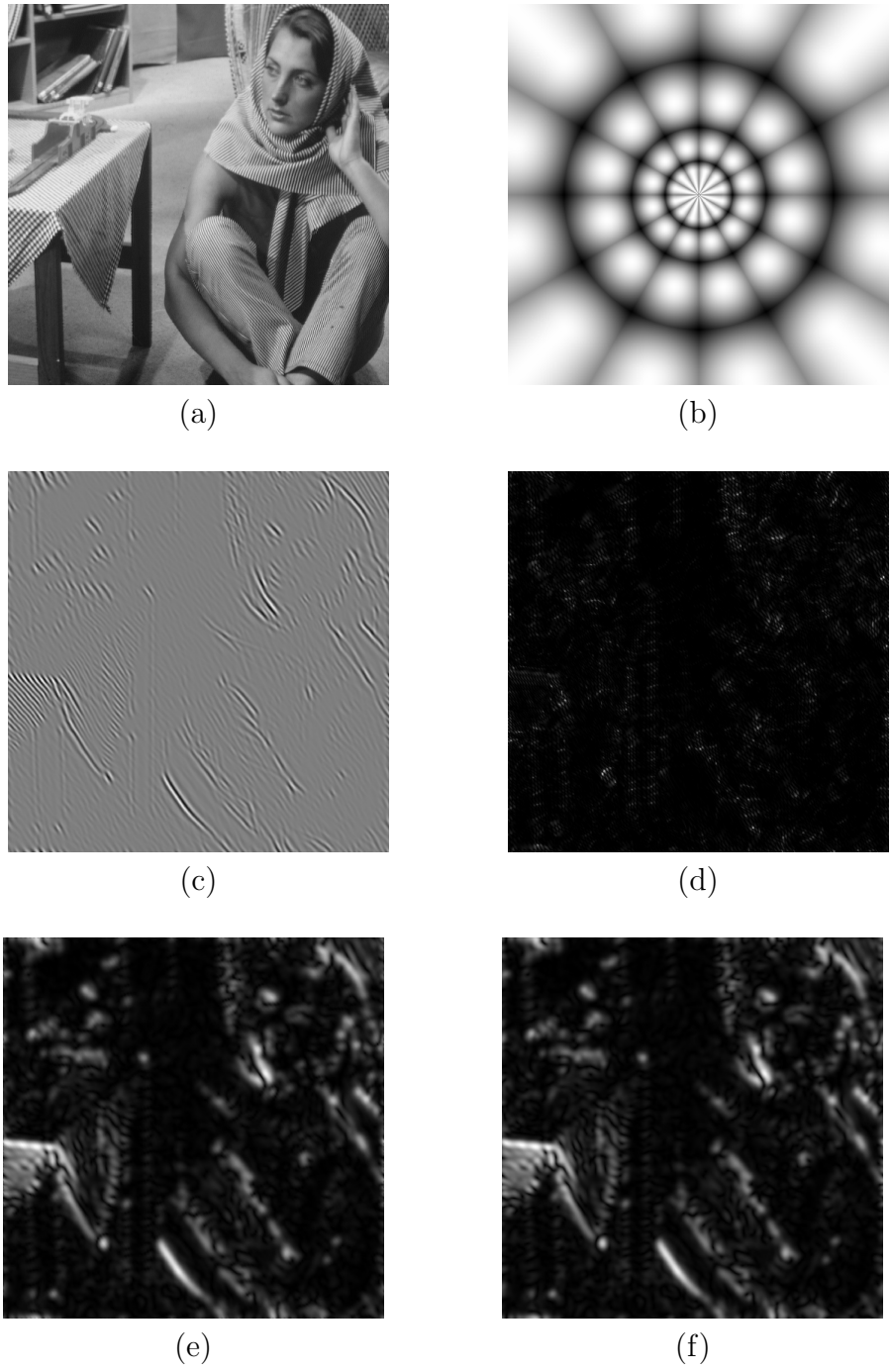


Figure 7.7: Computed AM using the monogenic signal and pHT. (a) Original barbara image. (b) Steerable pyramid filterbank. (c) One component of (a). (d) |(e) - (f)|. (e) Monogenic AM. (f) pHT AM.

7.3.3 Discussion and Conclusion

As noted in Section 3.2.1, the pHT kernel is not isotropic and does not have unity gain. However, for the particular signal model in (7.22), the pHT kernel no longer causes undesirable amplitude rippling because the spectrum $\widehat{s}(r, \theta)$ vanishes at the same places where the pHT kernel is zero. In other words, the pHT kernel can be interpreted as being effectively isotropic for signals of type (7.22). In addition, the pHT offers a simpler representation as the associated complex signal contains two components whereas the monogenic signal requires three. Because the pHT computes the imaginary image as $a(\mathbf{x}) \sin[\varphi(\mathbf{x})]$ instead of $|a(\mathbf{x}) \sin[\varphi(\mathbf{x})]|$ of the monogenic signal, we can use a simpler computation method in Section 3.2.1 to compute the FM function $\nabla\varphi(\mathbf{x})$ without having to estimate the local orientation $\phi(\mathbf{x})$ as is required in the pure monogenic signal model.

In this section, I proposed a new algorithm to compute the FM functions for the monogenic signal. I showed that the proposed algorithm is simpler than the monogenic signal model as it does not require the local orientation estimation step. In addition, I showed that in situations where a multi-scale multi-orientation signal decomposition is required to analyze a signal, both the pHT and the monogenic signal deliver similar AM and FM functions. However, the pHT provides simpler computations. Therefore, I advocated for use of the pHT in multi-scale multi-orientation AM-FM applications.

Chapter 8

Conclusions and Future Work

In this dissertation, I developed a new the multi-component perfect reconstruction AM-FM image transform called the xAMFM. I argued that the xAMFM is an attractive representation for image processing applications. First, I reasoned that the xAMFM is a perceptually motivated image transform. The xAMFM is equipped with a multi-orientation joint time-frequency localized filterbank which was designed to mimic the operation of the human visual system. Second, the xAMFM is able to capture nonstationary features in signals by computing the instantaneous amplitude modulation function (AM) and instantaneous frequency modulation function (FM). For images, the AM and FM offer intuitive interpretation of the local contrast and local texture structures. Finally, the xAMFM opens up a new research area where image processing filters are designed and operate in the AM-FM domain, rather than in the classical pixel domain or the Fourier domain. I showed that the the xAMFM can produce high fidelity image filtering results that are difficult or impossible to obtain with traditional LSI filtering in the pixel domain or in the Fourier domain for texture and orientation filtering applications.

In Chapter 2, I introduced the 1D AM-FM signal model as a viable representation for nonstationary signals. I discussed the limitations of the classical Fourier representation in cases where the signals of interest are nonstationary. I reviewed

many practical applications where importance nonstationary features of signals can be captured by the instantaneous frequency (IF, or FM). I then discussed three major computational techniques for obtaining the AM and FM functions for 1D signals, namely the Gabor analytic signal approach (AS), the Teager-Kaiser energy operator approach (TKEO), and the quasi-local method (QL). I evaluated the performance of these three methods in terms of the mean squared error between the computed AM and computed FM with respect to their corresponding true signals. Among the IF computation approaches, the AS remains a popular method because its implementation is straight forward and the method is more robust to noise compared to the TKEO.

In Chapter 3, I explored extensions of the 1D AM-FM signal model into multiple multidimensions, with emphasis on the 2D case. I first reasoned that the phase and IF play an important role in many image processing applications, such as image restoration, image segmentation, and optical flow estimation. We discussed the generalization of the 1D AM-FM approaches into 2D. Finally, I evaluated their effectiveness with respected to the mean squared error via simulations using both synthetic and real images. The partial Hilbert transform (pHT) and the monogenic signal produce satisfactory results with low MSE. In addition, the AM and FM computed from these two methods offer perceptually meaningful interpretations of the underlying image structures. The relationship between these two approaches is discussed Chapter 7.

In Chapter 4, I introduced the single component perfect reconstruction AM-FM image model. First, I argued that most previous AM-FM image processing techniques were limited to analysis applications. I then discussed the importance of phase unwrapping in the 2D analysis and reconstruction problems. The least-

squares approach is chosen as the phase reconstruction technique for the xAMFM because it is more resilient to noise than other discussed techniques. With perfect reconstruction FM transform, I defined the single component perfect reconstruction AM-FM image transform. I provided two algorithms for obtaining AM and FM functions from a single component image and reconstructing the original image from the AM and FM functions without errors. I verify the perfect reconstruction property by calculating the MSE between the reconstructed and original images.

In Chapter 5, I introduced the xAMFM for general images. I showed that the signal component AM-FM model in Chapter 4 is not sufficient to represent most images. As a result, I used a filterbank to decompose the image into K components. For this purpose, I modified the well-known steerable pyramid to create a full multi-scale, multi-orientation perfect reconstruction filterbank. In addition, I discussed the problems associated with the pHT wherein the computed AM and FM functions can show artifacts if the image component has frequency support perpendicular to the pHT filtering axis. I overcame this problem by rotating the axis of action of the pHT. Finally, I developed the xAMFM for general images. I verified that the xAMFM allows perfect reconstruction theoretically and experimentally.

In Chapter 6, I introduced the AM-FM image processing framework where filters may be designed to operate directly on the AM and FM functions of an image. I defined two classes of AM-FM image filters, namely the AM-based and the FM-based filters. I showed the performance gain of the AM-based filters over the traditional LSI filters in applications such as selective orientation attenuation, frequency selective filtering, image enhancement, and image fusion. For the FM-based filters, I performed geometric image transformation in the AM-FM domain. I compared the results of the AM-FM filters and the LSI filters using the peak signal

to noise ratio (PSNR) and the structure similarity index (SSIM). In both measures, the AM-FM filters outperform the LSI filters in the image scaling application.

In Chapter 7, I demonstrated the use of the xAMFM transform in image analysis applications. In addition, I showed the relationship between the pHT and the monogenic signal.

The main original contributions of this dissertation include the following.

- I proposed a perfect reconstruction FM algorithm. I used a least-squares approach to recover the phase signal from its gradient. In order to allow perfect reconstruction of the phase function, I enforced an initial condition on the reconstructed phase. The perfect reconstruction FM algorithm plays a critical role in the overall AM-FM transform.

- I constructed a perfect reconstruction multidimensional filterbank by modifying the well-known steerable pyramid. This modified filterbank ensures a true multi-scale multi-orientation signal decomposition. Such a decomposition is required for a perceptually meaningful AM-FM image representation.

- I overcame the problems associated with the partial Hilbert transform by rotating the direction of action of the pHT. This rotation results in artifact free filtering results in the AM-FM domain.

- I proposed the first practical framework for designing filters and performing signal processing directly in the modulation domain. I constructed two new classes of practical modulation domain filters. I showed that these modulation domain filters comparable traditional linear shift invariant (LSI) filters qualitatively and quantitatively in applications such as selective orienta-

tion filtering, selective frequency filtering, and fundamental image geometric transformations. In applications such as orientation and texture filtering and image magnification, the AM-FM outperforms the LSI filters.

- I provided extensions of the xAMFM for image decomposition problems. I illustrated that the AM-FM approach can successfully decompose an image into coherent components such as texture and structure components.

- I investigated the relationship between the two prominent AM-FM computation models, namely the partial Hilbert transform approach (pHT) and the monogenic signal. The established relationship helps unify these two AM-FM algorithms.

This dissertation lays a theoretical foundation for future nonlinear modulation domain image processing applications. For the first time, one can apply modulation domain filters to images and obtain high fidelity, predictable, and systematic results. The design of modulation domain filters is intuitive and simple, yet these filters produce superior results compared to those of pixel domain LSI filters.

Besides modulation domain image processing, this dissertation reopens several fundamental research problems in image processing. For instance, classical image analysis applications such as segmentation and edge detection can be reinvestigated in the modulation domain setting. In addition, as the modulation domain image transform possesses properties such as good correspondence with human visual perception and perfect reconstruction, it can potentially yield performance gains in image and video quality assessment and coding.

While the proposed modulation domain image filtering framework offers several advantages over the traditional pixel domain image processing, it is interesting to investigate its current limitations.

- The current transform is data independent. The transform assumes that the number of scales and the number of orientations per scale are predefined. In other words, the filterbank can produce more components than the original image actually has.

- The current phase reconstruction algorithm is *global*, meaning that changes of the gradient in a small neighborhood can lead to changes of the reconstructed phase in the entire image.

- The FM-based filters lack a mathematical stability analysis. Since the xAMFM transform creates nonlinear components, constraints must be enforced on the these components in order to produce stable and meaningful filtered outputs.

This dissertation opens up many future research directions. An obvious path is to overcome the limitations of the current work. For example, one can design data dependent algorithms to adaptively compute the AM-FM components instead of using a fixed filterbank. Another challenging topic is to study the performance of the xAMFM for corrupted signals.

Bibliography

[1] S. Acton, D. Mukherjee, J. Havlicek, and A. Bovik, "Oriented texture completion by AM-FM reaction-diffusion," *IEEE Trans. Image Process.*, vol. 10, no. 6, pp. 885–896, Jun. 2001.

[2] S. Anliney, "A property of the minimum vectors of a regularizing functional defined by means of the absolute norm," *IEEE Trans. Image Proc.*, vol. 45, no. 4, pp. 913–917, 1997.

[3] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform," *IEEE Trans. Image Proc.*, vol. 1, no. 2, pp. 205–220, Apr. 1992.

[4] J. Aujol, G. Gilboa, T. Chan, and S. Osher, "Structure-texture image decomposition – Modeling, algorithms and parameter selection," *Intl. J. Comput. Vis*, vol. 67, no. 1, pp. 111–136, 2006.

[5] J. L. Barron, D. J. Fleet, and S. S. Beauchemin, "Performance of optical flow techniques," *Int'l Journal of Comput. Vision*, vol. 12, no. 1, pp. 43–77, 1994.

[6] E. Bedrosian, "A product theorem for hilbert transforms," *Proc. IEEE*, vol. 51, no. 5, pp. 868–869, May 1963.

[7] M. Bertalmio, L. Vese, G. Sapiro, and S. Osher, "Simultaneous structure and texture image inpainting," *IEEE Trans. Image Proc.*, vol. 12, no. 8, pp. 882–889, Aug. 2003.

[8] J. M. Bioucas-Dias and G. Valadão, "Phase unwrapping via graph cuts," *IEEE Trans. Imag. Proc.*, vol. 16, no. 3, pp. 698–709, Mar. 2007.

[9] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal – Part I. Fundamentals," *Proc. IEEE*, vol. 80, no. 4, pp. 520–538, Apr. 1992.

[10] ——, "Estimating and interpreting the instantaneous frequency of a signal – Part II. Algorithms and applications," *Proc. IEEE*, vol. 80, no. 4, pp. 540–568, Apr. 1992.

[11] B. Boashash and L. White, "Instantaneous frequency estimation and automatic time-varying filtering," in *Proc. IEEE Int'l. Conf. Acoust., Speech, Signal Process.*, Albuquerque, NM, Apr. 03-06, 1990, pp. 1221–1224.

[12] A. C. Bovik, "Perceptual video processing: Seeing the future," *Proc. IEEE*, vol. 98, no. 11, pp. 1799–1803, Nov. 2010.

[13] 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, Jan. 1990.

[14] 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, Mar. 1992.

[15] A. C. Bovik and P. Maragos, "Conditions for positivity of an energy operator," *IEEE Trans. Signal Proc.*, vol. 42, no. 2, pp. 469–471, Feb. 1994.

[16] 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. 41, no. 12, pp. 3245–3265, Dec. 1993.

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

[18] A. Buades, T. Le, J.-M. Morel, and L. Vese, "Fast cartoon + texture image filters," *IEEE Trans. Image Proc.*, vol. 19, no. 8, pp. 1978–1986, Aug. 2010.

[19] T. Bülow and G. Sommer, "The hypercomplex signal - a novel extension of the analytic signal to the multidimensional case," *IEEE Trans. Signal Proc.*, vol. 49, no. 11, pp. 2844–2852, Dec. 2001.

[20] P. J. Burt and E. H. Adelson, "The Laplacian pyramid as a compact image coding," *IEEE Trans. Commun.*, vol. COM-31, no. 4, pp. 532–540, Apr. 1983.

[21] F. W. Campbell and J. G. Robson, "Application of fourier analysis to the visibility of gratings," *J. Physiol.*, vol. 197, pp. 551–566, 1968.

[22] K. R. Castleman, M. Schulze, and Q. Wu, "Simplified design of steerable pyramid filters," in *Proc. Int'l. Symp. Circuit and Syst.*, Montery, CA, May. 31 - Jun. 3, 1998, pp. 329–332.

[23] V. Chen and S. Qian, "Joint time-frequency transform for radar range-doppler imaging," *IEEE Trans. Aero. Elect. Syst.*, vol. 34, no. 2, pp. 486 –499, Apr. 1998.

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

[25] ——, *Time-frequency Analysis*. Prentice-Hall, NJ: Englewood Cliffs, 1995.

[26] M. Costantini, "A novel phase unwrapping method based on network programming," *IEEE Trans. Geos. Remot. Sensi.*, vol. 36, no. 3, pp. 813–821, May. 1998.

[27] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Proc.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.

[28] 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, Jul. 1985.

[29] M. N. Do and M. Vetterli, "The contourlet transform: An efficient directional multiresolution image representation," *IEEE Trans. Image Proc.*, vol. 14, no. 12, pp. 2091–2106, Dec. 2005.

[30] D. L. Donoho, "De-Noising by soft-thresholding," *IEEE Trans. Info. Theory*, vol. 41, no. 3, pp. 613–627, May 1995.

[31] G. Evangelopoulos and M. Maragos, "Texture modulation-constrained image decomposition," in *Proc. Int'l. Conf. Imag. Proc.*, San Diego, CA, Oct. 12-15, 2008, pp. 793–796.

[32] T. Ezzat, J. Bouvrie, and T. Poggio, "AM-FM demodulation of spectrograms using localized 2D max-Gabor analysis," in *Proc. IEEE Int'l. Conf. Acoust., Speech, Signal Process.*, vol. IV, Honolulu, HI, Apr. 15-20, 2007, pp. 1061–1064.

[33] H. Farid and E. Simoncelli, "Differentiation of discrete multidimensional signals," *IEEE Trans. Imag. Proc.*, vol. 13, no. 4, pp. 496–508, Apr. 2004.

[34] M. Felsberg and G. Sommer, "The monogenic signal," *IEEE Trans. Signal Proc.*, vol. 49, no. 12, pp. 3136–3144, Dec. 2001.

[35] D. J. Fleet and A. D. Jepson, "Computation of component image velocity from local phase information," *Int. J. Comput. Vis.*, vol. 5, no. 1, pp. 77–104, Dec. 1990.

[36] ——, "Stability of phase information," *IEEE Trans. Pattern Anal., Machine Intel.*, vol. 15, no. 12, pp. 1253–1268, Dec. 1993.

[37] R. Franzen, "Kodak lossless true color image suite," http://r0k.us/graphics/kodak/, [Online; accessed 19-April-2012].

[38] W. T. Freeman and E. H. Adelson, "The design and use of steerable filters," *IEEE Trans. Pattern Anal., Machine Intel.*, vol. 38, no. 2, pp. 587–607, March. 1992.

[39] D. L. Fried, "Least-squares fitting a wave-front distortion estimation to an array of phase-difference measurements," *J. Opt. Soc. Amer.*, vol. 67, no. 3, pp. 370–375, Mar. 1977.

[40] B. Friedlander and J. M. Francos, "Model based phase unwrapping of 2-D signals," *IEEE Trans. Imag. Proc.*, vol. 44, no. 12, pp. 2999–3007, Dec. 1996.

[41] G. Gabor, "Theory of communication," *J. Inst. Elec. Engr.*, vol. 93, pp. 429–457, 1946.

[42] D. C. Ghiglia and L. A. Romero, "Robust two-dimensional weighted and unweighted phase unwrapping that uses fast transforms and iterative methods," *J. Opt. Soc. Amer. A*, vol. 11, no. 1, pp. 107–117, Jan. 1994.

[43] G. Girolami and D. Vakman, "Instantaneous frequency estimation and measurement: a quasi-local method," *Meas. Sci. Technol.*, vol. 13, pp. 909–917, May 2002.

[44] R. M. Goldstein, H. A. Zebker, and C. L. Werner, "Satellite radar interferometry: Two-dimensional phase unwrapping," *Radio Science*, vol. 23, no. 4, pp. 713–720, July. 1988.

[45] S. L. Hahn, "Multidimensional complex signals with single-orthant spectra," *Proc. IEEE*, vol. 80, no. 8, pp. 1287–1300, Aug. 1992.

[46] ——, *Hilbert transforms in signal processing.* Boston: Artech House, 1996.

[47] H. M. Hanson, P. Maragos, and A. Potamianos, "A system for finding speech formants and modulations via energy separation," *IEEE Trans. Speech Audio Process.*, vol. 2, no. 3, pp. 436–443, Jul. 1994.

[48] 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, Jun. 1996.

[49] ——, "Multidimensional quasi-eigenfunction approximations and multicomponent AM-FM models," *IEEE Trans. Image Proc.*, vol. 9, no. 2, pp. 227–242, Feb. 2000.

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

[51] J. P. Havlicek, C. T. Nguyen, and M. Yeary, "Modulation domain infrared target models," in *Targets and Backgrounds XII: Characterization and Representation*, W. Watkins and D. Clement, Eds. Orlando, FL: SPIE v.6239, May. 6, 2006, pp. 2390D–1 – 62 390D–11.

[52] J. P. Havlicek, P. C. Tay, and A. C. Bovik, "AM-FM image models: Fundamental techniques and emerging trends," in *Handbook of Image and Video Processing*, 2nd ed., A. C. Bovik, Ed. Burlington, MA: Elsevier Academic Press, 2005, pp. 377–395.

[53] J. Havlicek, J. Havlicek, and A. Bovik, "The analytic image," in *Proc. IEEE Int'l Conf. Image Process.*, Santa Barbara, CA, Oct. 26-29 1997, pp. 446–449.

[54] S. Held, M. Storath, P. Massopust, and B. Forster, "Steerable wavelets frames based on the Riesz transform," *IEEE Trans. Imag. Proc.*, vol. 19, no. 3, pp. 653–667, Mar. 2010.

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

[56] R. H. Hudgin, "Wave-front reconstruction from compensated imaging," *J. Opt. Soc. Amer.*, vol. 67, no. 3, pp. 375–378, Mar. 1977.

[57] B. R. Hunt, "Matrix formulation of the reconstruction of phase values from phase differences," *J. Opt. Soc. Amer.*, vol. 69, no. 3, pp. 393–399, Mar. 1979.

[58] imagefusion.org, http://www.imagefusion.org, [Online; accessed 19-April-2012].

[59] J. Imberger and B. Boashash, "Application of the Wigner-Ville distribution to temperature gradient microstructure: A new technique to study small scale variations," *J. Phys. Oceannography*, vol. 16, no. 12, pp. 1997–2012, 1986.

[60] M. J. W. J. Portilla, V. Strela and E. P. Simoncelli, "Image denoising using scale mixtures of Gaussians in the wavelet domain," *IEEE Trans. Image Proc.*, vol. 12, no. 11, pp. 1338–1351, Nov. 2003.

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

[62] R. G. Keys, "Cubic convolution interpolation for digital image processing," *IEEE Trans. Acoust. Speech, Signal Proc.*, vol. 29, no. 6, pp. 1153–1160, Dec. 1981.

[63] N. Kitiyanan and J. P. Havlicek, "Modulation domain reference point detection for fingerprint recognition," in *Proc. IEEE Southwest Symp. Image Anal. & Interp.*, vol. II, Lake Tahoe, NV, Mar. 28-30 2004, pp. 147–151.

[64] H. Knutsson, "Representing local structure using tensors," in *Proc. 6th Scandinavian Conf. Image Anal.*, Oulu, Finland, Jun. 1989, pp. 244–251.

[65] J. J. Koenderink, "The structure of images," *Biological Cybernetics*, vol. 50, pp. 363–370, 1984.

[66] P. Kovesi, "Image features from phase congruency," *Videre*, vol. 1, no. 3, Summer 1999.

[67] K. G. Larkin, D. J. Bone, and M. A. Oldfield, "Natural demodulation of two-dimensional fringe patterns. I. General background on the spiral phase quadrature transform," *J. Opt. Soc. Am. A*, vol. 18, no. 8, pp. 1862–1870, Aug. 2001.

[68] K. Larkin and P. A. Fletcher, "A coherent framework for fingerprint analysis: are fingerprints holograms?" *Opt. Express.*, vol. 15, no. 14, pp. 8667–8677, Jun. 2007.

[69] Z. Liu, E. Blasch, Z. Xue, J. Zhao, R. Laganiére, and W. Wu, "Objective assessement of multiresolution image fusion algorithms for context enhancement in night vision: A comparative study," *IEEE Trans. Pattern. Anal. and Machi. Intell.*, vol. 34, no. 1, pp. 94–109, Jan. 2012.

[70] C. Loizou, V. Murray, M. Pattichis, M. Pantziaris, and C. Pattichis, "Multiscale amplitude-modulation frequency-modulation (AM-FM) texture analysis of ultrasound images of the intima and media layers in the carotid artery," *IEEE Trans. Info. Tech. Biomedicine*, vol. 15, no. 12, pp. 178–188, Mar. 2011.

[71] P. J. Loughlin and B. Tacer, "Comments on the interpretation of instantaneous frequency," *IEEE Signal Proc. Letters*, vol. 4, no. 5, pp. 123–125, May 1997.

[72] S. Lu and P. C. Doerschuk, "Nonlinear modeling and processing of speech based on sums of AM-FM formant models," *IEEE Trans. Signal Proc.*, vol. 44, no. 4, pp. 773–782, Apr. 1996.

[73] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal., Machine Intel.*, vol. 11, no. 7, pp. 674–693, July. 1989.

[74] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Proc.*, vol. 41, no. 12, pp. 3397–3415, Dec. 1993.

[75] L. Mandel, "Interpretation of instantaneous frequencies," *Am. J. Phys.*, vol. 42, pp. 840–846, 1974.

[76] B. S. Manjunath and W. Y. Ma, "Texture features for browsing and retrieval of image data," *IEEE Trans. Pattern Anal., Machine Intel.*, vol. 18, no. 8, pp. 837–842, Aug. 1996.

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

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

[79] ——, "On amplitude and frequency demodulation using energy operators," *IEEE Trans. Signal Proc.*, vol. 41, no. 4, pp. 1532–1550, Apr. 1993.

[80] D. Marr and E. Hildreth, "Theory of edge detection," *Proc. R. Soc. Lond. B*, vol. 207, pp. 187–217, 1980.

[81] W. Martin and P. Flandrin, "Wigner-Ville spectral analysis of nonstationary processes," *IEEE Trans. Acoust. Speech, Signal Proc.*, vol. 33, no. 6, pp. 1461–1470, Dec. 1985.

[82] S. Marčelja, "Mathematical description of the responses of simple cortical cells," *J. Opt. Soc. Am.*, vol. 70, no. 11, pp. 1297–1300, Nov. 1980.

[83] F. G. Meyer, A. Z. Averbuch, and R. R. Coifman, "Multilayered image representation: Application to image compression," *IEEE Trans. Image Proc.*, vol. 11, no. 9, pp. 1072–1080, Sept. 2002.

[84] M. Meyer, *Oscillating patterns in image processing and nonlinear evolution equations: The fifteenth Dean Jacqueline B. Lewis memorial lectures*. AMS, 2001.

[85] A. Middleditch and L. R. Wyatt, "An instantaneous-frequency filtering technique for high-frequency radar oceanography," *IEEE Journal of Oceanic Engr.*, vol. 31, no. 4, pp. 797–803, Oct. 2006.

[86] N. A. Mould, C. T. Nguyen, C. M. Johnston, and J. P. Havlicek, "Online consistency checking for am-fm target tracks," in *Proc. SPIE/IS&T Conf. Computational Imaging VI*, C. Bouman, E. Miller, and I. Pollack, Eds. San Jose, CA: SPIE, Jan. 26-31, 2008, pp. 681 413–1 – 681 413–12.

[87] V. Murray, P. Rodríguez, and M. Pattichis, "Multiscale AM-FM demodulation and image reconstruction methods with improved accuracy," *IEEE Trans. Imag. Proc.*, vol. 19, no. 5, pp. 1138–1152, 2010.

[88] C. T. Nguyen, P. Campbell, and J. P. Havlicek, "FM filters for modulation domain image processing," in *Proc. IEEE Int'l. Conf. Image Proc.*, Cairo, Egypt, Nov. 7-11, 2009, pp. 3973–3976.

[89] C. T. Nguyen and J. P. Havlicek, "AM-FM image filters," in *Proc. IEEE Int'l. Conf. Image Proc.*, San Diego, CA, Oct. 12-15, 2008, pp. 789–792.

[90] C. T. Nguyen, J. P. Havlicek, and M. Yeary, "Modulation domain template tracking," in *Proc. 4th Joint IEEE Int'l. Workshop Object Tracking, Class., in and Beyond the visible Spectrum, in conjunction with the 2007 IEEE Computer Soc. conf. Comput. Vision, Patern Recog.*, Minneapolis, Fe, MN, Jun. 17-22, 2007.

[91] C. T. Nguyen, R. A. Sivley, and J. P. Havlicek, "First results in perceptually-based AM-FM image filtering," in *Proc. IEEE Southwest Symp. Image Anal., Interp.*, Santa Fe, NM, Mar. 24-26, 2008, pp. 77–80.

[92] C. T. Nguyen, J. D. Williams, M. Özaydin, and J. P. Havlicek, "FM processing with generalized amplitude and phase: Application to modulation domain geometric image transformations," in *Proc. IEEE Int'l. Conf. Image Proc.*, Brussels, Belgium, Sep. 11-14, 2011, pp. 81–84.

[93] G. Nico, G. Palubinskas, and M. Datcu, "Bayesian approaches to phase unwrapping: Theoretical study," *IEEE Trans. Imag. Proc.*, vol. 48, no. 8, pp. 2545–2556, Sept. 2000.

[94] S. Olhede and G. Metikas, "The monogenic wavelet transform," *IEEE Trans. Signal Proc.*, vol. 57, no. 9, pp. 3426–3441, Sept. 2009.

[95] A. V. Oppenheim and J. S. Lim, "The importance of phase in signals," *Proc. IEEE*, vol. 69, no. 5, pp. 529–541, May. 1981.

[96] M. Pattichis and A. Bovik, "Analyzing image structure by multidimensional frequency modulation," *IEEE Trans. Pattern Anal., Machine Intel.*, vol. 29, no. 5, pp. 753–766, May 2007.

[97] F. Peyrin, Y. M. Zhu, and R. Goutte, "Extension of the notion of analytic signal for multidimensional signals. Application to images," in *Signal Processing III: Theories and Applications*, I. T. Young, *et al.*, Ed. Amsterdam, B. V. (North-Holland): Elsevier Science Publishers, 1986, pp. 677–680.

[98] J. Portilla and E. P. Simoncelli, "A parametric texture model based on joint statistics of complex wavelet coefficients," *Int'l. Journal of Comput. Vision.*, vol. 40, no. 1, pp. 49–71, Dec. 2000.

[99] J. Portilla, V. Strela, M. Wainwright, and E. P. Simoncelli, "Image denoising using scale mixtures of gaussians in the wavelet domain," *IEEE Trans. Image Proc.*, vol. 12, no. 11, pp. 1338–1351, Nov. 2003.

[100] M. D. Pritt, "Congruence in least-squares phase unwrapping," in *Proc. IEEE Int'l. Conf. Geosci. Remote Sensi.*, Singapore, Aug. 3-8, 1997, vol. II, pp. 875–877.

[101] G. Qu, D. Zhang, and P. Yan, "Information measure for performance of image fusion," *Elect. Letters.*, vol. 38, no. 7, pp. 313–315, Mar. 2002.

[102] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, vol. 60, pp. 259–268, 1992.

[103] E. Sejdić, L. Stanković, M. Daković, and J. Jang, "Instantaneous frequency estimation using the s-transform," *Signal Proc. Letters*, vol. 15, pp. 309–312, 2008.

[104] R. Shahidi and C. Moloney, "Decorrelating the structure and texture components of a variational decomposition model," *IEEE Trans. Image Proc.*, vol. 18, no. 2, pp. 299–309, Feb. 2009.

[105] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Trans. Image Proc.*, vol. 15, pp. 430–444, Feb. 2006.

[106] E. P. Simoncelli and W. T. Freeman, "The steerable pyramid: a flexible architecture for multi-scale derivative computation," in *Proc. IEEE Int'l. Conf. Image Proc.*, Washington, DC., Oct. 23-26, 1995, pp. 444–447.

[107] E. P. Simoncelli, W. T. Freeman, E. H. Adelson, and D. J. Heeger, "Shiftable multi-scale transform," *IEEE Trans. Info. Theory*, vol. 38, no. 2, pp. 587–607, March. 1992.

[108] R. A. Sivley and J. P. Havlicek, "Multidimensional phase unwrapping for consistent APF estimation," in *Proc. IEEE Int'l. Conf. Image Proc.*, Genoa, Italy, Sep. 11-14, 2005, vol. II, pp. 458–461.

[109] ——, "Perfect reconstruction AM-FM image models," in *Proc. IEEE Int'l. Conf. Image Proc.*, Atlanta, GA, Oct. 8-11, 2006, pp. 2125–2128.

[110] R. Sivley and J. Havlicek, "Perfect reconstruction AM-FM image models," in *Proc. IEEE Int'l. Conf. Image Process.*, Atlanta, GA, Oct. 8-11 2006, pp. 2125 – 2128.

[111] A. Skodras, C. Christopoulos, and T. Ebrahimi, "The JPEG 2000 still image compression standard," *IEEE Signal Proc. Mag.*, vol. 18, no. 5, pp. 36–58, Sep. 2001.

[112] U. Spagnolini, "2-D phase unwrapping and instantaneous frequency estimation," *IEEE Trans. Geos. Remot. Sensi.*, vol. 33, no. 3, pp. 579–589, May. 1995.

[113] J.-L. Starck, M. Elad, and D. L. Donoho, "Image decomposition via the combination of sparse representations and a variational approach," *IEEE Trans. Imag. Proc.*, vol. 14, no. 10, pp. 1570–1582, Oct. 2005.

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

[115] J. Strand, T. Taxt, and A. K. Jain, "Two-dimensional phase unwrapping using a block least-squares method," *IEEE Trans. Imag. Proc.*, vol. 8, no. 3, pp. 375–386, Mar. 1999.

[116] M. Sun, S. Qian, X. Yan, B. Baumann, X. Xia, R. Dahl, N. Ryan, and R. Sclabassi, "Localizing functional activity in the brain through time-frequency analysis and synthesis of the EEG," *Proc. IEEE*, vol. 84, no. 9, pp. 1302 –1311, Sept. 1996.

[117] H. Takajo and T. Takahashi, "Least-squares phase estimation from the phase difference," *J. Opt. Soc. Amer.*, vol. 5, no. 3, pp. 416–424, Mar. 1988.

[118] ——, "Noniterative method for obtaining the exact solution for the normal equation in least-squares phase estimation from the phase difference," *J. Opt. Soc. Amer.*, vol. 5, no. 11, pp. 1818–1827, Nov. 1988.

[119] M. T. Taner, F. Koehler, and R. E. Sheriff, "Complex seismic trace analysis," *Geophysics*, vol. 44, no. 6, pp. 1041–1063, 1979.

[120] P. Tsiakoulis, A. Potamianos, and D. Dimitriadis, "Short-time instantaneous frequency and bandwidth features for speech recognition," in *IEEE Workshop on Auto. Speech Recog. Understand.*, Albuquerque, NM, Nov. 13 - Dec. 17, 2009, pp. 103–106.

[121] M. Unser, A. Aldroubi, and M. Eden, "B-Spline signal processing: Part I–theory," *IEEE Trans. Image Process.*, vol. 41, no. 2, pp. 821–833, Feb. 1993.

[122] ——, "B-Spline signal processing: Part II–efficient design and applications," *IEEE Trans. Image Process.*, vol. 41, no. 2, pp. 834–848, Feb. 1993.

[123] M. Unser, D. Sage, and D. Ville, "Multiresolution monogenic signal analysis using the Riesz-Laplace wavelet transform," *IEEE Trans. Imag. Proc.*, vol. 18, no. 11, pp. 2402–2418, Nov. 2009.

[124] D. Vakman, "On the analytic signal, the Teager-Kaiser energy algorithm, and other methods for defining amplitude and frequency," *IEEE Trans. Signal Proc.*, vol. 44, no. 4, pp. 791–797, Apr. 1996.

[125] L. Vese and S. Osher, "Modeling textures with total variation minimization and oscillating patterns in image processing," *J. Sci. Comput.*, vol. 19, pp. 553–577, Oct. 2003.

[126] G. K. Wallace, "The JPEG still picture compression standard," *Commucations of the ACM*, vol. 34, no. 4, pp. 31–44, Apr. 1991.

[127] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Proc.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[128] C. S. Xydeas and V. Petrovic, "Image fusion performance measure," *Elect. Letters.*, vol. 36, no. 4, pp. 308–309, Feb. 2000.

[129] T. B. Yap, J. P. Havlicek, and V. DeBrunner, "Bayesian segmentation of AM-FM texture images," in *Proc. 35th IEEE Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, Nov. 4-7, 2001, pp. 1156–1160.

[130] Z. Zhidong and P. Min, "Instantaneous frequency estimation of diastolic murmurs based on EMD and TEO," in *IEEE Int'l. Conf. Bioinfo. Biomed. Engr.*, Wuhan, CN, Jul. 2007, pp. 829–832.

Appendices

Appendix A

Orientation Decomposition

Let M be the number of orientation, $M \in \mathbb{Z}$. We define $\tilde{S}$ as a sum of K components. We write $\tilde{S}$ as

$$\tilde{S} = \sum_{k=0}^{M-1} \cos^{2\ell}(\phi - \frac{\pi k}{K}) \tag{A.1}$$

We expand $\tilde{S}$ according to

$$\begin{aligned}
\tilde{S} &= \frac{1}{2^{2\ell}} \sum_{k=0}^{M-1} \left[e^{i(\phi - \frac{\pi k}{M})} + e^{-i(\phi - \frac{\pi k}{M})} \right]^{2l} \\
&= \frac{1}{2^{2l}} \sum_{k=0}^{M-1} \sum_{m=0}^{2l} \binom{2\ell}{m} e^{i(\phi - \frac{\pi k}{M})m} \times e^{-i(\phi - \frac{\pi k}{M})(2\ell - m)} \\
&= \frac{1}{2^{2l}} \sum_{k=0}^{M-1} \sum_{m=0}^{2l} \binom{2\ell}{m} e^{i2\phi(m-\ell)} \times e^{i(\ell-m)\frac{2\pi k}{M}} \\
&= \frac{1}{2^{2\ell}} \sum_{m=0}^{2\ell} \binom{2\ell}{m} e^{i2\phi(m-\ell)} \times \sum_{k=0}^{K-1} e^{i(\ell-m)\frac{2\pi k}{M}},
\end{aligned} \tag{A.2}$$

where $0 \leq m \leq 2\ell$ and $m \in Z$.

For $\ell = m$, S can be reduced to

$$\tilde{S} = \frac{M}{2^{2\ell}} \cdot \binom{2\ell}{\ell} = \frac{M}{2^{2\ell}} \left[\frac{(2\ell)!}{\ell!(\ell!)} \right].$$

For $l \neq m$, let the sum of the second term in (A.2) be B. We can see that B is a geometric series

$$B = \sum_{k=0}^{M-1} e^{i(l-m)\frac{2\pi k}{M}} = \frac{1 - e^{i(\ell-m)2\pi}}{1 - e^{\frac{i(\ell-m)2\pi}{M}}}. \tag{A.3}$$

As $m \in [0,\ 2\ell]$, then $l - m \in [-\ell,\ \ell]$. Therefore,

$$\frac{-\ell}{M} \leq \frac{\ell - m}{M} \leq \frac{\ell}{M}.$$

Hence, the denominator of (A.3) is non-zero if $M > \ell$. In addition, the numerator of (A.3) is always zero, because ℓ, $n \in \mathbb{Z}$. Therefore, $B = 0$ if $M > \ell$, $\forall \ell, m \in \mathbb{Z}$. Therefore, we arrive at the following result for $\tilde{S}$

$$\tilde{S} = \sum_{k=0}^{M-1} \cos^{2\ell}(\phi - \frac{\pi k}{K}) = \begin{cases} \frac{M}{2^{2\ell}} \left[\frac{(2\ell)!}{\ell!(\ell!)}\right] & \text{if } m = l \text{ and } M > l, \\ 0 & otherwise. \end{cases} \tag{A.4}$$

As discussed in Chapter 5, the oriented filter $G_k(\boldsymbol{\omega})$ has the frequency spectrum in polar form as

$$G_k(r, \theta) = \begin{cases} \alpha_k \left[\cos\left(\theta - \frac{k\pi}{M}\right)\right]^{M-1} & \text{if } \left|\theta - \frac{k\pi}{M}\right| < \frac{\pi}{2}, \\ 0 & \text{otherwise,} \end{cases} \tag{A.5}$$

where $0 \leq k \leq M - 1$ indicate orientation index and the constant α is defined as

$$\alpha = \frac{2^\ell \ell!}{\sqrt{M(2\ell)!}}. \tag{A.6}$$

Let $\ell \in \mathbb{Z}, \ell = M - 1$. We define the sum of K components $G_k(r, \theta)$ as S

$$\begin{aligned} S &= \sum_{k=0}^{M-1} |G_k(\boldsymbol{\omega})|^2 \\ &= \alpha \sum_{k=0}^{M-1} \cos^{2\ell}(\phi - \frac{\pi k}{K}). \end{aligned} \tag{A.7}$$

From (A.4), (A.6), and (A.7), we get

$$\begin{aligned} S &= \alpha^2 \frac{M}{2^{2\ell}} \left[\frac{(2\ell)!}{\ell!(\ell!)}\right] \\ &= \left[\frac{2^\ell \ell!}{\sqrt{M(2\ell)!}}\right]^2 \frac{M}{2^{2\ell}} \left[\frac{(2\ell)!}{\ell!(\ell!)}\right] \\ &= 1. \end{aligned} \tag{A.8}$$

Appendix B

2D Least-squares Phase Unwrapping

Let $\varphi(m,n)$ be a discrete unwrapped phase. Let $\rho(m,n)$ be the measured phase difference. m, n are indices in the rectangular grid $[0,\ M-1]\times[0,\ N-1]$. The 2D least-squares phase unwrapping aims to solve

$$\overbrace{\Big(\varphi(m+1,n)-2\varphi(m,n)+\varphi(m-1,n)\Big)}^{\varphi_{mm}(m,n)}+\overbrace{\Big(\varphi(m,n+1)-2\varphi(m,n)+\varphi(m,n-1)\Big)}^{\varphi_{nn}(m,n)}=\Big(\rho_m(m,n)-\rho_m(m-1,n)\Big)+\Big(\rho_n(m,n)-\rho_n(m,n-1)\Big), \tag{B.1}$$

where $\varphi_{mm}(m,n)$ and $\varphi_{nn}(m,n)$ are the discrete approximation of the second derivatives in the vertical direction and the horizontal direction, respectively. The sum of $\varphi_{mm}(m,n)$ and $\varphi_{nn}(m,n)$ is the discrete Laplacian of the image $\varphi(m,n)$. The equation (B.1) means that the second derivative of the unwrapped phase and that of the wrapped phase must be equal.

Let Φ be the 2D DCT transform of $\varphi(m,n)$. $\varphi(\mathbf{x})$ can be represented from the DCT coefficients as

$$\varphi(m,n)=\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}w(i,j)\Phi(i,j)\cos\left[\frac{\pi}{2M}i(2m+1)\right]\cos\left[\frac{\pi}{2N}j(2n+1)\right], \tag{B.2}$$

where

$$w(i,j) = \begin{cases} 0.25 & \text{if } i = j = 0, \\ 0.5 & \text{if } i = 0 \text{ and } j \neq 0, \\ 0.5 & \text{if } j = 0 \text{ and } i \neq 0, \\ 1.0 & \text{otherwise.} \end{cases}$$

Let Γ be the 2D DCT transform of $\gamma(m,n)$. $\gamma(m,n)$ is represented by DCT coefficients as in (B.2).

Substituting the definition in (B.2) to each element of the left hand side (LHS) in (B.1) and expanding the first cosine term, we have

$$\begin{aligned} \varphi(m+1,n) &= \sum_{i=0}^{M-1}\sum_{j=0}^{N-1} w(i,j)\Phi(i,j)\cos\left[\frac{\pi}{2M}i(2(m+1)+1)\right]\cos\left[\frac{\pi}{2N}j(2n+1)\right] \\ &= \sum_{i=0}^{M-1}\sum_{j=0}^{N-1} w(i,j)\Phi(i,j)\cos\left[\frac{\pi}{2M}i(2m+1)+\frac{\pi}{M}i\right]\cos\left[\frac{\pi}{2N}j(2n+1)\right] \\ &= \sum_{i=0}^{M-1}\sum_{j=0}^{N-1} w(i,j)\Phi(i,j)\Bigg(\cos\left[\frac{\pi}{2M}i(2m+1)\right]\cos\left[\frac{\pi}{M}i\right]\cos\left[\frac{\pi}{2N}j(2n+1)\right]- \\ &\quad \sin\left[\frac{\pi}{2M}i(2m+1)\right]\sin\left[\frac{\pi}{M}i\right]\cos\left[\frac{\pi}{2N}j(2n+1)\right]\Bigg). \end{aligned} \tag{B.3}$$

$$\begin{aligned} \varphi(m-1,n) &= \sum_{i=0}^{M-1}\sum_{j=0}^{N-1} w(i,j)\Phi(i,j)\cos\left[\frac{\pi}{2M}i(2(m-1)+1)\right]\cos\left[\frac{\pi}{2N}j(2n+1)\right] \\ &= \sum_{i=0}^{M-1}\sum_{j=0}^{N-1} w(i,j)\Phi(i,j)\cos\left[\frac{\pi}{2M}i(2m+1)-\frac{\pi}{M}i\right]\cos\left[\frac{\pi}{2N}j(2n+1)\right] \\ &= \sum_{i=0}^{M-1}\sum_{j=0}^{N-1} w(i,j)\Phi(i,j)\Bigg(\cos\left[\frac{\pi}{2M}i(2m+1)\right]\cos\left[\frac{\pi}{M}i\right]\cos\left[\frac{\pi}{2N}j(2n+1)\right]+ \\ &\quad \sin\left[\frac{\pi}{2M}i(2m+1)\right]\sin\left[\frac{\pi}{M}i\right]\cos\left[\frac{\pi}{2N}j(2n+1)\right]\Bigg). \end{aligned} \tag{B.4}$$

Repeating the same expansion for every element on the left hand side of (B.1) and simplifying the overall sum, we arrive at the frequency representation of the LHS in DCT domain

$$\begin{aligned} \text{LHS} &= \sum_{i=0}^{M-1}\sum_{j=0}^{N-1} w(i,j)\Phi(i,j)\cos\left[\frac{\pi}{2M}i(2m+1)\right]\times \\ &\quad \cos\left[\frac{\pi}{2N}j(2n+1)\right]\left(2\cos\left[\frac{\pi}{M}i\right]+2\cos\left[\frac{\pi}{N}j\right]-4\right) \end{aligned} \tag{B.5}$$

Repeating the same expansion for every element on the right hand side (RHS) of (B.1) and simplifying the overall sum, we arrive at the frequency representation of the RHS in DCT domain

$$\text{RHS} = \sum_{i=0}^{M-1}\sum_{j=0}^{N-1} w(i,j)\Gamma(i,j)\cos\left[\frac{\pi}{2M}i(2m+1)\right]\cos\left[\frac{\pi}{2N}j(2n+1)\right] \tag{B.6}$$

From (B.5) and (B.6), we have the DCT relationship according to

$$\Phi(i,j) = \frac{\Gamma(i,j)}{2\cos\left(\frac{\pi}{M}i\right)+2\cos\left(\frac{\pi}{N}j\right)-4}. \tag{B.7}$$