COMPLEX DIRECTIONAL WAVELET TRANSFORMS: REPRESENTATION, STATISTICAL MODELING AND APPLICATIONS

by

AN PHUOC NHU VO

Presented to the Faculty of the Graduate School of

The University of Texas at Arlington in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT ARLINGTON

December 2008

To my parents Vo Hanh & Nguyen Thi Ngoc Oanh

and my husband Nha Hoang Quan Nguyen

who encourage and inspire me to pursue doctoral study.

ACKNOWLEDGEMENTS

I would like to thank my supervising professor Dr. Soontorn Oraintara for constantly motivating and encouraging me, and also for his invaluable advice during the course of my doctoral studies. I wish to thank my committee members, Dr. K. R. Rao, Dr. Michael Manry, Dr. Qilian Liang and Dr. Seoung B. Kim for their interest in my research and for taking time to serve in my dissertation committee.

I am grateful to all the teachers who taught me during the years I spent in school, first in Vietnam and then in the Unites States. To all of the MSP lab alumni and current members with whom I shared long discussions and many late night working hours, I would like to give my sincere thanks. In particular I would like to thank Truong Nguyen who has helped me when I first started my PhD program. Finally, I would like to dedicate this work to my parents, sisters and husband for their sacrifice, encouragement and patience. I also thank several of my friends who have helped me throughout my career.

October 10, 2008

ABSTRACT

COMPLEX DIRECTIONAL WAVELET TRANSFORMS:
REPRESENTATION, STATISTICAL MODELING
AND APPLICATIONS

AN PHUOC NHU VO, Ph.D.
The University of Texas at Arlington, 2008

Supervising Professor: Soontorn Oraintara

The thesis presents an new image decomposition for feature extraction, which is called the pyramidal dual-tree directional filter bank (PDTDFB). The image representation has an overcomplete ratio of less than 8/3 and uses a separable filter bank implementation structure. We discuss how to utilize both magnitude and phase information obtained from the PDTDFB for the purpose of texture image retrieval. The relative phase, which is the difference of phases between two adjacent complex coefficients, has a linear relationship with the angle of dominant orientation within a subband. This information is incorporated to form a new feature vector called CDFB-RP. Another application of PDTDFB is texture segmentation. A new feature extraction method is proposed for texture segmentation. The approach is based on incorporating the phase information obtained from complex filter banks. The PDTDFB is used to decompose a texture image in order to provide complex subband coefficients. The local mean direction, extracted from the phases of the coefficients, is defined as additional features for classification and segmentation.

We proposed a modified version of the PDTDFB for image denoising. Unlike the previous approach, the new FB provides an approximately tight-frame decomposition. Then we proposed the complex Gaussian scale mixture (CGSM) for modeling the distribution of

complex directional wavelet coefficients. The statistical model is then used to obtain the denoised coefficients from the noisy image decomposition by Bayes least squares estimator. Performance of the denoised images using the PDTDFB is compared with the conventional transforms including the orthogonal wavelet, the contourlet and the steerable pyramid.

A new approach which exploits the probabilistic properties from the phase information of two-dimensional complex wavelet coefficients for the image modeling is developed. Definition, property and statistics of relative phase of the complex coefficients are studied in detail. We proposed von Mises and wrapped Cauchy for the probability density function (pdf) of the relative phase in the complex wavelet domain. The von Mises and wrapped Cauchy models are compared, and the simulation results show that the wrapped Cauchy fits well with the peaky and heavy-tailed pdf of the relative phase and the von Mises fits well with the pdf which is in Gaussian shape. For most of the test images, the wrapped Cauchy model is more accurate than the von Mises, when images are decomposed by different complex wavelet transforms including the DTCWT, the PDTDFB and a modified version of curvelet.

With the assumptions of the Gaussian image model as well as the Gaussian scale mixture (GSM), the marginal and joint distributions for the phase of the complex wavelet coefficients are studied in detail. From these hypotheses, we then derive the probability density function of the relative phase (RP-PDF) in complex wavelet domain. We propose the maximum-likelihood method to estimate two RP-PDF parameters. The RP-PDF fits well with behaviors of the relative phase from various real images including texture images as well as natural images. The RP-PDF model is compared with the von Mises and wrapped Cauchy distributions. The experimental results, in which the real images are decomposed by various complex wavelets such as the DTCWT, the PDTDFB and the curvelet, show that the RP-PDF model for relative phase is more accurate than the others.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Problem Statement

The multiscale and multidirectional transform is a tool that has been used widely in the last decade for image processing. The image data is represented by a new set of coefficients organized by location, orientation, and scale. Natural images typically contain many geometrical features, such as edges and textures, so the directional bases may better represent those features. This dissertation investigates:

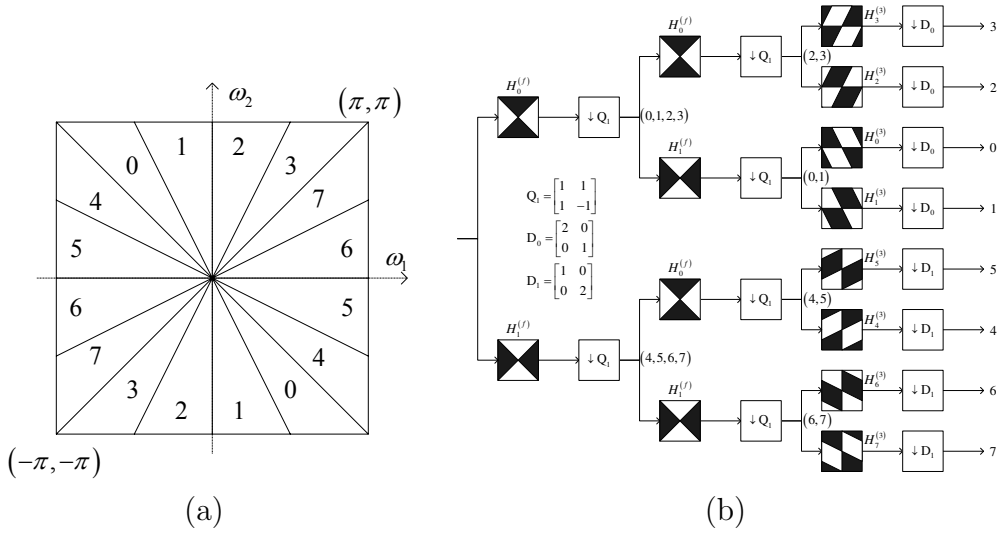
(a) how to extract the dominant features such as an edge from both magnitude and phase of transform coefficients and

(b) how to efficiently model an image in order to extract valuable image information for the purpose of analysis, e.g. image restoration, classification and segmentation.

The image retrieval problem has recently become more important and necessary because of the rapid growth of multimedia databases and digital libraries. Different search engines use different features to retrieve an image. In Chapter 2 we discuss how magnitude and phase information of the pyramidal dual-tree directional filter bank (PDTDFB) coefficients can be used to classify texture images. Texture segmentation is also an important application for pattern recognition and image analysis. It has been studied intensively and many different features have been proposed to be used as attributes in segmentation. Chapter 4 discusses how to extract a new feature for texture segmentation from relative phase of complex coefficients in the complex wavelet domain.

Many applications in image processing such as image compression and image denoising can benefit from a statistical model to characterize the image in the transform domain. A clean, precise probability model which can sufficiently describe typical images becomes essential. In Chapter 3, the complex Gaussian scale mixture (CGSM) model for

the PDTDFB coefficients is proposed and used for image denoising through Bayes least squares estimator.



(a) (b)

Figure 1.1. (a) The frequency supports of the 8-band DFB, and (b) the binary tree structure of the DFB in [1]. The black and white regions signify the passband and stopband of the 2-D filters used in the tree structure.

It is well known that an advantage of the complex wavelets compared to the real-valued wavelets is that they contain both magnitude and phase information. The magnitude of a complex coefficient represents the strength of feature such as edges while the phase indicates its location. In many applications, only the real part or the magnitude of the complex wavelet coefficient and its statistical model is used although the phase information can hold important statistical properties of the image. The property and statistical model of phase and its benefits have not been studied deeply. An essential question raised is how to *efficiently* utilize the phase information of complex wavelet coefficients for image modeling. In Chapter 5, we develop a new approach to exploit the statistical properties from the phase information of two-dimensional complex wavelet coefficients. A new probability density is derived for the relative phase distribution in the complex wavelet domain with the assumptions of Gaussian model as well as Gaussian scale mixture model in Chapter 6. The parameters of the new model are estimated using the maximum-likelihood

estimator. The new relative phase model is more accurate than the popular von Mises and Wrapped Cauchy distributions.

1.2 Literature Review

In this section, the literature review on complex directional wavelet transforms and statistical modeling of images is presented. Among many directional transforms, the PDTDFB and the uniform curvelet transform which are complex transforms are chosen to be transforms of choice. The current state of the art on statistical modeling of images and the motivation for the proposed models are presented.

1.2.1 Real Directional Wavelets

1.2.1.1 The directional filter bank

The directional filter bank (DFB), of which subband partitioning is presented in Fig. 1.1(a), has been introduced by Bamberger and Smith [1]. A major property of the DFB is its ability to extract 2-D directional information of an image, which is important in image analysis and other applications. The DFB is maximally decimated and perfect reconstruction. This means that the total number of subbands' coefficients is the same as that of the original image, and they can be used to reconstruct the original image without error.

The DFB can be implemented by a tree structure consisting of three levels of two-band systems as illustrated in Fig. 1.1(b). Each level can be implemented by using separable polyphase filters, which make the structure very computationally efficient. However, there are several difficulties in applying the DFB in image analysis. Its first drawback, as one can see in Fig. 1.1(a), is the way the low frequency band is divided. All subbands meet at DC and thus the DC component of the image spreads to all of the directional subbands. The other constraints are due to the fact that the filter bank (FB) is maximally decimated. It has been pointed out that the main problem of the discrete wavelet transform

(DWT) in image analysis and pattern recognition is that the transform is not translation invariant [3]. The DFB, similar to other biorthogonal transforms, is a translation-variant decomposition. Moreover, based on the theory of multirate filter bank, one can show that the FB does not have the permissibility property [4], which means that the Fourier transforms of the directional filters have peaks in the stopband regions. This problem seriously affects the directionality of the filters' spatial impulse responses.

1.2.1.2 Discrete contourlet transform

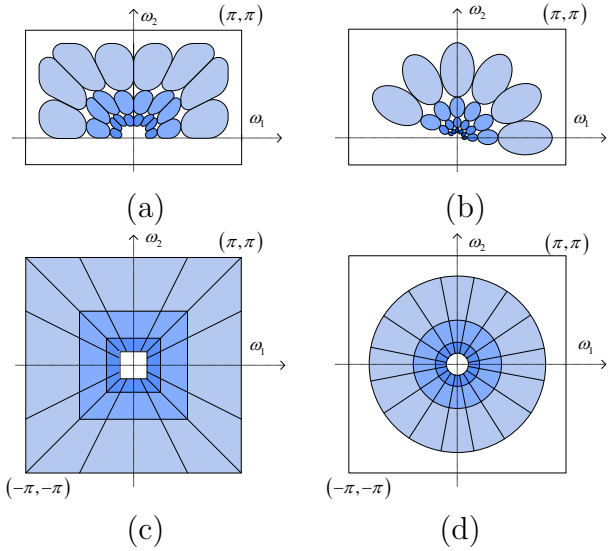


Figure 1.2. Frequency partition of directional filter banks: (a) PDTDFB ($S = 3$; $K = 8$), (b) Gabor wavelets ($S = 4$; $K = 6$) , (c) contourlet transform ($S = 3$; $K = 8$) and (d) steerable pyramid ($S = 3$; $K = 8$).

Discrete contourlet transform or pyramidal DFB is a combination of a Laplacian pyramid and a DFB [5]. Bandpass images from the Laplacian pyramid are fed into a DFB so that directional information can be captured. The low frequency component is separated from the directional components. After decimation, the decomposition can be reiterated with the same DFB in the lowpass band to form a pyramid structure. The contourlet transform provides a multiscale directional decomposition. Fig. 1.2(c) presents the frequency decompositions by the contourlet transform with $S = 3$ and $K = 8$. We use

$s \in \{1, 2, ..., S\}$ to indicate the scale index and $k \in \{1, 2, ..., K\}$ to indicate the orientation index with $K = 2^n$. Its redundancy ratio is less than 4/3 because the directional subbands are also decimated.

1.2.1.3 Steerable pyramid

Steerable pyramid [6] is another multiscale multi-directional representation. The basic filters are translations and rotations of a single function except for the inner lowpass subband and the outer residual subband. Furthermore, a filter at any orientation can be computed as a linear combination of the basic filters. The image is decomposed into one decimated lowpass subband and a set of undecimated directional subbands. This decomposition is reiterated in the lowpass subband. Thus the steerable pyramid is constructed as a recursive pyramid. Because the directional subbands are undecimated, there are $4K/3$ times as many coefficients in the representation as in the original image. Fig. 1.2(d) shows the frequency domain decomposition performed by the steerable pyramid transform.

1.2.2 Complex Directional Wavelets

1.2.2.1 Gabor wavelets

A generic two dimensional Gabor function and its Fourier transform are given by:

$$g(x, y) = (\frac{1}{2\pi\sigma_x\sigma_y})exp\left[-\frac{1}{2}(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}) + 2\pi jWx\right], \tag{1.1}$$

$$G(u, v) = exp\left[-\frac{1}{2}\left(\frac{(u - W)^2}{\sigma_u^2} + \frac{v^2}{\sigma_v^2}\right)\right], \tag{1.2}$$

where $\sigma_u = 1/(2\pi\sigma_x)$ and $\sigma_v = 1/(2\pi\sigma_y)$ are the bandwidths of the filter and W is the central frequency. To obtain a Gabor filter bank with K orientations and S scales, this function is dilated and rotated as follows

$$g_{mn}(x, y) = a^{-m}g(x', y'), \tag{1.3}$$

$$x' = a^{-m}(xcos\theta + ysin\theta), \tag{1.4}$$

$$y' = a^{-m}(-xsin\theta + ycos\theta), \tag{1.5}$$

where $\theta = n\pi/K$, n = 1, 2,..., K, and m = 0,1,...,S-1. Given a certain number of scales and orientations, the scaling factor a and bandwidths of the filters are chosen to ensure that the half-peak magnitude supports of the frequency responses touch each other as shown in Fig. 1.2(b). In [7], the authors found that a Gabor filter bank with 4 scales and 6 orientations at each scale performed best. In this case, the scaling factor a is 2. After decomposing an image, 24 subbands are formed as follows:

$$subI_k(x, y) = g_{mn}(x, y) * I(x, y), \tag{1.6}$$

where $I(x,y)$ is the input image, k = 1,2,...,24, n = 1,2,...,6, m = 0,...,3, and $*$ is 2D convolution. As one can see, this is an overcomplete representation with a redundant ratio of $KS = 24$.

1.2.2.2 The dual-tree complex wavelet transform

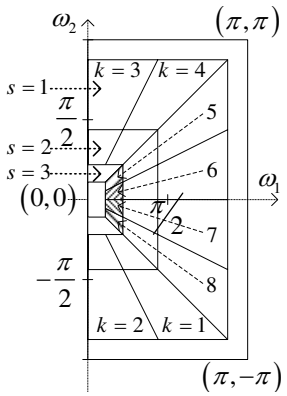


Figure 1.3. The essential frequency supports of the complex filters in the three-level eight-band PDTDFB decomposition.

Kingsbury proposed the dual-tree complex wavelet transform (DTCWT) which provides both shift-invariance and good directional selectivity [8, 9]. The filters employed in the two trees are designed in such a way that the aliasing in one branch in the first tree is approximately canceled by the corresponding branch in the second tree. The design of

the DTCWT in [9] focuses on the design of a two-channel FB having filters satisfying the half-sample phase delay condition with respect to an existing FB. Directional selectivity is achieved in the 2-D case by combining the outputs of the FB in such a way that the equivalent complex filters have supports in only one quadrant of the frequency plane [8]. This method of increasing directionality is equivalent to attaching quadrant FBs [10] to the highpass subband of the DWT. However the DTDWT can have only six directional subbands. We refer to [3] for an excellent tutorial overview on the DTCWT.

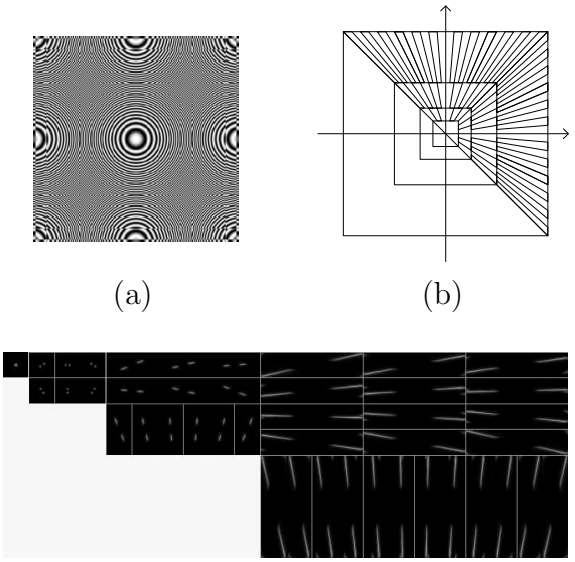
1.2.2.3 The pyramidal dual-tree directional filter bank

Recently, a novel filter bank named pyramidal dual-tree directional filter bank (PDTDFB) is proposed in [11, 2]. Figs. 1.2(a) and 1.3 show an example of the PTDFB decomposition when $S = 3$ and $K = 8$. The FB provides a shifttable multiresolution, multidirectional decomposition of two-dimensional (2-D) signal. One resolution level of the PDTDFB consists of a Laplacian pyramid and a pair of directional filter banks (DFBs), designated as primal and dual DFBs. The PDTDFB is a perfect reconstruction filter bank, which means that the synthesis side of the FB implements an inverse operator for the decomposition of 2-D signals done by the analysis FB. In Chapter 2, how the dominant features and useful information can be explored from the PDTDFB coefficients will be investigated.

1.2.2.4 The uniform curvelet transform

A modified multiresolution and multidirectional discrete transform [12] that will be used in the experiments in Chapters 5 and 6 borrows the ideas from two recently introduced discrete transforms, which are the PDTDFB [11, 2] and the fast discrete curvelet transform [13]. The transform is essentially a filter bank implementation in the frequency domain.

Similar to the PDTDFB, the directional filters have a one-side support in the frequency domain, making the subband coefficients complex. This discrete transform also has some similarities to the wrapping-based fast discrete curvelet transform (FDCT) [13]



(a) (b)

(c)

Figure 1.4. (a) The zoneplate image, (b) Frequency support of the curvelet functions, and (c) Magnitude of complex coefficients.

in the sense that both are defined based on windowing in the DFT domain. The main difference is that in the wrapping-based FDCT, the redundancy of the transform is reduced by wrapping the frequency domain of the subbands, while in this implementation, the redundancy is reduced by decimating the subbands by diagonal integer matrices. By this construction, the modified curvelet basis functions are located on a uniform integer grid at each resolution, while the basis of the FDCT is located on a non-integer grid. Fig. 1.4 is an example of the zoneplate image decomposed by our discrete curvelet transform. The decomposition has three directional scales, with $N = 6$, 12 and 24 at three resolutions. Fig. 1.4(c) shows the magnitude of the complex coefficients in transform domain.

1.2.3 Statistical Modeling of Images

The majority of applications in image processing such as denoising and compression relies on statistical models to characterize images because exact mathematical models are unfeasible or too complex. Natural images basically have common characteristics and occupy an extremely tiny space of all images. The key of statistical representation is to

capture these characteristics in a small number of parameters. A clean, precise probability model which can sufficiently describe typical images become essential.

Images have been modeled using the wavelet transform [14]. It was considered to be a good decorrelator for an image and the wavelet coefficients within a subband were assumed to be independent and identically distributed. With this assumption, the structure of transform coefficients is simple to model and the natural images are modeled by the marginal wavelet model [15] whose distribution is a two-parameter generalized Gaussian density (GGD). GGD is a suitable member of the class of peaky and heavy-tailed non-Gaussian distribution for modeling the marginal behavior of the wavelet coefficients. Although this marginal wavelet model is significantly more powerful than the classical Gaussian model [16], some important characteristics of the image are not captured by the marginal models because they do not take into account the interdependencies between different subbands of a given image and between neighbor coefficients within a subband. A recent work has found that the amplitude of coefficients of similar position, orientation and scale are highly correlated [17].

A number of researcher have developed joint statistical models in wavelet domain [18, 19, 20, 21]. The higher order dependencies may be modeled by hidden Markov model which is a simple parametric model for local dependencies with a set of hidden random variables that govern the parameters. Wavelet hidden models have proven to be useful for statistical image processing. The hidden Markov model captures the key attributes of the jointly non-Gaussian statistics of the wavelet coefficients of a typical image. In [18], the Markovian dependencies between the hidden state variables are introduced to characterize the key dependencies between the wavelet coefficients and to match the non-Gaussian distribution. This model is a useful tool in a number of applications. The major drawbacks of the hidden Markov model based on the orthogonal wavelets are their limited ability in capturing directional information, shift-variant and the expensive computations for iterative training. In [19], the authors develop a model for the neighborhoods of oriented pyramid coefficients based on a Gaussian scale mixture (GSM) assumption which is the

product of a Gaussian random vector and an independent hidden random scalar multiplier. This model can account for both marginal and pairwise joint distributions of wavelet coefficients. The GSM model suggests a hidden Markov structure for natural images with two-state discrete multipliers, corresponding to a finite mixture of Gaussian [22].

The GSM model [22] is used to model groups of neighborhood wavelet coefficients. The components of a GSM vector are highly dependent, but the dependency between coefficients decreases as their spatial separation increases. Thus, a tree GSM model that specifies probabilistic relations between the multipliers is used. The transform coefficients are linked indirectly by their shared dependency on the hidden multipliers.

The coefficients within each local neighborhood around a reference coefficient of a subband are characterized by GSM model

$$\mathbf{x} \stackrel{\Delta}{=} \sqrt{z}\mathbf{u}, \tag{1.7}$$

$$p_{\mathbf{x}}(\mathbf{x}) = \int p(\mathbf{x}|z)p_z(z)dz, \tag{1.8}$$

$$p_{\mathbf{x}}(\mathbf{x}) = \int \frac{\exp\left(-\mathbf{x}^H(z\mathbf{C_u})^{-\mathbf{1}}\mathbf{x}\right)}{(\pi)^N|z\mathbf{C_u}|}p_z(z)dz, \tag{1.9}$$

where $\mathbf{x}$ is a GSM random vector, $\mathbf{u}$ is a zero-mean Gaussian random vector, and $\mathbf{C_u} = E[\mathbf{u}\mathbf{u}^H]$ is the covariance matrix of $\mathbf{u}$. $\sqrt{z}$ is known as an independent positive scalar random variable which has density $p(z)$ and N is the dimensionality of $\mathbf{x}$ and $\mathbf{u}$.

The GSM density is symmetric and zero-mean, and has heavier tails than the Gaussian density. An important property of the GSM model is that the density of $\mathbf{x}$ is Gaussian when conditioned on z. The normalized vector $\mathbf{x}/\sqrt{z}$ is Gaussian. The probability density of the multiplier $p_z(z)$ can be Gaussian, generalized Gaussian, lognormal density or can be found by using the maximum likelihood approach from an observed set of neighborhood vectors,

$$\widehat{p}_z(z) = \arg\max_{p_z(z)} \sum_{m=1}^{M} \log_e \int_0^{\infty} p(\mathbf{x}_m|z)p_z(z)dz, \tag{1.10}$$

where the sum is over the neighborhoods. The GSM model describes the shape of wavelet coefficient marginal distributions and the correlation between the amplitudes of neighbor

coefficients. This model can account for both marginal and pairwise joint distributions of wavelet coefficients. The GSM estimator is more complex than Gaussian or wavelet marginal but it is significantly better both visually and in terms of mean squared error in the denoising application [17]. However, the GSM model is used for modeling real coefficients.

Inspired by this approach, we propose a complex Gaussian scale mixture (CGSM) for modeling the complex coefficients in Chapter 3. The CGSM [23] is applied to image denoising through the Bayes least squares estimator.

1.3 Thesis Outline

In Chapter 2, we propose a new feature for texture image retrieval. We develop image denoising algorithm in the PDTDFB domain based on the CGSM in Chapter 3. Another application of the PDTDFB is texture segmentation. A new feature extraction method is proposed for texture segmentation in Chapter 4. Properties and statistics of relative phase of the complex coefficients are studied in detail. We proposed the von Mises and wrapped Cauchy distributions for the modeling of the relative phase in complex wavelet domain in Chapter 5. With the assumption of the Gaussian image model as well as the GSM in real wavelet domain, the marginal and joint distributions of phases of the complex coefficients in complex wavelet domain are studied, and from these hypotheses, we then derive the probability density function for the relative phase in the complex wavelet domain in Chapter 6. Summary and future works are listed in Chapter 7.

CHAPTER 2
COMPLEX DIRECTIONAL FILTER BANK FOR TEXTURE IMAGE RETRIEVAL

2.1 Introduction

The image retrieval problem has recently become more important and necessary because of the rapid growth of multimedia databases and digital libraries. Content-based image retrieval is currently an attractive research area. This chapter discusses how the magnitude and phase information of the recently proposed complex directional filter bank (CDFB) [24] coefficients can be used to classify texture images.

One of the approaches to texture feature extraction is the filter bank approach that decomposes a texture image into subbands using a linear transform or a filter bank (FB). In [25], texture classification performances of various FB methods are compared, and the conclusion is that no method performs well in all kinds of textures. Several previous works extract texture features based on wavelet packet signatures [26] and tree-structured wavelet transform [27]. Although these methods allow for a multiresolution decomposition, they are limited in directional selectivity and may not be suitable for images with geometric information such as textures.

The 2-D Gabor transform [28, 29, 30, 31] is a directional decomposition that optimally achieves joint resolution in space and spatial frequency. The Gabor filters are widely used in feature extraction for processing texture images [7]. It is the standardized method for feature extraction in estimating the 'Homogeneous Texture Descriptor' in MPEG-7 [32]. In [7], the retrieval performance of the Gabor wavelet and other multiresolution techniques such as wavelet transform and tree-structured wavelet transform are compared. Previous studies indicate that the Gabor feature yields the best correct texture retrieval rate [7]. Despite the high performance compared to others, the 2-D Gabor transform produces an overcomplete representation for images and is very computationally intensive. The Gabor

wavelets are not orthogonal. They are typically defined in frequency domain and can be viewed as bandpass filters. There is redundancy of information with a ratio of KS in the filtered images, where S is the number of scales and K is the number of orientations per scale.

Another multiresolution multi-directional image representation method is the steerable pyramid [6]. The basis filters of the steerable pyramid are translations and rotations of a single function, and a filter at any orientation can be computed as a linear combination of the basis filters. At the same scale and position, the power of the coefficients corresponding to different orientations is invariant to rotations of the input signal. This property can be used in rotation invariant texture recognition [33] and retrieval [34]. The steerable pyramid can be designed to produce any number of undecimated directional subbands, but this decomposition is significantly overcomplete by a factor of $4K/3$, where K is the number of orientations.

Many other multiresolution multi-directional image representations like the octave-band DFB [35], the multiscale DFB [36], the contourlet [37] and the complex wavelets [38, 39], also have been used to extract feature vectors. As compared to the conventional real wavelet transform, the main advantage of the complex wavelets is the shiftable property which means that the energy of the output is shift invariant. This produces a texture feature which is more robust to translation in the image [38]. The 2-D complex wavelet with an overcomplete ratio of four is less redundant than the Gabor transform and the steerable pyramid. However, the complex wavelet transform can be designed to produce any number of scales but the number of orientations is limited to six.

In most of the above classification methods, though some of the filters are complex, only the real part or the magnitude of the transform coefficients is used in texture discrimination. The phase information has not been explicitly utilized in image processing applications. One of the earliest works that points out the importance of phase information is [40], with the famous example where the main image structure is reconstructed by using only the phase of the Fourier coefficients. The phase holds crucial information

about image structures and features. The image features such as edges and shadows are determined by analyzing the phase of the harmonic components [41] or computing the phase congruency [42]. The phase congruency matrix provides a quantitative measure of the significance of the edge at each pixel of the image, and yields high quality in edge detection. However, its usage in classification is not efficient because the phase congruency matrix has the same size as the image.

Some previous works have used the Gabor phase in image classification applications such as the iris and palmprint identification [43, 44] and the face recognition [45]. These methods are based on the quadrant bit coding (QBC) extracted from the complex Gabor coefficients. Each pixel in each subband image will be encoded to two bits according to the quadrant in which the Gabor phase angle lies. The feature vector is created from these coding bits over all subband images and the classification is achieved by the Hamming distance. In the iris and palmprint identification [43, 44], the feature vector is chosen to have the length of 256 bytes. To model the QBC based on Gabor phase pattern more efficiently, in [45], the spatial histogram has been proposed to model the encoded Gabor phases. Although this classification method achieves high performance in face recognition, it also has the high-dimensional histogram features.

Some other applications exploit the local phase information across scales of the complex wavelet such as the description of image texture [46], the detection of blurred images [47] and object recognition [48]. In [46] the local phase behavior is captured by the cross-correlation of complex coefficients of bands at adjacent scales (fine-scale and coarse-scale), and this statistical measurement distinguishes edges from lines and helps in representing gradients due to shading and lighting effects. In [47], the local phase coherence relationship across scale and space has been suggested and the phases of the finest scale coefficients can be well predicted from those of the coarser scale coefficients. The disruption of this local phase is an effective measure for blur detection. This across scale relationships are also captured using modified product of coefficients at adjacent scales and has been used in [48]. Another investigation of local phase in the same orientation and the same

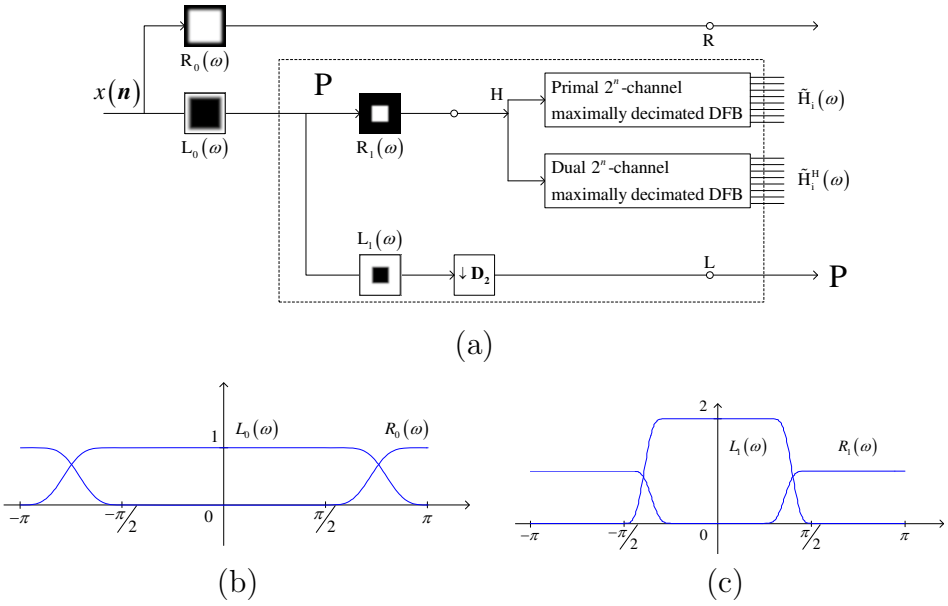


Figure 2.1. (a) The FB implemented the pyramidal CDFB image representation. Block P is reiterated to create multi-level decomposition. Slices of the 2-D frequency responses of: (b) $R_0(\boldsymbol{z})$ and $L_0(\boldsymbol{z})$, and (c) $R_1(\boldsymbol{z})$ and $L_1(\boldsymbol{z})$.

scale is based on the dual-tree complex wavelet transform [49]. The feature orientations are determined by the phase difference between the adjacent coefficients in six fixed directional subbands. The local phase has been used in some applications such as the description of image texture [46], the detection of blurred images [47], the object recognition [48] and the face recognition [50]. Another investigation of local phase is the SameLevel Product transform (SLP) [49]. The SLP transform extracts phase information from the complex wavelet transform. All of the above approaches exploit phase information of the complex wavelet for some applications. However, how to use the phase feature for image retrieval have not been investigated. In this chapter, we introduce how to use phase information of the CDFB coefficients and its applications in texture image retrieval [51, 24].

This chapter has three main contributions. First, a modified image decomposition for feature extraction based on CDFB is proposed. The CDFB decomposition is based on the analysis side of the pyramid dual-tree directional filter bank proposed in [11]. Since the purpose of the CDFB decomposition is for feature extraction, perfect reconstruction condition is not required. By combining magnitude and phase information of the CDFB

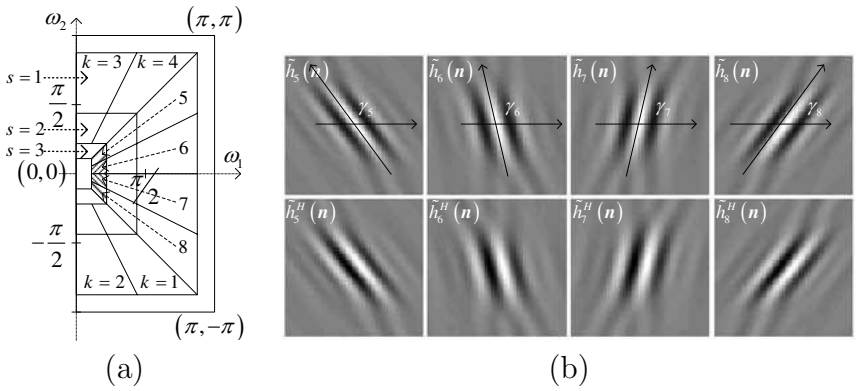
coefficients, a feature vector formed by relative phase of the CDFB coefficients (CDFB-RP) is proposed to discriminate texture images. How phase information can be incorporated explicitly to improve the classification rate is discussed. Finally, a comparison of five different feature extraction schemes in texture retrieval is presented. We show that our proposed CDFB-RP can achieve higher retrieval accuracy than other transforms, while requiring much less storage memory and computation time.

In the next section, the CDFB representation for feature extraction is described. We showed the linear relationship between the RP of the complex coefficients and the angle of dominant orientation of texture image in section 2.3. The procedure to retrieve and classify texture images and the experimental results are presented in section 2.4. We discuss the results and conclude the chapter in section 2.5.

2.2 The Complex Directional Filter Bank

The (energy) shiftable complex directional pyramid is a new image decomposition, which is recently introduced in [2]. The CDFB pyramid provides a shiftable multiresolution, multidirectional decomposition of two-dimensional (2-D) signals. It is a perfect reconstruction filter bank, which means that the synthesis side of the FB is the inverse operator for the decomposition of 2-D signals done by the analysis FB. In this section, we described the CDFB decomposition for feature extraction based on the analysis side of the CDFB pyramid. The FB used in this work to provide the CDFB decomposition does not have the residual highpass subband as in the original CDFB pyramid. Another difference is that the designed FB used in this work is not constrained by the perfect reconstruction condition.

The CDFB, shown in Fig. 2.1(a), is an interactive multiscale and multidirectional FB. Each resolution level consists of a two-channel FB and a pair of directional filter banks (DFBs). The purpose of the CDFB is to provide a shiftable and scalable multiresolution directional decomposition. This transform has some similarities with the complex version of the shiftable pyramid [6], while maintaining a much lower redundant ratio.

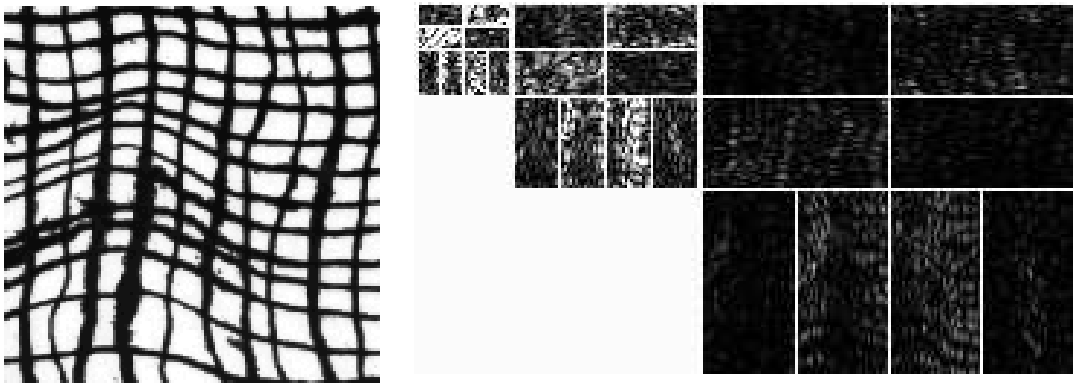


(a) (b)

Figure 2.2. The essential frequency supports of (a) the complex filters in the three-level eight-band CDFB decomposition, and (b) corresponding spatial impulse responses.

According to Fig. 2.1(a), the input image is first filtered by the lowpass filter $L_0(\boldsymbol{z})$, while the narrow band highpass information (R_0) is disregarded, before passing through the first level of a multiresolution FB. This two-channel FB has two filters, highpass $R_1(\boldsymbol{z})$ and lowpass $L_1(\boldsymbol{z})$. Slices of the frequency responses of these filters at $\omega_2 = 0$ are illustrated in Fig. 2.1(b) and (c). The high frequency component at the output of the filter $R_1(\boldsymbol{z})$ is then decomposed by the pair of primal and dual DFBs, resulting in the highest resolution directional decomposition. The low frequency component, after decimation by $\boldsymbol{D}_2 = 2\boldsymbol{I}$, is fed into the second level decomposition for the second resolution. The block P shows one level of the CDFB, where the 2×2^n decimated outputs of the two DFBs are the real and imaginary parts of 2^n complex-valued subbands. Fig. 2.3 shows an example of the CDFB decomposition when $S = 3$ and $K = 8$. For more details of the construction of the CDFB, the reader is referred to [2].

The most important property of the CDFB is that all complex directional subbands are shiftable in the sense that there is no significant aliasing in the decimated complex subbands. Therefore, each complex directional subband provides a shiftable description of image in a specific scale and direction. Note that this description is also very parsimonious because the decimation ratio for the subband is increasing with the number of directions and with the higher scale (lower resolution). An example of the frequency supports of a three-level eight-band CDFB is shown in Fig. 2.2.



(a) (b)

Figure 2.3. An example of the CDFB decomposition: (a) a 128×128 image extract from the D104 texture in Brodatz database, and (b) the CDFB decomposition with three level and eight directional band. Subimages are amplitude of complex-valued subband coefficients.

By construction of the CDFB, each pair of corresponding directional filters has the Hilbert transform relationship [2]. Therefore, the equivalent directional complex filter for each subband has a one-sided frequency support, as illustrated in Fig. 2.2(a). In spatial domain, the real part of the complex filter is symmetric while the imaginary part is anti-symmetric as in Fig. 2.2(b). The amplitude and phase information of a complex coefficient provides local information on the directional feature of the image at a specific scale and direction. The objective of the next section is to understand the relation between the phase information and a typical edge so that it can be added to the feature vector.

2.3 Relative Phase of the CDFB Coefficient

2.3.1 Linear Relationship between Local Phase and Distance

It has been stated in [46] that the local phase varies linearly with the distance from features and in [49], the authors also have observed that the phase of a 1-D dual-tree DWT coefficient is consistently linear with respect to the feature offset (distance to the step). However, the proof for this relationship has not been given. Here, we will show that the phase in the vicinity of the feature such as a step or a ramp has a linear relationship with the distance.

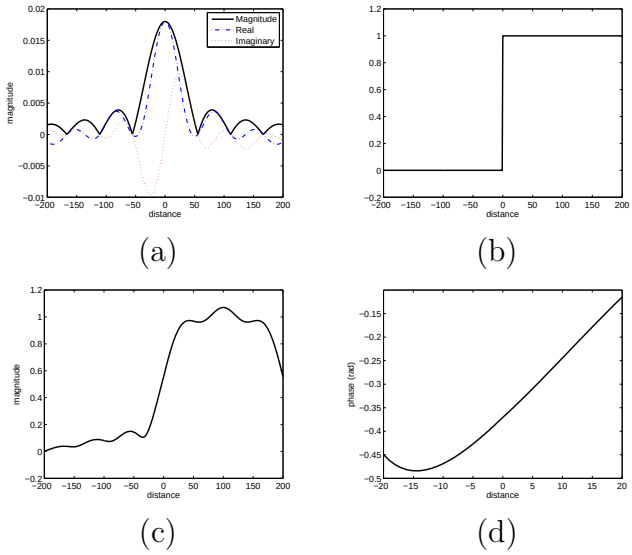


Figure 2.4. Linear relationship between local phase and distance. (a) Complex filter $h(t)$, (b) step function $x(t)$, (c) magnitude of $y(t)$ and (d) phase of $y(t)$ in the vicinity of the step (near t = 0).

Let the input $x(t) = u(t)$ be a unit step signal as shown in Fig. 2.4(b), and the ideal complex filter $h(t)$ (Fig. 2.4(a)) has one-sided frequency support $H(\omega)$ as:

$$H(\omega) = \begin{cases} 1 & 0 \leq \omega_1 \leq \omega \leq \omega_2, \\ 0 & \text{otherwise.} \end{cases}$$

The output $y(t)$ can be expressed as:

$$\begin{aligned} y(t) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} H(\omega)X(\omega)e^{j\omega t}d\omega = \frac{1}{2\pi}\int_{\omega_1}^{\omega_2}\frac{1}{j\omega}e^{j\omega t}d\omega, \\ &= \frac{1}{2\pi}\int_{\omega_1}^{\omega_2}\frac{1}{j\omega}\sum_{n=0}^{\infty}\frac{(j\omega t)^n}{n!}d\omega, \\ &= \frac{1}{2\pi j}\left(\ln\left|\frac{\omega_2}{\omega_1}\right| + \sum_{n=1}^{\infty}\frac{(jt\omega_2)^n - (jt\omega_1)^n}{n!n}\right). \end{aligned}$$

The magnitude and phase of $y(t)$ are shown in Figs. 2.4(c) and (d), respectively. Because the output $y(t)$ is considered in the vicinity of the step, i.e. $|\omega t| \ll 1$, hence $y(t)$ can be approximated by:

$$y(t) \approx \frac{1}{2\pi j}\left(\ln\left|\frac{\omega_2}{\omega_1}\right| + jt(\omega_2 - \omega_1)\right),$$

and the phase of $y(t)$ can be approximated by:

$$\begin{aligned}\angle y(t) &\approx \tan^{-1}\left(\frac{\omega_2-\omega_1}{\ln|\frac{\omega_2}{\omega_1}|}t\right)-\frac{\pi}{2},\\ &= \tan^{-1}\left(\frac{\omega_2(1-\frac{\omega_1}{\omega_2})}{\ln|(\frac{\omega_1}{\omega_2})^{-1}|}t\right)-\frac{\pi}{2}.\end{aligned}$$

Let $p = 1-\frac{\omega_1}{\omega_2}$. Hence $p \in [0\ 1]$. In addition, we have the inequality: $1-p \le e^{-p}$ for $0 \le p \le 1$. Therefore,

$$\begin{aligned}\frac{\omega_2(1-\frac{\omega_1}{\omega_2})t}{\ln|(\frac{\omega_1}{\omega_2})^{-1}|} &= \frac{\omega_2 pt}{\ln|(1-p)^{-1}|},\\ &\le \frac{\omega_2 pt}{\ln|e^p|} = \omega_2 t \ll 1,\end{aligned}$$

and the phase of $y(t)$ can be approximated by:

$$\angle y(t) \approx \left(\frac{\omega_2-\omega_1}{\ln|\frac{\omega_2}{\omega_1}|}\right)t-\frac{\pi}{2}.$$

Similarly, let the input $x(t) = tu(t)$ be a unit ramp signal and the ideal complex filter $H(\omega)$ be defined as above. In this case, the phase of $y(t)$ can be approximated by:

$$\begin{aligned}\angle y(t) &\approx \tan^{-1}\left(\frac{\omega_1\omega_2\ln|\frac{\omega_2}{\omega_1}|}{\omega_2-\omega_1}t\right),\\ &\approx \left(\frac{\omega_1\omega_2\ln|\frac{\omega_2}{\omega_1}|}{\omega_2-\omega_1}\right)t.\end{aligned}$$

From above formula, we can see that the phase of $y(t)$ in the vicinity of the features such as the steps or ramps (at $t = 0$) is linear with t (the distance to feature). Fig. 2.4(d) shows this linear relationship near the step at $t = 0$.

2.3.2 Relative Phase Feature for Image Retrieval

$$p_{sk}(i,j) = \begin{cases}\angle y_{sk}(i,j) - \angle y_{sk}(i,j+1) \text{ if } 1 \le k \le \frac{K}{2},\\ \angle y_{sk}(i,j) - \angle y_{sk}(i+1,j) \text{ if } \frac{K}{2} < k \le K.\end{cases} \tag{2.1}$$

We denote γ_k as the center angle of the CDFB subband k. The angle difference between two consecutive DFB subbands is approximated $\frac{\pi}{K}$. Thus subband k contains directional information at angles $\theta_k = \gamma_k + \alpha$, where $-\frac{\pi}{2K} < \alpha < \frac{\pi}{2K}$.

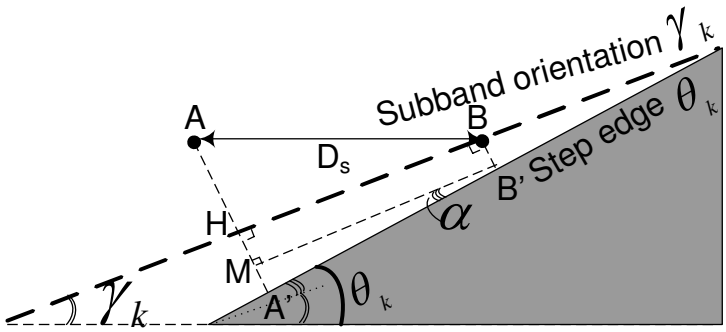


Figure 2.5. Relationship between the angle θ_k of an edge and the distances from two horizontally adjacent coefficients located at A and B to the edge in the direction normal to the subband orientation k ($1 \leq k \leq \frac{K}{4}$) at some arbitrary scale.

Let us consider an edge at angle θ_k in the supported region of subband k with $1 \leq k \leq \frac{K}{4}$. In this case, the center angle of subband γ_k is an acute angle. Assume that the two horizontally adjacent coefficients A and B are located in the neighborhood of an edge as shown in Fig. 2.5. AA' and BB' represent the distances from A and B to the edge in the direction normal to subband orientation, respectively. The distance between A and B at scale s is $D_s = 2^s$. We determine the angle of the edge θ_k by determining α in terms of AA' and BB':

$$\begin{aligned} \tan\alpha &= \frac{MA'}{MB'} = \frac{AA' - HM - AH}{MB'}, \\ &= \frac{AA' - BB' - D_s \sin\gamma_k}{D_s \cos\gamma_k} = \frac{AA' - BB'}{D_s \cos\gamma_k} - \tan\gamma_k. \end{aligned}$$

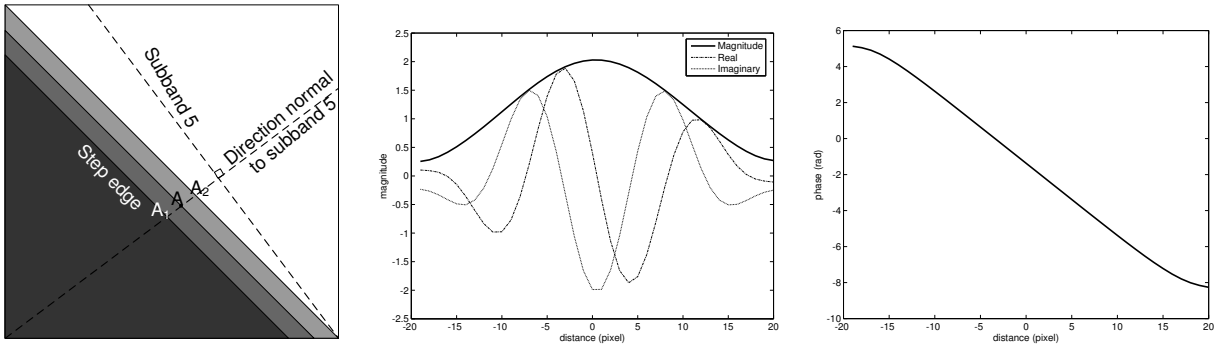
If $K \geq 8$, $|\alpha| \leq \frac{\pi}{16}$ and $\alpha \approx \tan\alpha$. Hence, the feature orientation θ_k can be approximated by:

$$\theta_k \approx \gamma_k - \tan\gamma_k + \frac{AA' - BB'}{D_s \cos\gamma_k}. \tag{2.2}$$

Similarly, the feature orientation θ_k of the other subbands can be approximated by:

$$\theta_k \approx \begin{cases} \gamma_k - \tan\gamma_k - \frac{AA'-BB'}{D_s \cos\gamma_k} & \text{if } \frac{K}{4} < k \leq \frac{K}{2}, \\ \gamma_k + \cot\gamma_k + \frac{AA'-BB'}{D_s \sin\gamma_k} & \text{if } \frac{K}{2} < k \leq \frac{3K}{4}, \\ \gamma_k + \cot\gamma_k - \frac{AA'-BB'}{D_s \sin\gamma_k} & \text{if } \frac{3K}{4} < k \leq K. \end{cases} \tag{2.3}$$

In the 2-D case, the behavior of the phase is a straight forward extension of the 1-D case as shown in 2.3.1, and the proof is omitted due to limited space. Figs. 2.6(b)



(a) (b) (c)

Figure 2.6. Translation of an edge from A_1 to A_2 for subband $s = 3$ and $k = 5$ when $\theta_k = 135°$: (a) edge translation and coefficient at A for different positions of the edge, (b) magnitude and (c) phase.

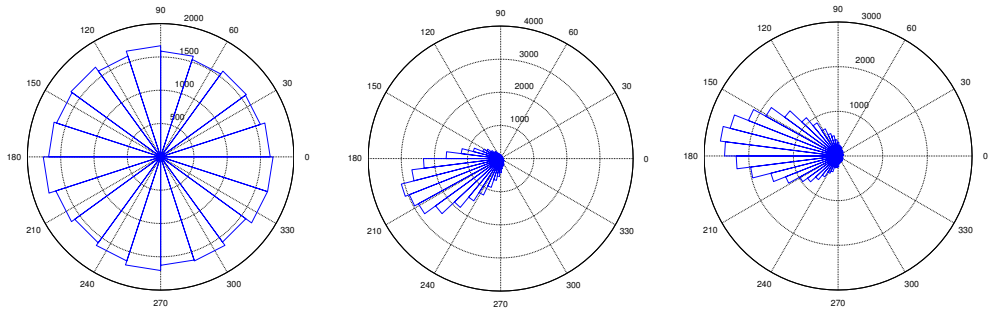
and (c) show the magnitude and phase of a CDFB coefficient at A (Fig. 2.6(a)) when the edge (angle $\theta_k = 135^0$) is translated from A_1 to A_2. The x-axis represents the translation distance of the edge in horizontal direction. When A lies on the edge, $x = 0$. The distance from A to the edge is approximately $x \sin \theta_k$. We observe that when the edge moves from A_1 to A_2, the corresponding phase at A (Fig. 2.6(c)) varies linearly with respect to the distance to the edge in the direction normal to the subband orientation (γ_k). Hence, the phase at A can be estimated by:

$$\angle y_{sk}(A) = a_{sk} AA' + b_{sk}. \tag{2.4}$$

The slope a_{sk} and the intercept b_{sk} are constants for each scale s and orientation k. Therefore the term $(AA' - BB')$ in (2.2) and (2.3) can be computed from the difference of the phases at A and B:

$$AA' - BB' = \frac{\angle y_{sk}(A) - \angle y_{sk}(B)}{a_{sk}}, \tag{2.5}$$

when a_{sk} can be estimated from the phase data in Fig. 2.6(c). From (2.2), (2.3) and (2.5), we can conclude that the feature orientation θ_k of all CDFB subbands is linearly proportional to the RP $(\angle y_{sk}(A) - \angle y_{sk}(B))$. Because the RP can represent a dominant direction within a directional subband, we use it as a feature in the texture image retrieval application.



(a) (b) (c)

Figure 2.7. The phase histogram at scale $s = 1$ and orientation $k = 6$ for: (a) the CDFB phases of image 'Bark.0000', (b) the RP phases of image 'Bark.0000' and (c) the RP phases of image 'Metal.0002'.

Fig. 2.7(a) shows the uniform distribution of CDFB phases of the texture 'Bark.0000' in the VisTex collection at the finest scale $s = 1$ and orientation $k = 6$. This distribution can not inform us any information of the edges, while the distribution of RPs has a particular direction as in Fig. 2.7(b). The circular mean m^c of RPs as later defined in (4.4) determines the mean direction of the dominant orientations θ_k in subband k, and the circular variance σ^c in (2.10) determines the measure of dispersion for these dominant orientations. Note that the RP distributions of different images have different parameters m_c and σ_c. The RP distribution of the texture 'Metal.0002' is in Fig. 2.7(c) with $m_c = 2.99$ and $\sigma_c = 0.33$. These two parameters are used to form the RP feature vector in the next section.

2.4 Classification Method and Numerical Experiments

In this section, a detailed comparison of five different feature extraction schemes in texture retrieval is presented. Compared with Gabor filters, the CDFB is a compromise between two competing requirements for image representation: shiftability and low redundancy of information. We also include two other multiresolution directional decompositions in feature extraction. One is the contourlet decomposition [52], which is very similar to the CDFB in terms of directionality of the filters and implementation. The other is the steerable pyramid [6], whose highpass directional filters are steerable.

2.4.1 Texture Image Database and Feature Database

We select two groups of image textures for our experiment. The texture database used in the first experiment contains 116 texture images from the Brodatz album [53, 25] and the USC database, which was used in [7]. The second group of textures has 40 images from the VisTex databases used in [34]. Each of these 512×512 images is divided into sixteen 128×128 non-overlapping sub-images, thus creating a database of 1856 texture samples in the first experiment, and 640 in the second one. Each original image is treated as a single class and therefore there are 16 samples from each of the 116 (or 40) classes. To reduce the intensity correlation, all images are normalized to have zero mean and unit variance. For each image in the database, all the four decompositions are applied. The RP matrix of each CDFB subband is created as in (4.1), and their corresponding feature vectors are computed.

2.4.2 Texture Feature Extraction

Each image in the database is applied to the following four decompositions: the steerable pyramid, the contourlet transform, the 2-D Gabor transform and the CDFB. The Gabor wavelet is applied with four scales and six orientations per scale, while the other transforms have three scales of eight orientations. For each subband, the mean and standard deviation of the absolute values of the coefficients are calculated as follows:

$$m(y) = \frac{1}{L} \sum_{i,j} |y(i,j)|, \tag{2.6}$$

$$\sigma(y) = \left(\frac{1}{L} \sum_{i,j} (|y(i,j)| - m(y))^2 \right)^{\frac{1}{2}}, \tag{2.7}$$

where $y_k(i,j)$ are the coefficients and L is the number of coefficients. Since there are twenty-four subbands, feature vectors are constructed by $m(y)$ and $\sigma(y)$ as

$$f_y = [m_1(y), m_2(y), \ldots m_{24}(y), \sigma_1(y), \sigma_2(y), \ldots, \sigma_{24}(y)]. \tag{2.8}$$

To construct the feature vector for CDFB-RP, the RP matrix of each CDFB subband is created as in (4.1).

The mean of the magnitudes of the CDFB coefficients, the circular mean and the circular variance of the RPs which are computed by:

$$m^c(y) = \tan^{-1}\left(\frac{\sum_{i,j}\sin p(i,j)}{\sum_{i,j}\cos p(i,j)}\right), \text{ and} \tag{2.9}$$

$$\sigma^c(y) = 1 - \frac{\sqrt{\left(\sum_{i,j}\sin p(i,j)\right)^2 + \left(\sum_{i,j}\cos p(i,j)\right)^2}}{L}. \tag{2.10}$$

are used to form the CDFB-RP feature vector, where $p(i,j)$ is the $(i,j)^{th}$ element in the RP matrix and L is the number of elements. In order to obtain a feature vector which has the same number of features as the Gabor or other feature vectors, the CDFB-RP feature vector is formed by twenty-four means of the magnitudes of the CDFB coefficients (all subbands), sixteen circular means of RPs (from finest subbands) and eight circular variances of RPs (from finest subbands):

$$f_y = [m_1(y), \ldots, m_{24}(y), m_1^c(y), \ldots, m_{16}^c(y), \sigma_1^c(y), \ldots, \sigma_8^c(y)].$$

2.4.3 Distance Measure and Query Processing

The query pattern can be any one of the texture patterns from the image databases. Let f_x and f_y be two feature vectors obtained from one of the four transforms. The distance between them is defined by

$$d(f_x, f_y) = \sum_{k=1}^{24}\left(\left|\frac{m_k(x) - m_k(y)}{\alpha(m_k)}\right| + \left|\frac{\sigma_k(x) - \sigma_k(y)}{\alpha(\sigma_k)}\right|\right), \tag{2.11}$$

where $\alpha(m_k)$ and $\alpha(\sigma_k)$ are the standard deviations of $m_k(\cdot)$ and $\sigma_k(\cdot)$ of the entire database. Let f_x and f_y be two CDFB-RP feature vectors. The distance between them is defined as

$$\begin{aligned} d(f_x, f_y) = &\sum_{k=1}^{24}\left|\frac{m_k(x) - m_k(y)}{\alpha(m_k)}\right| + \sum_{k=1}^{16}\left|\frac{m_k^c(x) - m_k^c(y)}{\alpha(m_k^c)}\right| \\ &+ \sum_{k=1}^{8}\left|\frac{\sigma_k^c(x) - \sigma_k^c(y)}{\alpha(\sigma_k^c)}\right|, \end{aligned} \tag{2.12}$$

where $\alpha(m_k)$, $\alpha(m_k^c)$ and $\alpha(\sigma_k^c)$ are the standard deviations of $m_k(\cdot)$, $m_k^c(\cdot)$ and $\sigma_k^c(\cdot)$ of the entire database.

Table 2.1. Average retrieval accuracy of 116 texture images in the Brodatz database

	Gabor	**Contour**	**Steer**	**CDFB**	**CDFB-RP**
N = 15	74.01	69.37	71.23	72.60	**75.23**
N = 40	83.78	81.21	82.54	83.10	**86.46**
N = 80	88.17	86.79	87.96	87.90	**90.84**

Table 2.2. Average retrieval accuracy of 40 texture images in the Vistex database

	Gabor	**Contour**	**Steer**	**CDFB**	**CDFB-RP**
N = 15	80.81	75.45	74.65	79.29	**82.26**
N = 40	91.09	88.46	87.94	90.40	**92.07**
N = 65	93.89	92.32	91.47	93.32	**94.69**

For each query image, N nearest neighbors are selected, and the number of these textures belonging to the same class as the query texture, except for itself, is counted. This number (less than or equal to fifteen) divided by fifteen is defined as the retrieval rate. The performance of the entire class is obtained by averaging this rate over the sixteen members which belong to the same class of texture. The average of all classes is the overall performance of the transform.

2.4.4 Experimental Results

Table 2.1 summarizes the overall retrieval rates using different directional transforms for the first experiment. If only the top 15 texture images that are nearest to the search texture are considered, the CDFB-RP gives the best overall retrieval performance of 75.23 %. The CDFB and the Gabor wavelet are at 72.6 % and 74.01 %, while the contourlet and the steerable pyramid are at 69.37 % and 71.23 %. Fig. 2.8 shows the overall performances for the case of N from 20 to 100. It is clear that the CDFB-RP is consistently better than all other feature vectors. Fig. 2.10 shows an example of the images browsed using the CDFB-RP features.

In the second experiment with 40 VisTex textures, the retrieval rates for the five directional decompositions are in Table 2.2. Fig. 2.9 shows the overall performances for

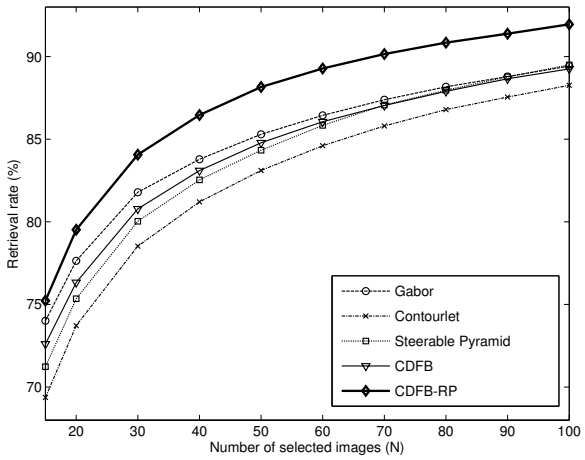


Figure 2.8. Average retrieval rate of Brodatz database according to the number of top images considered.

the case of N from 20 to 65. We observe that with $N = 15$, the overall retrieval performance of the CDFB-RP is highest at 82.26%. The performances of the CDFB (79.29%) is very close to that of the Gabor wavelet (80.81%), while they are significantly higher than those of the contourlet (75.45%) and steerable pyramid (74.65%).

In order to evaluate the efficiency of the overcomplete representations in estimating the feature vectors of the textures, we compare the CPU time used to calculate the feature vectors (Table 2.3). The computation is done in MATLAB. As we can see, the Gabor features take the longest time to compute, which is 0.274 second. However, it can be reduced if we decimate the lowpass subband of the Gabor filters. The CDFB-RP features take 0.047 second, which is much less time than Gabor. The feature extraction time for CDFB-RP is approximately equal to 0.048 second for the steerable pyramid feature, but more time than the contourlet features (0.025 s) and the CDFB features (0.036 s).

The proposed relative phase features can be applied to other complex wavelets such as the complex Gabor filter bank. The Gabor-RP feature achieves the performance slightly lower than that of the CDFB-RP for texture image retrieval. However, the 2-D Gabor transform produces an over complete representation for images and is very computationally intensive. Therefore, the Gabor phase features requires more computation time than the CDFB phase features, as shown in Table 2.3.

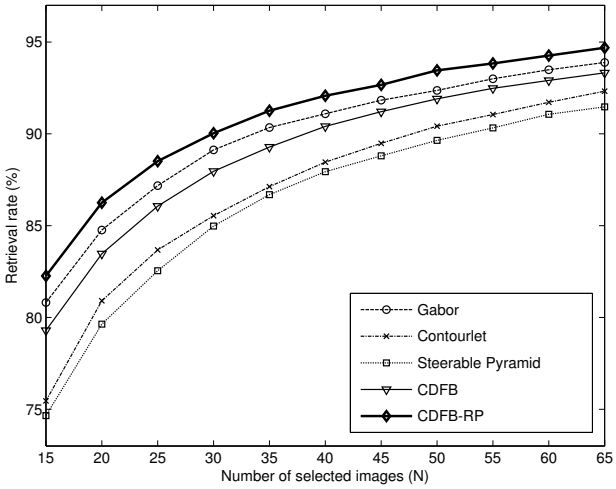


Figure 2.9. Average retrieval rate according to the number of top images considered when the database is 40 VisTex textures.

Table 2.3. Feature vector length, redundant ratio and feature extraction time for the various texture features

	Gabor	**Contour**	**Steer**	**CDFB**	**CDFB-RP**
Feature vector length	48	48	48	48	48
Redundant ratio	24	4/3	32/3	8/3	8/3
Feature extraction time (s)	0.274	0.025	0.048	0.036	0.047

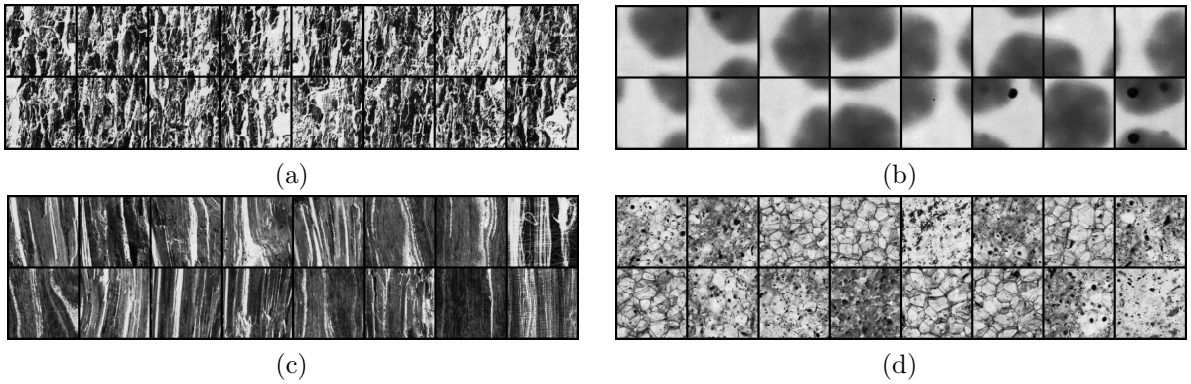


Figure 2.10. Examples of texture images retrieved using CDFB-RP features. The average retrieval rate is shown in the parentheses and the first image represents the query image: (a) D12 (96.25 %), (b) D48 (98.75 %), (c) D72 (43.33 %), (d) D73 (42.08 %). Although D72 and D73 have the average retrieval performance low, the incorrect images look similar to the query image.

2.5 Summary

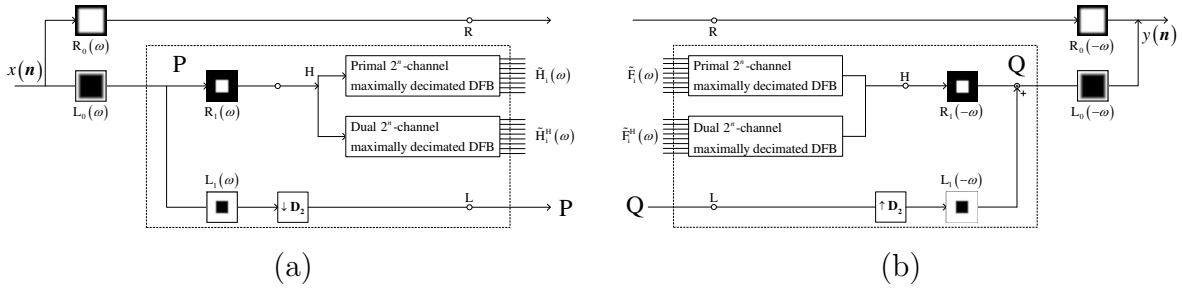
A new image feature, which we call CDFB-RP, is proposed for feature extraction. The feature is calculated based on the CDFB decomposition and the phase difference between complex coefficients. It is shown that this phase difference has a linear relationship with the angle of dominant orientation within a subband. The CDFB-RP takes advantage of this attractive property to characterize the orientations of the images. Moreover, the CDFB has several nice properties such as multiscale and multi-directional decomposition, shiftable subband, and efficient implementation. Finally, the CDFB-RP feature extraction scheme is employed in texture image retrieval. The phases make a great contribution to the success of the new feature extraction scheme. Compared to other magnitude based methods such as the 2-D Gabor wavelet, the contourlet, the steerable pyramid and the CDFB, the CDFB-RP yields best overall performance in classification rate while keeping the complexity relatively low. The overcomplete ratio of the CDFB-RP is bounded by 8/3 which is much less than those of the Gabor wavelet ($KS = 24$) and the steerable pyramid ($4K/3 = 32/3$), and is only twice of that of the contourlet (4/3). A fast decomposition structure and low redundancy make CDFB-RP more efficient in searching and browsing texture images.

CHAPTER 3
IMAGE DENOISING USING SHIFTABLE DIRECTIONAL PYRAMID AND SCALE MIXTURE OF COMPLEX GAUSSIAN

3.1 Introduction

Many applications in image processing such as image compression, image denoising can benefit from a statistical model to characterize the image in the transform domain. A clean, precise probability model which can describe sufficiently typical images becomes essential. In this chapter, the complex Gaussian scale mixture model for complex coefficients of the pyramidal dual-tree directional filter bank (PDTDFB [2]) is proposed for image denoising through Bayes least squares estimator.

There are many works on the statistics of decomposition coefficients of the wavelet transform [14]. The wavelet coefficients within a subband were often assumed to be independent and identically distributed. With this assumption, the wavelet coefficients can be modeled by the marginal model whose distribution is a two-parameter generalized Gaussian density (GGD). The GGD is very suitable distribution for the peaky and heavy-tailed non-Gaussian statistic of typical image wavelet decomposition. Although this wavelet marginal model is more powerful than the Gaussian model, it does not take into account the dependencies between different subbands as well as between a coefficient and its neighboring coefficients of the same subband. A number of researchers have developed successfully joint statistical models in wavelet domain [18, 22, 54]. The Hidden Markov Tree was introduced in [18] to model the wavelet decomposition. A bivariate probability density function has been proposed to model the statistical dependencies between a wavelet coefficient and its parent [54]. In [22], the author developed a model for neighborhoods of oriented pyramid coefficients based on a Gaussian scale mixture (GSM) which is the product of a Gaussian random vector and an independent hidden random scalar multiplier. This model can account for both marginal and pairwise joint distributions of wavelet coefficients. The GSM



(a) (b)

Figure 3.1. A shiftable pyramid [2]. (a) The analysis side, and (b) Synthesis side. Similar P and Q blocks can be reiterated at lower scale to decompose an image into a multiscale representation.

estimator is more complex than Gaussian or wavelet marginal estimators but it provides significantly higher performances in terms of both mean squared error and visual quality in image denoising [17].

In image analysis applications, an overcomplete, multiresolution and multidirectional representation usually provides much better performance compared to the critical representation such as the discrete wavelet transform (DWT). Examples of multiscale and multidirectional decompositions include Gabor filter bank, steerable pyramid [6], complex wavelet, contourlet transform [5] and the PDTDFB decomposition [2]. In this chapter, we develop image denoising algorithm in PDTDFB domain based on the complex Gaussian scale mixture (CGSM) [23].

3.2 Scale Mixture of Complex Gaussian Distributions

3.2.1 Complex Gaussian Distributions

A complex Gaussian distribution is defined as follows [55][56]:

$$f(\mathbf{u}) = \frac{\exp\left(-\mathbf{u}^H(\mathbf{C_u})^{-\mathbf{1}}\mathbf{u}\right)}{(\pi)^N|\mathbf{C_u}|}, \tag{3.1}$$

where $\mathbf{u} \in \mathbf{C}^{\mathbf{N}\mathrm{x}1}$ is a vector of complex stochastic variable defined as $\mathbf{u} = \left(u_1, u_2, ...u_N\right)^T$, and $u_n = x_n + jy_n$, $x_n, y_n \in R^1$, real-valued variables normally distributed. It will be assumed that $E[u_n] = E[x_n] + jE[y_n] = 0$, where $E[.]$ is expectation operator. In these

expressions, the superscript T denotes transposition, the superscript H denotes complex conjugate transposition, and $j = \sqrt{-1}$. Furthermore, $\mathbf{C_u} \in \mathbf{C}^{N\text{x}N}$ is the complex covariance matrix defined as $\mathbf{C_u} = E[\mathbf{u}\mathbf{u}^H]$. By definition, $\mathbf{C_u}$ is positive definite and Hermitian symmetric, hence, its inverse exists.

3.2.2 Complex Gaussian Scale Mixture for Complex Coefficient Model

A statistic model based on Gaussian scale mixture distribution, which is the product of a Gaussian random vector and an independent hidden random scalar multiplier, is developed in [22]. This model can account for both marginal and pairwise joint distributions of real wavelet coefficients. Inspired by this approach, we define a complex Gaussian scale mixture for modeling the complex coefficients. If two random variables (x, z) have a jointly distribution, a mixture of $f(x)$ distributions is defined [57] as follows

$$h(x) = \int f(x|z)g(z)dz, \tag{3.2}$$

where $f(x)$ and $g(z)$ are the probability density functions of x, z.

Suppose that the vector $\mathbf{u}$ has a complex Gaussian distribution and that scalar real variable $\sqrt{z}$ has some distribution on $(0, \infty)$ with a density $p_z(z)$ $(z > 0)$. We refer $\mathbf{x} \overset{\Delta}{=} \sqrt{z}\mathbf{u}$ as the scale mixtures of complex Gaussian distribution as follows

$$p_{\mathbf{x}}(\mathbf{x}) = \int p(\mathbf{x}|z)p_z(z)dz, \tag{3.3}$$

$$p_{\mathbf{x}}(\mathbf{x}) = \int \frac{\exp\left(-\mathbf{x}^H(z\mathbf{C_u})^{-\mathbf{1}}\mathbf{x}\right)}{(\pi)^N|z\mathbf{C_u}|}p_z(z)dz, \tag{3.4}$$

where $\mathbf{C_u} = E[\mathbf{u}\mathbf{u}^H]$ is the complex covariance matrix of $\mathbf{u}$, $\mathbf{u} = \left(u_1, u_2, ...u_N\right)^T$, and N is the dimensionality of $\mathbf{u}$ and $\mathbf{x}$. The density of $\mathbf{x}$ is complex Gaussian when conditioned on z, and the variable z is known as the multiplier. We assume that the coefficients $\mathbf{x} = \mathbf{x}_{real} + j\mathbf{x}_{imag}$ within each local neighborhood around a reference coefficient of a complex subband are characterized by a CGSM model. In general, the neighborhood may include coefficients from other subbands, as well as from the same subbands. Similar to the

GSM, the probability density of the multiplier $p_z(z)$ can be founded by using maximum log likelihood approach for estimating a nonparametric $p_z(z)$ from an observed set of M neighborhood vectors.

$$\widehat{p}_z(z) = \arg\max_{p_z(z)} \sum_{m=1}^{M} \log_e \int_0^{\infty} p(\mathbf{x}_m|z)p_z(z)dz. \tag{3.5}$$

The PDTDFB coefficients are linked indirectly by their shared dependency on the hidden multipliers z. Hence, the CGSM model can describe the shape of complex wavelet coefficient distributions and the correlation between neighbor coefficients. Furthermore, the CGSM model captures both magnitude and phase information of the natural image.

3.3 Image Denoising Application

3.3.1 Thresholding

3.3.1.1 Simulations

In the first set of experiments, the images are denoised by using the hard thresholding method. We decompose images into subbands using the PDTDFB as in Fig 3.1, DWT and contourlet FBs. All the three decompositions have four resolution levels, and the lowest resolution subbands (coarse signal) are kept unchanged. An orthogonal and nearly linear phase *symlet* FB of length 10 is used in the DWT decomposition. In the contoulet and PDTDFB decomposition, the two lower resolution levels are also *symlet* wavelet, while the two higher levels are directional FBs with 32 subbands and 16 subbands at the highest resolution and the next lower resolution, respectively. The added noise is Gaussian and white with variance σ^2. The threshold is set at three times of the standard deviation of the noise in the subbands. The noise variance in the wavelet subband is also σ^2, since the wavelet is orthonormal. The noise variances in the directional subbands of the PDTDFB decomposition are estimated by

$$\begin{aligned} \sigma_k^2 &= \frac{1}{4\pi^2} \int \sigma^2 \left|F_k(\boldsymbol{\omega})\right|^2 d\boldsymbol{\omega} \\ &= \sigma^2 \sum_{\boldsymbol{n}\in\mathcal{Z}^2} \left|f_k(\boldsymbol{n})\right|^2 \end{aligned} \tag{3.6}$$

where $f_k(\boldsymbol{n})$ and $F_k(\boldsymbol{\omega})$ are the spatial and frequency responses of the kth considered directional filter.

3.3.1.2 Results

We compare the PSNR values of the denoising results using different transforms with the hard thresholding method as shown in Table 3.1. The PDTDFB transform provides higher PSNR values than the wavelet and contourlet transforms for three images. It is evident that the PDTDFB is consistently better than the wavelet and contourlet transforms when the standard deviation of the input noise is varying between $\sigma = 15$ and $\sigma = 100$.

3.3.2 Bayes Least Squares Estimator

One of the best methods for image denoising is Bayes least squares estimator based on the Gaussian scale mixture model (BLS-GSM) presented in [17]. For each neighborhood, the reference coefficient at center of the neighborhood is estimated from $\mathbf{y}$, the set of observed coefficients. The subband coefficients are real numbers, and the probability density function is a function of the real variable. The BLS-GSM method is used to estimate the real subband coefficients. However, the shiftable complex directional pyramid decomposes an image into the subbands whose coefficients are complex values. Our purpose here is to develop the BLS-GSM algorithm for estimating the complex coefficients.

Let $\mathbf{y}$ be the vector corresponding to a neighborhood of N observed complex coefficients

$$\mathbf{y} = \mathbf{x} + \mathbf{w}, \tag{3.7}$$

where $\mathbf{x}$ is a original coefficient vector and $\mathbf{w}$ is a noise vector in the transform domain.

We make a simplifying assumption that $\mathbf{w}$ is a zero-mean complex Gaussian vector and $\mathbf{x}$ is a CGSM vector as shown in (3.4). It is well know that the Bayes least squares estimation is the conditional expectation when $\mathbf{x}$ and $\mathbf{y}$ are real random vectors as follows

$$\widehat{\mathbf{x}} = E[\mathbf{x}|\mathbf{y}]. \tag{3.8}$$

$$E[\mathbf{x}|\mathbf{y}] = \int_{\mathbf{0}}^{\infty} p(z|\mathbf{y}) E[\mathbf{x}|\mathbf{y}, z] dz, \tag{3.9}$$

where the scalar real variable z has some distribution on $(0, \infty)$ with a density $p_z(z)$ $(z > 0)$. When conditioned on z as shown in [17]

$$E[\mathbf{x}|\mathbf{y}, z] = z\mathbf{C_u}(z\mathbf{C_u} + \mathbf{C_w})^{-\mathbf{1}}\mathbf{y}. \tag{3.10}$$

It is possible to show that the same results (3.8), (3.9), (3.10) are valid for the complex random variables. However, the covariance matrices $\mathbf{C_u} = E[\mathbf{u}\mathbf{u}^H]$ and $\mathbf{C_w} = E[\mathbf{w}\mathbf{w}^H]$ are positive definite and Hermitian symmetric ($c_{k_1,k_2} = c^*_{k_2,k_1}$), and the density of the observed neighborhood vector conditioned on z is a zero-mean complex Gaussian, with covariance $\mathbf{C_{y|z}} = z\mathbf{C_u} + \mathbf{C_w}$

$$p(\mathbf{y}|z) = \frac{\exp\left(-\mathbf{y}^H(z\mathbf{C_u} + \mathbf{C_w})^{-\mathbf{1}}\mathbf{y}\right)}{(\pi)^N |z\mathbf{C_u} + \mathbf{C_w}|}, \tag{3.11}$$

For estimating $\mathbf{x}$, $p(z|\mathbf{y})$ as in (3.9) is computed as follows

$$p(z|\mathbf{y}) = \frac{p(\mathbf{y}|z)p_z(z)}{\int_{\mathbf{0}}^{\infty} p(\mathbf{y}|\alpha)p_z(\alpha)d\alpha}, \tag{3.12}$$

and $p_z(z) \propto 1/z$ as shown in [17] is applied to the experiments in this chapter.

3.3.2.1 Simulations

We decompose the images into subbands using a shiftable complex directional pyramid [2]. The representation consists of oriented bandpass bands at 3 scales (16 orientations in the finest scale, 8 orientations in the coarse scale and 4 orientations in the coarsest scale), highpass residual band, and one lowpass band. Each subband except for the lowpass residual band is denoised by using the BLS estimator described above. The denoised image is reconstructed from the processed subbands and the lowpass band. We assume the image is corrupted by independent additive Gaussian noise. The Lena and Barbara images of size 512×512 are used in this simulation.

Table 3.1. PSNR values of the image denoising experiments using hard thresholding method

Image	σ	**PSNR**	**DWT**	**contourlet**	**PDTDFB**
Lena	15	24.61	30.03	30.32	**31.98**
	20	22.11	28.67	29.11	**30.83**
	30	20.17	26.67	27.41	**28.95**
	50	14.15	24.05	25.10	**26.39**
	75	10.63	21.73	23.12	**24.04**
	100	8.13	19.84	21.68	**22.33**
Barbara	15	24.61	27.31	27.59	**29.32**
	20	22.11	25.70	26.38	**28.06**
	30	20.17	23.62	24.68	**26.52**
	50	14.15	21.52	22.70	**24.31**
	75	10.63	19.86	21.13	**22.45**
	100	8.13	18.31	19.99	**20.99**
Peppers	15	24.61	29.87	29.95	**31.09**
	20	22.11	28.51	28.92	**30.00**
	30	20.17	26.52	27.17	**28.46**
	50	14.15	23.87	24.94	**26.10**
	75	10.63	21.56	23.05	**23.83**
	100	8.13	19.58	21.47	**22.18**

Table 3.2. PSNR values of the image denoising experiments using Bayes least squares estimator

Image	σ	**PSNR**	**FS** [17]	**DWT**	**PDTDFB**
Lena	15	24.61	**33.90**	32.39	33.56
	20	22.11	**32.66**	31.04	32.31
	25	20.17	**31.69**	30.01	31.33
	50	14.15	**28.61**	27.01	28.04
	75	10.63	**26.84**	24.41	26.06
	100	8.13	**25.64**	23.21	24.60
Barbara	15	24.61	**31.86**	29.88	**31.86**
	20	22.11	30.32	28.24	**30.45**
	25	20.17	29.13	27.05	**29.38**
	50	14.15	25.48	23.82	**26.04**
	75	10.63	23.65	22.32	**24.16**
	100	8.13	22.61	21.44	**22.73**

Table 3.3. SSIM values of the image denoising experiments using Bayes least squares estimator

Image	σ	**SSIM**	**FS**	**DWT**	**PDTDFB**
Lena	15	0.45	**0.89**	0.86	0.88
	20	0.34	**0.87**	0.83	0.86
	25	0.27	**0.85**	0.80	0.84
	50	0.11	**0.78**	0.69	0.77
	75	0.06	**0.73**	0.59	0.71
	100	0.04	**0.69**	0.53	0.65
Barbara	15	0.58	**0.90**	0.86	**0.90**
	20	0.48	**0.87**	0.82	**0.87**
	25	0.40	0.84	0.78	**0.85**
	50	0.20	0.70	0.62	**0.73**
	75	0.11	0.61	0.53	**0.65**
	100	0.07	0.53	0.47	**0.58**

We obtain the neighborhood noise covariance $\mathbf{C_w}$ by decomposing a random noise image which has normal distribution with mean zero, standard deviation σ and dimensionality as original image in to shiftable complex pyramid subbands. This image has the same power spectrum as the noise. Given $\mathbf{C_w}$, the covariance $\mathbf{C_u}$ can be computed from the observation covariance matrix $\mathbf{C_y} = E[z]\mathbf{C_u} + \mathbf{C_w}$. Set $E[z] = 1$, hence, $\mathbf{C_u} = \mathbf{C_y} - \mathbf{C_w}$.

3.3.2.2 Results

Tables 3.2 and 3.3 show the PSNR and SSIM [58] values of the denoising results when the standard deviation of the input noise is varying between $\sigma = 15$ and $\sigma = 100$. The quality of the denoised images of our proposed denoising method are compared to those of BLS-GSM method [17] in wavelet domain including orthogonal wavelet (DWT) and full steerable pyramid (FS). Our method performs better than the orthogonal wavelet in terms of mean squared error (MSE) and perceptual image quality (SSIM) and is comparable to the steerable pyramid with several noise levels σ from 15 to 100. This is significant since the PDTDFB has much lower overcomplete ratio compared to the steerable pyramid. The PDTDFB redundancy ratio approximately 11/3, while the redundancy ratio of the

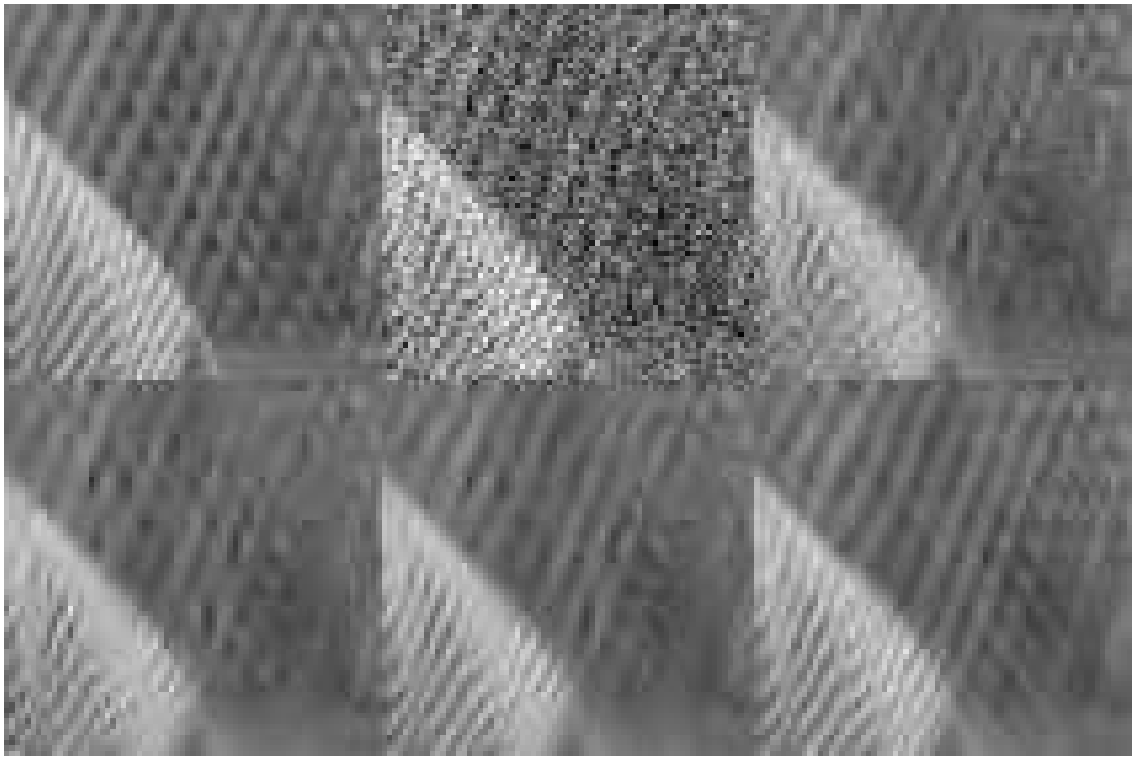


Figure 3.2. Comparison of denoising results on Barbara. From left to right and top to bottom: Original image, Noisy image ($\sigma = 25, PSNR = 20.17$), DWT($PSNR = 27.05$), UDWT ($PSNR = 28.06$), FS ($PSNR = 29.13$), and PDTDFB ($PSNR = 29.38$).

steerable pyramid with 8 orientations is 32/3. The proposed denoising method could achieve high quality image denoising, recover very fine details, e.g. texture. Figs. 3.2 and 3.3 show the zoom-in image denoising results of Barbara image with different denoising methods. In these figures, UDWT denotes the undecimated discrete wavelet transform.

3.4 Summary

The modified version of the PDTDFB has been proposed for image denoising. The shiftable properties of this approximately tight-frame decomposition benefits the image denoising application. In comparison to the existing transforms including the wavelet, the contourlet, the PDTDFB yields the best image denoising performance with the thresholding method. Although the overcomplete ratio of the PDTDFB is much lower than this of steerable pyramid, by combining the CGSM model with BLS estimator, PDTDFB could

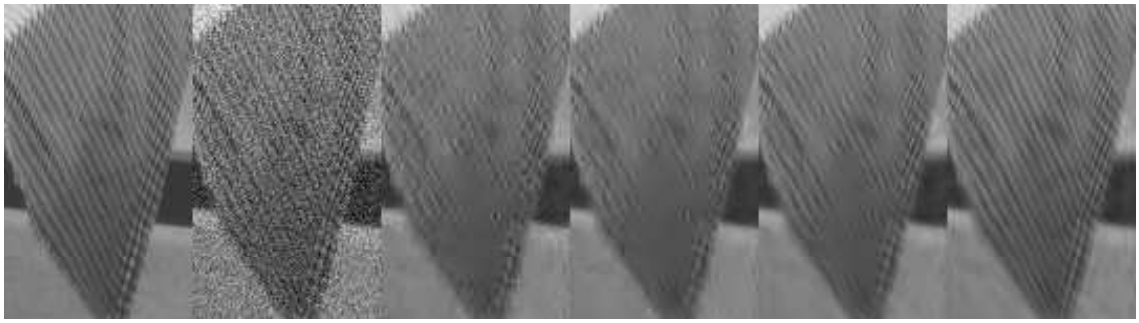


Figure 3.3. Comparison of denoising results on Barbara. From left to right: Original image, Noisy image($\sigma = 25, PSNR = 20.17$), DWT ($PSNR = 27.05$), UDWT ($PSNR = 28.06$), FS ($PSNR = 29.13$), and PDTDFB ($PSNR = 29.38$).

achieve the denoised image quality comparable to steerable pyramid with the BLS-GSM algorithm.

CHAPTER 4
USING PHASE AND MAGNITUDE INFORMATION OF THE COMPLEX DIRECTIONAL FILTER BANK FOR TEXTURE SEGMENTATION

4.1 Introduction

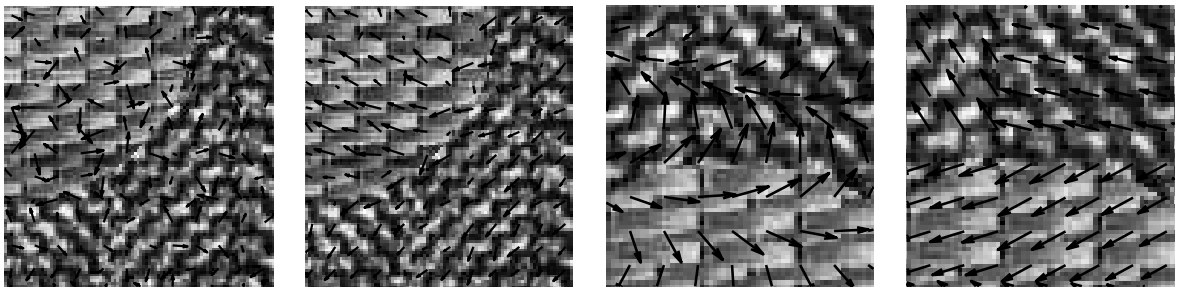
Texture segmentation is an important application for pattern recognition and image analysis. It has been studied intensively and many different features have been proposed to be used as attributes in segmentation. This chapter discusses how magnitude and phase information of the complex directional filter bank (CDFB) [2] coefficients can be used to segment texture images.

One of the common approach in texture segmentation is to use a filter bank to decompose a texture image into subbands [59, 25]. The subband images, after some nonlinear processing, are then used to form local features in segmentation. The Gabor filters are widely used in feature extraction for texture segmentation [30, 31, 60, 61]. In [30], the complex Gabor filters are used to extract features in order to find the boundaries between textures by comparing the channel magnitude responses and detecting discontinuities from large variations in channel phase responses. In [31] a fixed set of Gabor filters are proposed for texture segmentation. The filter selection method is based on reconstruction of the input image from the subband images. Texture classification performances of many different multi-channel filtering approaches have been compared in [25], and the conclusion is that no method performs well in all kinds of textures. The wavelet frame [62] and quadrature mirror filter bank (QMF) were two of the best filters. Textures are also characterized by features extracted from images modeled in multiresolution domain as Markov random field [63], hidden Markov tree [64] and coupled Markov random field [65]. Some other works characterize texture by its rotation invariant roughness using fractal dimension [66], adaptive scale fixing [67] and local spectral histogram [68] .

In most of the above segmentation methods, though the phase holds the crucial information about image structures and features [40], only the real part or the magnitude of the transform coefficients is used in texture discrimination. The image features such as edges and shadows are determined by analyzing the phase of the harmonic components [41] or computing the phase congruency [42]. Some other applications exploit the local phase information across scales of the complex wavelet such as the description of image texture [46], the detection of blurred images [47] and object recognition [48]. Another investigation of local phase in the same orientation and the same scale is based on the dual-tree complex wavelet transform [49].

Recently, the complex directional filter bank is proposed for texture image retrieval [24]. When both magnitude and phase information in the complex coefficients are used, the classification system has the best performance compared to other magnitude based methods. The phase information is used successfully in texture image retrieval. In this chapter, we discuss how to develop the magnitude and phase based feature extraction method which is proposed in [24] to make it suitable for texture segmentation [69].

4.2 Complex Directional Filter Bank and Local Mean Direction Feature for Texture Segmenatation



(a) (b) (c) (d)

Figure 4.1. D17D55 image cropped. (a) CDFB phase ($s = 2, k = 2$), (b) CDFB-LMD ($s = 2, k = 2$), (c) CDFB phase ($s = 5, k = 3$), and (d) CDFB-LMD ($s = 5, k = 3$).

Since the phase difference of two neighboring coefficients or relative phase (RP) can represent a dominant direction θ_k within a directional subband, it is used in the texture

image retrieval application. For more details and the cases of other values of k, the reader is referred to Chapter 2.

Note that the RP feature proposed in Chapter 2 for image retrieval are the circular mean of RPs which determines the global mean direction of the dominant orientations θ_k in the whole subband k and the circular variance determines the measure of global dispersion for these dominant orientations. Since the feature vector in Chapter 2 represents the global information, we are not able to use this RP feature to discriminate the pixels in the segmentation problem. In this section, we discuss how to develop the magnitude and phase based feature extraction method to make it suitable for texture segmentation.

The RP of the CDFB at the spatial location (k_1, k_2) of scale s and orientation k is calculated from the corresponding subband coefficients as

$$p_{sk}(k_1,k_2) = \begin{cases} \angle y_{sk}(k_1,k_2) - \angle y_{sk}(k_1,k_2+1) \text{ if } 1 \le k \le \frac{K}{2}, \\ \angle y_{sk}(k_1,k_2) - \angle y_{sk}(k_1+1,k_2) \text{ if } \frac{K}{2} < k \le K, \end{cases} \tag{4.1}$$

where $y_{sk}(k_1, k_2)$ is the subband complex coefficient at position (k_1, k_2), $s = 1, 2, ..., S$ and $k = 1, 2, ..., K$. Note that, among the K orientations, the first half ($1 \le k \le \frac{K}{2}$) are more horizontal (vertical in frequency domain) and thus the RP is calculated in the vertical direction whereas the other half ($\frac{K}{2} < k \le K$) are more vertical and it is calculated in the horizontal direction.

Consider when $1 < k \le \frac{K}{4}$. Suppose that there is an edge of angle θ_k at position (k_1, k_2). This angle can be estimated from the RP $p_{sk}(k_1, k_2)$ by

$$\theta_k \approx \gamma_k - \tan\gamma_k + \frac{p_{sk}(k_1,k_2)}{a_{sk} D_s \cos\gamma_k}, \tag{4.2}$$

where γ_k is the center angle of the subband, and a_{sk} is a constant estimated in the neighborhood of (k_1, k_2). In our segmentation algorithm, the RP $p(k_1, k_2)$ is replaced by the local mean direction (LMD) of the RPs of the coefficients in the local window, i.e.

$$\theta_k \approx \gamma_k - \tan\gamma_k + \frac{\hat{p}_{sk}(k_1,k_2)}{a_{sk} D_s \cos\gamma_k}, \tag{4.3}$$

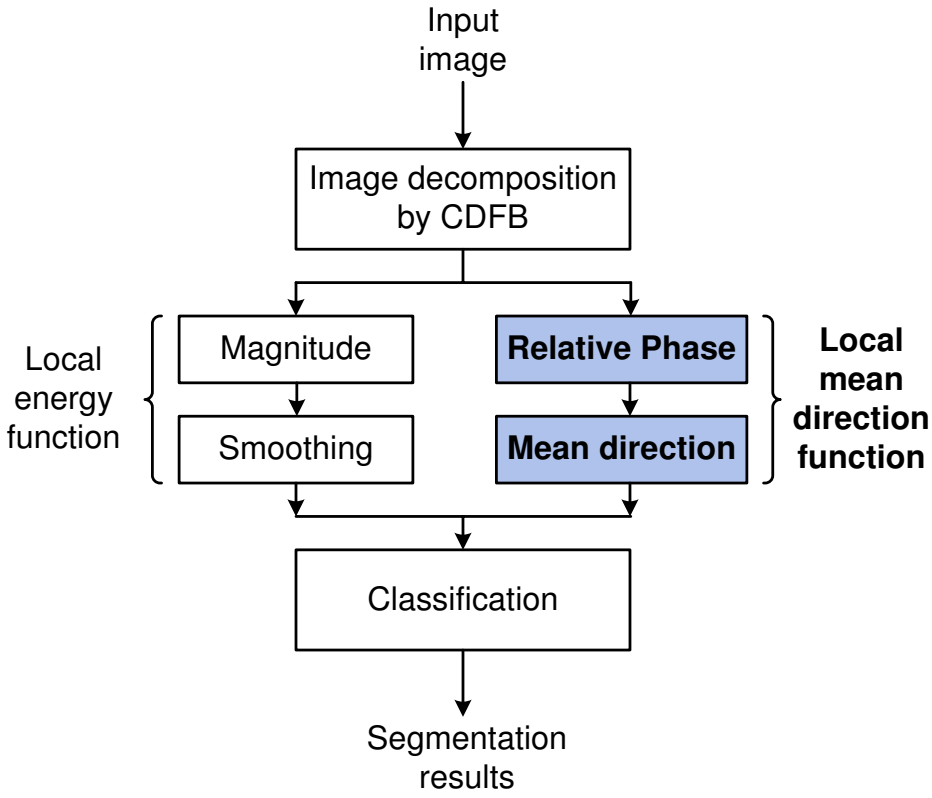


Figure 4.2. The block diagram for the CDFB-LMD-based segmentation.

where $\hat{p}_{sk}(k_1, k_2)$ is the LMD at (k_1, k_2) which is defined by

$$\hat{p}_{sk}(k_1, k_2) \quad = \quad \arctan \frac{\sum_{k_1,k_2 \in W} \sin p(k_1, k_2)}{\sum_{k_1,k_2 \in W} \cos p(k_1, k_2)}. \tag{4.4}$$

The LMD is also computed by (4.4) for all other values of k. We can see that the angle of the edge can be estimated from the RP $p_{sk}(k_1, k_2)$ in the neighbor of an edge. Hence the RP $p_{sk}(k_1, k_2)$ is approximately constant in the vicinity of an edge, and when evaluated in directional subbands, the LMDs can be used to identify the edge direction. The quality of texture segmentation depends on the window size. As the large window is used, the sufficient amount of information is captured. However, a small window is necessary to accurately locate the boundaries between texture regions. This suggests that the selection of window size could possibly be based on the contents of the image. The images with larger texture would require larger window sizes whereas finer textures would require smaller windows. In our experiments, the 5×5 window is used for the texture images and the 3×3 window is used for the real images.

Fig. 4.1 illustrates the difference between the phase and the LMD. Each image is a combination between two different types of textures (D17 and D55). In Figs. 4.1(a) and (c), each arrow represents the magnitude (length) and phase (angle) of the CDFB coefficient, while in Figs. 4.1(b) and (d), the angle represents the estimated angle calculated from the LMDs above. It is clear that the distinction of the LMDs between the two types of textures D17 and D55 can be captured and recognized in Figs. 4.1(b) and (d) while no clear distinction is obtained from the phase of CDFB coefficients.

Since the LMD determines the local mean direction of the dominant orientations θ_k in the small window, it provides local information of the subband images and can discriminates the pixels. Therefore, the LMD will be used as an additional feature for texture segmentation in the next section.

4.3 Proposed Segmentation Method

The framework of the proposed method is shown in Fig. 4.2. A texture image is first decomposed by the CDFB. The complex-valued subband images obtained from the CDFB are then used to extract local magnitude and phase features. For magnitude, a local smoothing window of size $M \times M$ is applied to the magnitude of the complex coefficients. Therefore, each pixel of the output of the window represents an average magnitude of the $M \times M$ block. For phase, a local mean direction at a pixel is determined by the phases of the complex coefficients in the local $M \times M$ block.

The classification is tested on two groups of synthetic texture images from the Brodatz album [53] in our experiments. The texture images used in the first experiment contains fifteen images, each composing of two different types of textures. The second group of textures consists of five images where each has five types of textures. The size of these images is 256×256. For each experiment, each image in the database is decomposed by the following three decomposition methods: the 2-D Gabor transform, the CWT and the CDFB. The numbers of scales and orientations are chosen to give the best results. The Gabor transform is applied with six scales and eight orientations per scale while the CWT

has six scales of six orientations (Recall that the number of orientations of the CWT is fixed to six.), and the CDFB has five or six scales of eight orientations.

The local energy is computed over a window by

$$e_k(k_1, k_2) = \frac{1}{N} \sum_{(m,n)\in W} w_{mn} |y_k(k_1 - m, k_2 - n)|, \tag{4.5}$$

where N is the number of elements in the window W. In our experiments, the running average window is used although any other type of windows such as the Gaussian weighted window can also be applied [31]. In texture segmentation, neighboring pixels are very likely to belong to the same texture category. We include the spatial coordinates of the pixel as additional features as proposed in [31] for the five-texture image segmentation. For the case of two-texture image segmentation, the spatial coordinates are not used because the same performance is obtained when adding these two coordinate features. The LMD features of the subbands are estimated by (4.4).

When using filter banks, the most frequent segmentation method is to search for class prototypes in the feature space by clustering the feature vectors according to certain rules. Each image pixel is classified by determining the class prototype which is closest to its feature vector. In our experiments, we use the standard K-means clustering algorithm used in [62]. Before clustering each feature is normalized to have a zero mean and a constant variance. Some other complex classification methods may also be applied to further improve the segmentation results. However, this is beyond the scope of the work and might smear the effect of the LMD feature.

4.4 Simulation Results

Table 4.1 summarizes classification error rates using different directional transforms for the fifteen two-texture images. The average error rates of the methods after using phase are consistently much lower than when using magnitude alone. The combination of CDFB and LMD (CDFB-LMD) yields the best overall segmentation performance of 2.0%. This corresponds to an improvement of 35.4% from using only magnitude. The improvements

after phase incorporation for the cases of the Gabor transform (Gabor-LMD) and the CWT (CWT-LMD) are 27.7% and 31.3% respectively.

In the second experiment with five five-texture images, the average classification error rates for the CDFB and CDFB-LMD are summarized in Table 4.2. The average error rate for the case of CDFB-LMD is 2.6%, much lower than that of the case of CDFB which is 4.7%. In this case, the improvement of the CDFB-LMD after using phase information is 46%. Fig. 4.3 shows percentage errors of each of the five five-texture images for CDFB and CDFB-LMD. It is evident that the CDFB-LMD yields a nice improvement compared to the CDFB which is based on magnitude information alone. Using this classification result, the segmentation results for the five images are shown in Fig. 4.4. It is clear that the phase information can complement to magnitude information because higher classification accuracy and better boundaries are achieved. Fig. 4.5 shows the segmentation results of some real images. The color images are converted to the gray-scale images before segmentation.

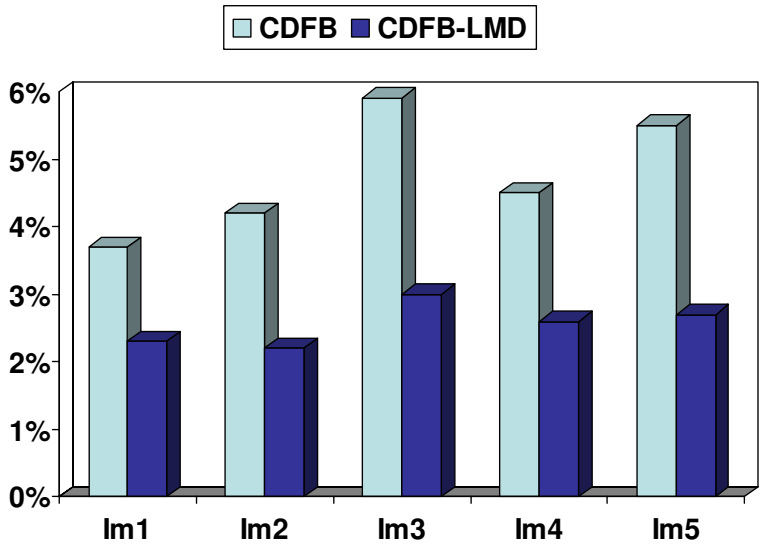


Figure 4.3. Percentage errors for CDFB and CDFB-LMD in segmentation of five-texture images.

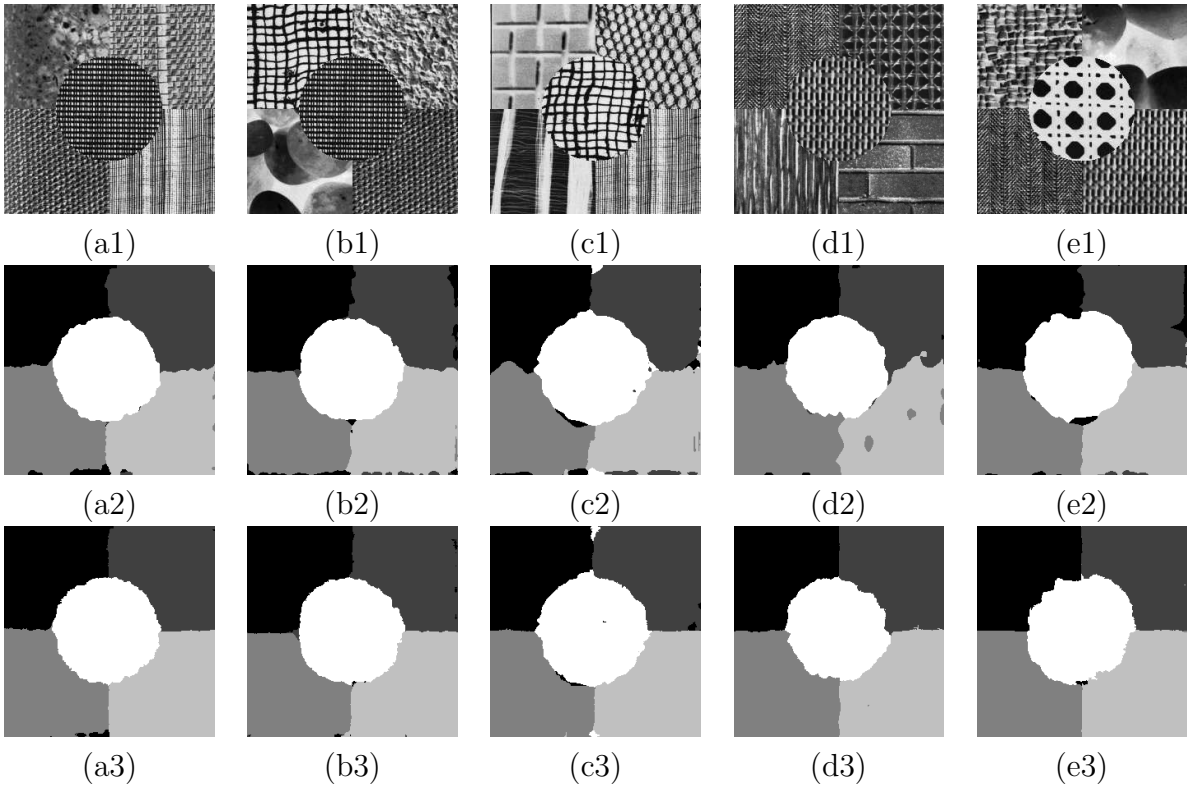


Figure 4.4. Segmentation of five-texture images. (a1)-(e1): the 256×256 synthetic images from Brodatz album; (a1) Im1 image composed of D73, D85, D77, D106 and D21; (b1) Im2: D104, D4, D30, D77 and D21; (c1) Im3: D1, D36, D51, D106 and D104; (d1) Im4: D16, D52, D68, D94 and D53; and (e1) Im5: D84, D30, D16, D53 and D101. (a2)-(e2): Segmentation results using CDFB. (a3)-(e3): Segmentation results using CDFB-LMD.

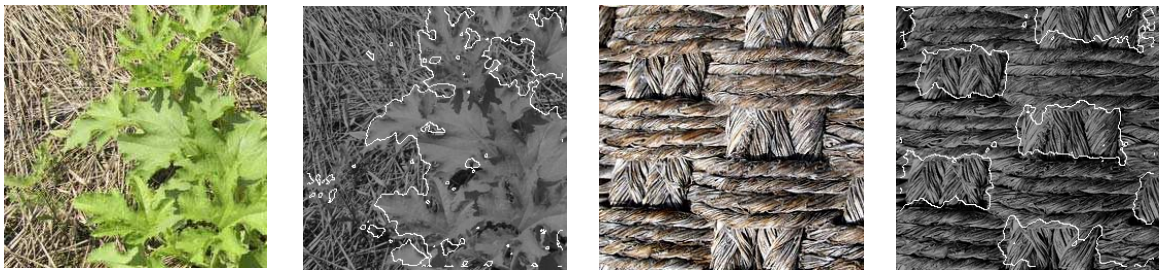


Figure 4.5. Segmentation of real images. (a) the pumpkinplant image, (c) the island texture, (b) and (d) segmentation results using CDFB-LMD.

Table 4.1. Comparison of different feature extraction schemes in segmentation of two-texture images

	Gabor/+LMD	**CWT/+LMD**	**CDFB / +LMD**
# scales	6 /6	6 /6	5 /5
# orientations	8 /8	6 /6	8 /8
# features	48 /96	36 /72	40 /80
D21-79	3.3 /2.1	2.8 /2.4	2.7 /1.3
D68-6	2.8 /1.3	3.3 /2.2	1.6 /1.4
D101-106	1.6 /1.4	2.1 /1.6	3.0 /1.6
D12-8	2.3 /2.2	2.7 /2.1	3.5 /2.3
D105-16	1.9 /1.2	1.4 /1.1	2.8 /1.7
D21-57	2.1 /1.9	2.1 /1.7	2.2 /1.4
D103-26	2.6 /2.4	2.0 /1.3	2.7 /2.1
D17-72	2.0 /1.8	3.7 /4.0	2.8 /2.2
D54-56	12.7 /5.6	12.9 /5.4	7.5 /4.8
D21-4	2.5 /2.2	2.9 /2.3	3.2 /1.7
D17-55	1.7 /2.2	3.9 /2.1	4.2 /2.2
D21-77	1.8 /1.2	1.6 /0.7	1.2 /1.0
D73-106	1.1 /1.2	0.8 /0.9	2.2 /2.1
D21-83	3.7 /3.1	3.9 /2.9	4.2 /2.2
D85-11	1.6 /1.6	4.7 /4.2	3.4 /2.5
Ave error	2.9%/2.1%	3.4% /2.3%	3.1% /2.0%
Improvement after using phase	**27.7%**	**31.3%**	**35.4%**

4.5 Summary

A new image feature is proposed for texture image segmentation. In addition to magnitude information typically used in many other segmentation methods, phase information is incorporated to further improve the performance. The phase information is extracted from the phase difference of neighboring subband coefficients of the complex transforms including the Gabor transform, the CWT and the CDFB. From the experiments, the seg-

Table 4.2. Comparison of different feature extraction schemes in segmentation of five-texture images

	Gabor	**CDFB**	**CDFB-LMD**
# scales	6	6	6
# orientations	8	8	8
# features	50	50	98
Ave errors	4.4%	4.7%	2.6%
Improvement			**46.0%**

mentation results after using phase are much better ($27.7\% - 46\%$ improvement) than using magnitude alone in term of classification error rate with different texture images. In this chapter and Chapter 2, phase information has been *explicitely* and successfully utilized for the applications of texture image segmentation and texture image retrieval. It would be interesting to see if phase information can be beneficial in other image processing applications.

CHAPTER 5
A STUDY OF RELATIVE PHASE IN COMPLEX WAVELET DOMAIN: PROPERTY, STATISTICS AND APPLICATION

5.1 Introduction

It is well known that an advantage of complex wavelets compared to conventional real-valued wavelets is that they provide both magnitude and phase information. The magnitude of a complex coefficient describes the strength of feature such as edges while the phase indicates the location of the feature. In many image processing applications, only the real part or the magnitude of complex wavelet coefficient and its statistical model are used, although the phase information can hold important statistical properties of the image. The property and statistical model of phase and its benefits have not been studied deeply. An essential question raised is how to efficiently utilize the phase information of complex wavelet coefficients for image modeling. In this chapter, we develop a new approach to exploit the statistical properties from the phase information of two-dimensional complex wavelet coefficients of an image [70, 71, 72].

There have been several works on the statistics of real coefficients of the wavelet transform [14, 16]. The wavelet coefficients within a subband which have the peaky and heavy-tailed non-Gaussian distribution can be modeled by a two-parameter generalized Gaussian density (GGD) [14, 73, 74]. The GGD is applied to image denoising as a prior density and outperforms the Gaussian model [75, 73, 74]. Although the GGD wavelet marginal model is more powerful than the Gaussian model, it does not take into account the dependencies between different subbands as well as between a coefficient and its neighboring coefficients in the same subband.

Various researchers have developed joint statistical models in the wavelet domain [18, 20, 22, 54]. A bivariate probability density function has been proposed to model the statistical dependencies between a wavelet coefficient and its parent [54]. The hidden

Markov model with a two-state hidden multiplier variable was introduced in [18, 20] to capture the key features of the joint probability density of the wavelet coefficients. A continuous multiplier variable has been proposed to governed the local variance [76, 77, 22, 46]. In [22, 17], the authors developed a model for the neighborhoods of oriented pyramid coefficients based on a Gaussian scale mixture (GSM) which is the product of a Gaussian random vector and an independent hidden random scalar multiplier. The GSM estimator provides significantly higher performances in terms of both mean square error and visual quality in image denoising [17]. In most of the above statistical models, only the real part of the coefficients is modeled.

In [78, 79, 80, 81], the magnitude distribution of the dual-tree complex wavelet transform (DTCWT) coefficients is modeled by a scaled mixture of Rayleigh distribution. However, in [82], the authors show that for a Gaussian distributed signal, the magnitude distribution of the DTCWT coefficients can be modeled by the Rayleigh distribution for only the second and higher level decompositions and by the generalized Gamma for the first-level decomposition. In [82], the density function of the phase components is not uniform for the first-level decomposition, and it is uniform for the second and higher level decompositions. Therefore, the moments of the phase distribution are computed. However, for natural images, distribution of phase components is uniform in all subbands of complex wavelet transforms such as the pyramidal dual-tree directional filter bank (PDTDFB) [2] and the modified curvelet transform [12].

In some applications a knowledge of the statistical properties of phase components is used, such as edge detection [83], image painting and reconstruction [84, 85]. Some other applications exploit the local phase information across scales of the complex wavelet such as the description of texture images [46], the detection of blurred images [47], object recognition [48] and the face recognition [50]. In [46] the local phase behavior is captured by the cross-correlation of complex coefficients of bands at adjacent scales (fine-scale and coarse-scale), and this statistical measurement distinguishes edges from lines and helps in representing gradients due to shading and lighting effects. In [47], the local phase

coherence relationship across scale and space has been suggested and the phases of the finest scale coefficients can be well predicted from those of the coarser scale coefficients. The disruption of this local phase is an effective measure for blur detection. This across scale relationships are also captured using modified product of coefficients at adjacent scales and have been used in [48, 86]. Another investigation of local phase in the same orientation and the same scale is based on the DTCWT and the feature orientations are determined by the phase difference between the adjacent coefficients in six fixed directional subbands [49].

The statistical model for phases has not been explicitly utilized in image processing applications. Therefore a statistical model for the phases of the complex wavelet coefficients can be beneficial to the developments in the image processing community. In this chapter, we develop a new studying approach of the phase difference of two neighboring complex wavelet coefficients called relative phase. The statistics of relative phase is studied. Two simple periodic distribution, the von Mises and wrapped Cauchy are proposed to capture the relative phase statistics. The parameters of both models are estimated by maximizing the likelihood function. We compare the performances of the two models using real images which include the standard natural and texture images.

5.2 Modified Curvelet Transform

The modified multiresolution and multidirectional discrete transform borrows the ideas from the two recently introduced discrete transforms, which are the pyramidal dual-tree directional filter bank (PDTDFB) [2] and the fast discrete curvelet transform [13]. We define in the frequency plane a set of N 2-D directional filters $\phi_i(\boldsymbol{z}), i = 1, .., N$ and a lowpass filter $\phi_0(\boldsymbol{z})$ in such a way that the directional subbands and the lowpass subband can be decimated without aliasing. The decimation ratio for the lowpass band is $diag\{2, 2\}$, and the decimation ratios for the first and second $N/2$ directional filters are $diag\{M/2, 2\}$ and $diag\{2, M/2\}$, respectively, where M is a power of two number and is linearly proportional to N. The defined filters in the frequency domain are real-

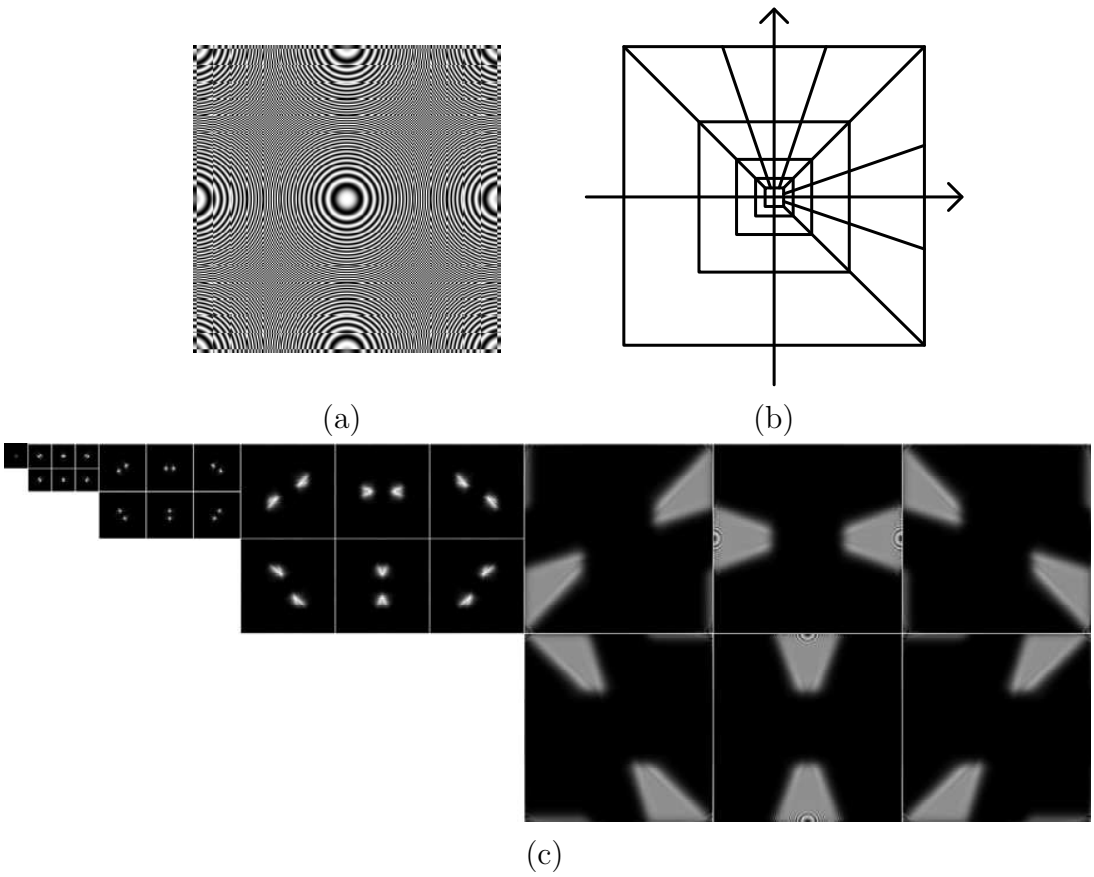


Figure 5.1. (a) The zoneplate image, (b) Frequency support of the curvelet functions, and (c) Magnitude of complex coefficients in the transform domain with 4 scales and 6 orientations.

valued functions and satisfy the perfect reconstruction conditions, taking into account the decimation ratio: $\frac{1}{4}\phi_0^2(\boldsymbol{\omega}) + \frac{1}{2M}\sum_{i=1}^{N}\phi_i^2(\boldsymbol{\omega}) + \phi_i^2(-\boldsymbol{\omega}) = 1$.

Similar to the PDTDFB, the directional filters have one-sided support in the frequency domain, making the subband coefficients complex. The same transform with different values of M and N is applied iteratively at the lowpass subband to create a multiresolution decomposition. In the reconstruction procedure, the final complex components are simply discarded. We can interpret this as a dual-tree FB structure [9]. The discrete transform also has some similarities to the wrapping-based fast discrete curvelet transform (FDCT) [13] in the sense that both are defined based on windowing in the DFT domain. The main difference is that in the wrapping-based FDCT, the redundancy of the transform is reduced by wrapping the frequency domain of the subbands, while in the

implementation, the redundancy is reduced by decimating the subbands by diagonal integer matrices. By this construction, the curvelet basis functions are located on a uniform integer grid at each resolution, while the basis of the FDCT is located on a non-integer grid. Fig. 5.1 is an example of the zoneplate image decomposed by the discrete curvelet transform. The decomposition has four directional scales, with $N = 6$ at each scale. The Fig. 5.1(c) shows the magnitude of the complex coefficients in the transform domain. For detailed construction of this new directional transform, the reader is referred to [12].

5.3 Relative Phase and Its Property In Complex Wavelet Domain

Definition 5.3.1. *The relative phase at a spatial location (k_1, k_2) within a particular complex subband is defined as the phase difference of two adjacent complex wavelet coefficients, e.g.,*

$$\begin{aligned} \theta(k_1, k_2) &= \angle y(k_1, k_2) - \angle y(k_1, k_2 + 1), \\ or\ \ \theta(k_1, k_2) &= \angle y(k_1, k_2) - \angle y(k_1 + 1, k_2), \end{aligned} \tag{5.1}$$

where $y(k_1, k_2)$ is the coefficient at position (k_1, k_2).

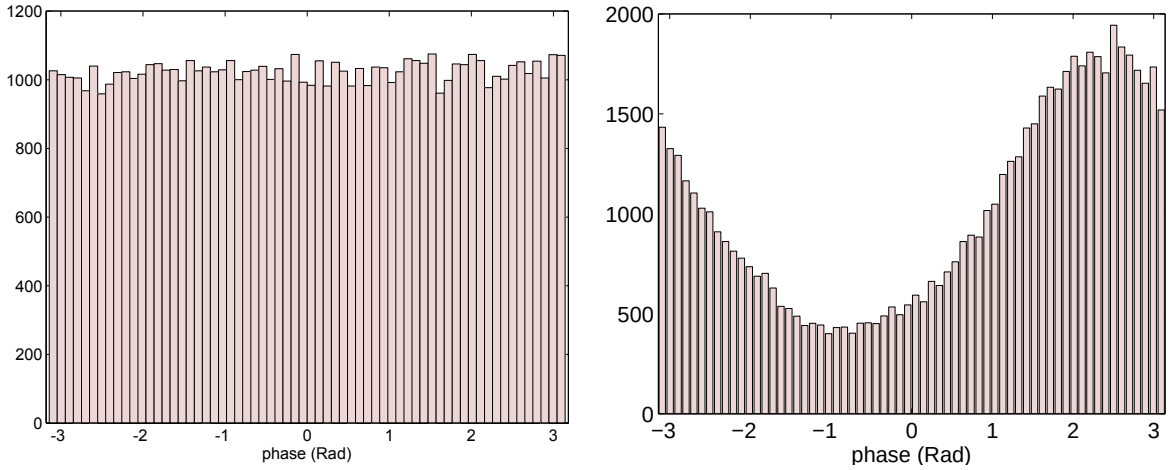
Property 5.3.1. *The feature orientation α of an edge in the supported region of a subband is linearly proportional to the relative phase of two adjacent complex wavelet coefficients as follows*

$$\alpha \approx a + b\ \theta(k_1, k_2), \tag{5.2}$$

where a and b are constant and can be estimated for each subband. It is noted that the CDFB is an example to prove this property of relative phase (see Section 5.3) which also agrees with other complex wavelet transforms such as the DTCWT and the complex curvelet transform.

5.4 Statistics of Relative Phases In Complex Wavelet Domain

In this section, we will observe behaviors of relative phase within a subband and propose its probability density functions.



(a) (b)

Figure 5.2. Histograms at a particular wavelet subband of Barbara image. (a) Phases of complex coefficients and (b) Relative phase.

5.4.1 Descriptive Statistics

Since phase is circular data, we will need some measures like circular mean, circular variance to describe the circular distribution. These will be useful in making comparison between circular distributions.

5.4.1.1 Measure of center and circular mean

Let θ_1, θ_2, ..., θ_n be a set of circular observations given in terms of angles. Consider the polar to rectangular transformation for each observation $(\cos\theta_i, \sin\theta_i)$, $i = 1, 2, ..., n$. We obtain the resultant vector of these n unit vectors by summing them to get

$$R = (C, S) = (\sum_{i=1}^{n} \cos\theta_i, \sum_{i=1}^{n} \sin\theta_i). \tag{5.3}$$

The direction of this resultant vector R is the circular mean direction, denoted by $\bar{\theta}$ and defined as follows

$$\bar{\theta} = \arg\left(\sum_{i=1}^{n} \cos\theta_i + j\sum_{i=1}^{n} \sin\theta_i\right). \tag{5.4}$$

So, the circular mean direction is given by

$$\bar{\theta} = \arctan\left(\frac{\sum_{i=1}^{n} \sin\theta_i}{\sum_{i=1}^{n} \cos\theta_i}\right), \tag{5.5}$$

where arctan is the four quadrant inverse tangent. Using function $atan2.m$ in Matlab, we can find the four quadrant inverse tangent of the real parts of the elements of C and S, where $-\pi \leq \bar{\theta} = \text{atan2(S,C)} \leq \pi$.

5.4.1.2 Measure of dispersion and circular variance

The direction $\bar{\theta}$ of the vector resultant R provides a mean direction. However, we will want to know how concentrated circular data is towards this center. If all angles point in the same direction, the concentration is high. Conversely, if the data spread over the circle, there will be no concentration. The dispersion D of the points corresponding to θ_i on the unit circle is defined as

$$D = \frac{1}{n}\sum_{i=1}^{n}(1 - \cos(\theta_i - \alpha)), \tag{5.6}$$

where α is an angle of an arbitrary point on the unit circle.

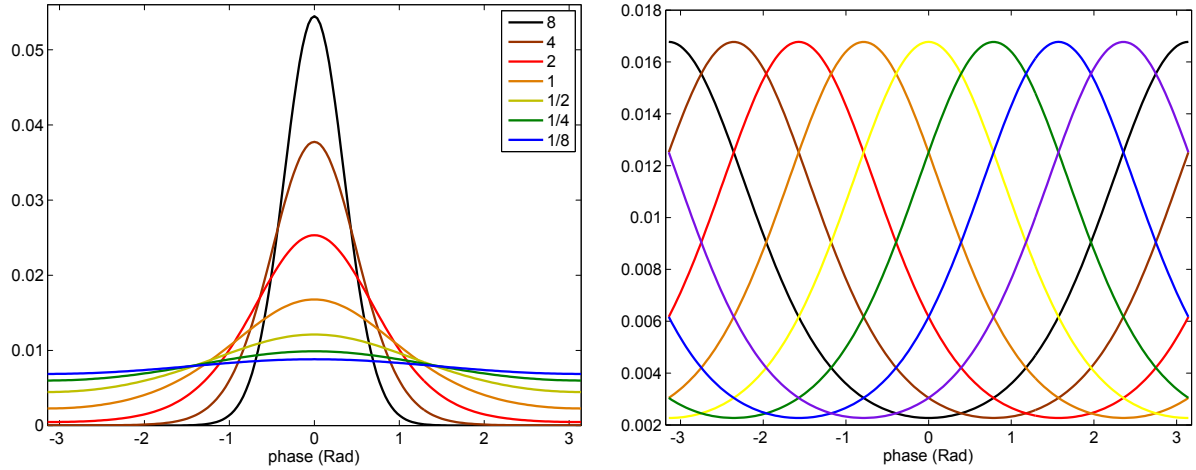
The dispersion D will be minimized at $\alpha = \bar{\theta}$. The circular variance is defined as

$$S = 1 - \bar{R} = 1 - \frac{1}{n}\sum_{i=1}^{n}(1 - \cos(\theta_i - \bar{\theta})), \tag{5.7}$$

where $\bar{R}$ is the mean resultant length, and $0 \leq S \leq 1$. If the n observed directions are tightly clustered about the mean direction θ_0 then circular variance S will be nearly zero. On the other hand, if the directions are widely dispersed then S will be nearly one.

5.4.2 Circular Probability Distributions of Relative Phase

Our objective in this subsection is to find a statistical model which is able to capture the relative phase behaviors in the complex wavelet domain. The distribution for phases of complex coefficients is uniform as shown in Fig. 5.2(a). This uniform distribution of phases cannot yield any information of the image. Fig. 5.2(b) displays the histogram of the relative phase in one subband showing a (periodic) bell shape similar to the Gaussian distribution. Hence we propose the von Mises for modeling the probability density function (pdf) of the relative phase in the complex wavelet domain.



(a) (b)

Figure 5.3. Von Mises Distributions. (a) $\nu = (1/8, 1/4, ...8)$ and $\mu = 0$, and (b) $\nu = 1$ and $\mu = (-\pi, -3\pi/4, ..., 3\pi/4)$.

5.4.2.1 Von Mises distribution

An angular random variable θ has the von Mises distribution with the parameters μ and ν $VM(\mu, \nu)$ [87, 88] if its probability density function (pdf) has the form

$$p(\theta; \mu, \nu) = \frac{1}{2\pi I_0(\nu)} e^{\nu \cos(\theta - \mu)}, \tag{5.8}$$

where I_0 denotes the modified Bessel function of the first kind and the zero-th order which can be defined by $I_0(\nu) = \frac{1}{\pi} \int_0^{\pi} e^{\nu \cos\theta} d\theta$. The von Mises density functions in (5.8) with various values of ν and μ are depicted in Fig. 5.3.

The parameter μ is the mean direction and the parameter ν is known as the concentration parameter. Note that $VM(\mu, \nu)$ and $VM(\mu + \pi, -\nu)$ have the same distribution. For this model, we set the values of ν to be non-negative, and the ranges of θ and μ are $[-\pi\ \pi]$. The VMs distribution is unimodal with two parameters μ and ν, and is symmetrical about mean direction $\theta = \mu$. The larger the value of concentration parameter ν, the denser the clustering around the mean direction μ. For $\nu = 0$, the von Mises distribution tends to the uniform distribution. As $\nu \to \infty$ it becomes concentrated at the point $\theta = \mu_0$. The parameters of the von Mises distribution are estimated using maximum-likelihood as shown in Appendix A.1.

Fig. 5.5 shows the empirical histogram of the relative phase in a particular complex wavelet subband for four different natural images, along with the best fitting of the von Mises distribution. Fitting was performed by maximum-likelihood estimator. The von Mises distribution fits very well with the histograms in the first and second plots which is similar to Gaussian except that it is periodic with a period of 2π. We can conclude that the von Mises fits very well with the relative phase pdfs which are in Gaussian shapes.

In Fig. 5.5(c) and (d), the marginal distributions of the relative phase in these subbands are in peaky and heavy tailed shapes. Evidently, the von Mises distribution is not accurate enough to model these peaky, heavy-tailed pdfs of relative phase. In the next subsection, the wrapped Cauchy distribution is proposed to deal with this type of data.

5.4.2.2 Wrapped Cauchy distribution

The wrapped Cauchy distribution is obtained by wrapping the Cauchy distribution on the real line with density $f(x)$ around the circle, where

$$f(x) = \frac{\sigma}{\pi(\sigma^2 + (x-\mu)^2)}, \qquad -\infty < x < \infty. \tag{5.9}$$

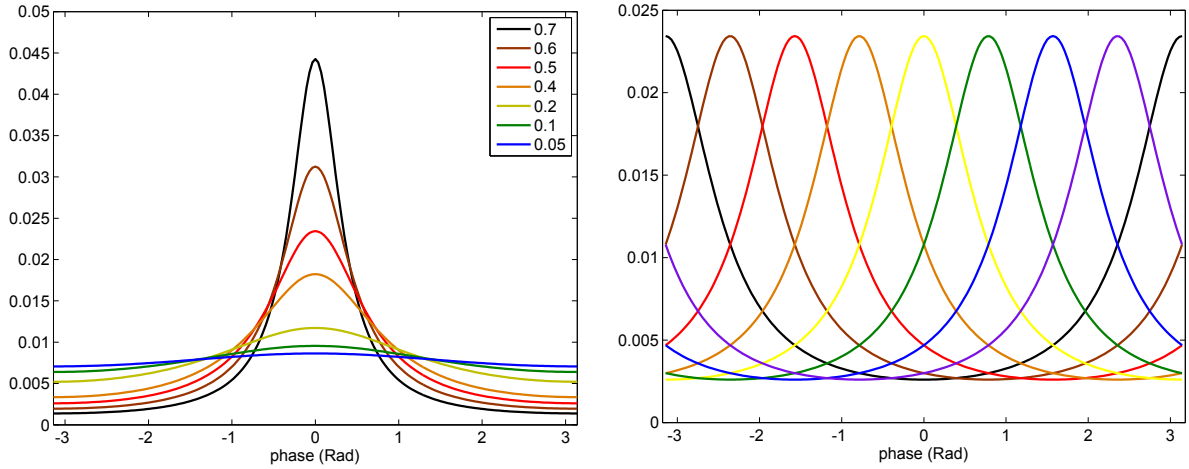
It has the probability density function [87, 88]

$$p(\theta) = \frac{1}{2\pi}\frac{1-\rho^2}{1+\rho^2-2\rho\cos(\theta-\mu)}, \qquad -\pi \le \theta \le \pi, \tag{5.10}$$

where $\rho = e^{-\sigma}$. The wrapped Cauchy density functions in (5.8) with various values of ρ and μ are shown in Fig. 5.4.

The wrapped Cauchy distribution is again unimodal and symmetric with the location parameter $-\pi \le \mu \le \pi$ and the scale parameter $0 \le \rho < 1$. As $\rho \to 0$, it tends to the uniform distribution and as $\rho \to 1$ it is concentrated at the point $\theta = \mu_0$. The maximum likelihood estimation for this model can be obtained by a recursive algorithm (see Appendix A.2).

Fig. 5.5 shows that the wrapped Cauchy distribution does not fits very well with the histograms in the first and the second plots as the von Mises (Fig. 5.5(a) and (b)) because



(a) (b)

Figure 5.4. Wrapped Cauchy Distributions. (a) $\rho = (0.05, 0.1, ...0.7)$ and $\mu = 0$, and (b) $\rho = 0.5$ and $\mu = (-\pi, -3\pi/4, ..., 3\pi/4)$.

the pdfs of the relative phase in these subband are in Gaussian shapes. However, it fits very well with the histograms in Fig. 5.5(c) and (d) because the pdfs of relative phase in these subbands are peaky and heavy tailed. Hence, we propose the wrapped Cauchy to model for relative phase pdfs which are in peaky and heavy-tailed shapes and Von Mises for relative phase pdfs which are in Gaussian shapes.

5.4.2.3 Comparison of results

This section details the comparison of the von Mises with the wrapped Cauchy distributions for relative phase in different complex wavelet transforms. We evaluate the von Mises and the wrapped Cauchy for many real images including texture images as well as standard images, which are decomposed by various complex wavelet transforms such as the DTDWT [8], the PDTDFB [2] and the modified curvelet [12]. The data for evaluation and their information such as the number of subbands, the size of subband and the number of bins are shown in Table 5.1.

The simulation results in Table 5.1 show that for most of the test images, the wrapped Cauchy model is more accurate than the von Mises. For several images, the accuracy of the fitted wrapped Cauchy model is slightly lower than or approximates that of the von

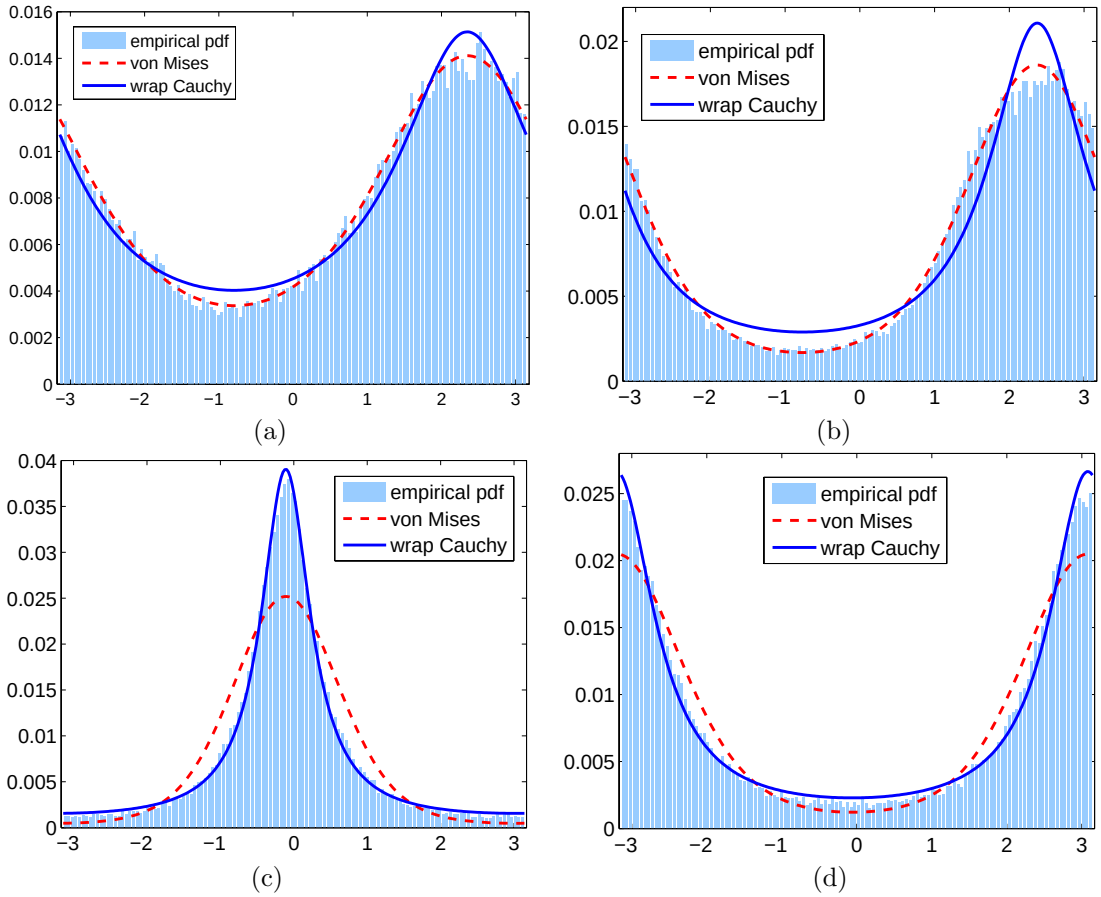
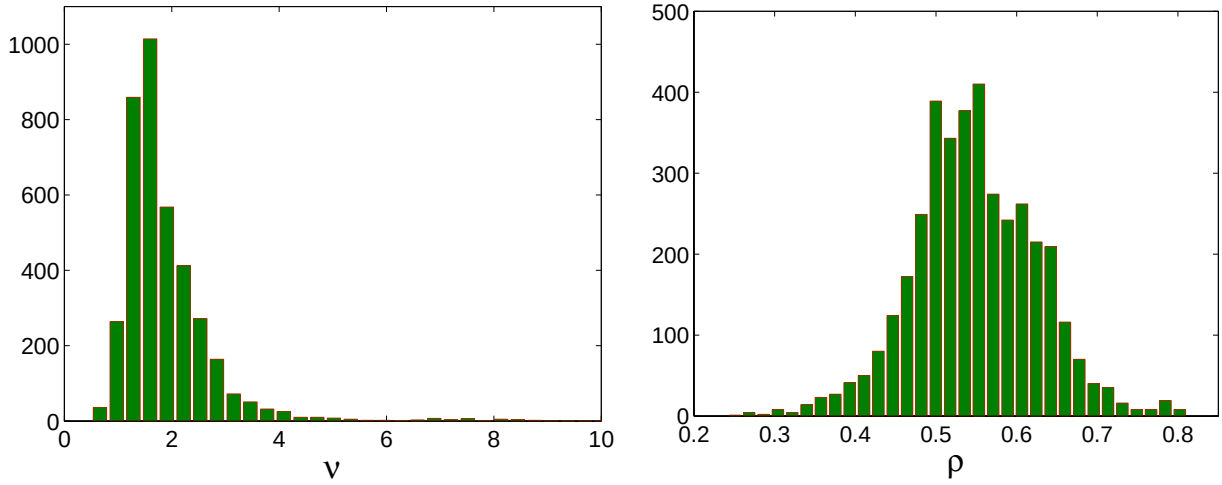


Figure 5.5. The von Mises and the wrapped Cauchy distribution fitted to the empirical histograms of relative phase at a particular finest complex wavelet subband. (a) Lena, $\Delta H/H = 0.0004$ (vMises) and $\Delta H/H = 0.0015$ (wCauchy); (b) Barbara,$\Delta H/H = 0.0008$ (vMises) and $\Delta H/H = 0.0069$ (wCauchy); (c) Boat, $\Delta H/H = 0.0178$ (vMises) and $\Delta H/H = 0.0016$ (wCauchy); and (d) Fingerprint, $\Delta H/H = 0.0050$ (vMises), and $\Delta H/H = 0.0016$ (wCauchy).

Mises. For all of the above complex wavelet transforms, if compared with the von Mises, the wrapped Cauchy is much better in fitting the relative phase pdf for all texture images and the fingerprint image. The last row in Table 5.1 presents the average $\Delta H/H$ of all test images. It is clear that compared with von Mises, the wrapped Cauchy is much better in fitting model for relative phase of complex wavelet coefficients.

5.4.2.4 Range of estimated parameters

It is of interest to know the common range for the values of ν in the von Mises model for texture images as well as of ρ in the wrapped Cauchy model. Fig. 5.6 shows the



(a) (b)

Figure 5.6. Histogram of estimated values for the parameters ν and ρ of 3840 finest subbands of size 64×64 from 640 texture images of size 128×128. (a) ν (von Mises) and (b) ρ (wrapped Cauchy).

histograms of the estimated parameters for 3840 finest subbands of size 64×64 from 640 texture images of size 128×128 described in Section 6.6.3. The curvelet transform is used to decompose texture images. For the von Mises model, most of values of the estimated ν parameter are from around 1 to 3, while the ρ values of the wrapped Cauchy model are from 0.45 to 0.65.

5.5 Application to Texture Image Retrieval

In this section, a new feature extraction method is proposed for texture image retrieval. The approach bases on incorporating the phase information obtained from a complex wavelet transform. A complex wavelet transform is used to decompose a texture image in order to provide complex subband coefficients. The relative phase statistics, extracted from the phases of the coefficients, is defined as additional features for classification. Performance of the proposed method is compared with the traditional methods. In order to illustrate the improvement of the relative phase-based feature in any complex wavelet transform, we use three multiresolution directional complex wavelet transforms including the DTCWT [8], the PDTDFB [2] and the curvelet transform [12]. These complex

Table 5.1. Average relative entropy of model and histogram, as a fraction of the total entropy of the histogram ($\Delta H/H$) for von Mises (VM) and wrapped Cauchy (WC) in various complex wavelet domains.

				DTDWT		PDTDFB		UCURV	
Texture	Bands	Size	Bins	VM	WC	VM	WC	VM	WC
640	3840	64×64	32	0.0042	**0.0027**	0.0139	**0.0065**	0.0103	**0.0045**
40	240	128×128	64	0.0028	**0.0015**	0.0110	**0.0042**	0.0078	**0.0031**
40	240	256×256	128	0.0020	**0.0010**	0.0086	**0.0029**	0.0060	**0.0024**
Lena	6	256×256	128	**0.0006**	0.0011	0.0024	**0.0022**	**0.0015**	**0.0015**
Barbara	6	256×256	128	**0.0013**	0.0017	**0.0026**	0.0042	**0.0023**	0.0047
Boat	6	256×256	128	**0.0020**	0.0025	0.0044	**0.0031**	0.0041	**0.0019**
Fingerp	6	256×256	128	0.0014	**0.0007**	0.0098	**0.0029**	0.0092	**0.0029**
Peppers	6	128×128	64	**0.0013**	0.0020	0.0045	**0.0028**	0.0029	**0.0023**
House	6	128×128	64	0.0024	**0.0018**	**0.0043**	0.0073	**0.0019**	0.0052
Cameram	6	128×128	64	**0.0019**	0.0022	0.0033	**0.0025**	0.0043	**0.0027**
Average	4362			0.0020	**0.0017**	0.0065	**0.0040**	0.0050	**0.0031**

transforms are similar to each other in terms of directionality of the filters and one-sided supports in the frequency domain. A comparison of various features such as energy feature [7] [39], GGD-based feature [89], WD-HMM based feature [34] and our proposed feature in texture retrieval is presented.

5.5.1 Texture Image Database and Feature Database

In our experiment, the texture database contains 40 texture images from the VisTex databases used in [34] [89] [39]. Each of these 512×512 images is divided into sixteen 128×128 non-overlapping sub-images, thus creating a database of 640 texture samples. Each original image is treated as a single class and therefore there are 16 samples from each of 40 classes. To reduce the intensity correlation, all images are normalized to have zero mean and unit variance. After that, three decompositions are applied for each image in the database, and their corresponding feature vectors are computed.

5.5.2 New Texture Image Feature Using Statistics of Relative Phase

In this image retrieval application, we propose using statistics of relative phase to extract feature for texture images. This feature extraction method is a statistical approach.

(a) (b)

Figure 5.7. Two sub-images with the size of 128×128: (a) Leaves.0011 and (b) Bark.0000.

When compared with the von Mises, the wrapped Cauchy is much better in fitting the relative phase pdf for all 640 sub-images from Vistex data in Table. 5.1. Therefore, two parameters of the wrapped Cauchy model for relative phase pdfs in each subband will be estimated.

Some examples of both images and their extracted features plus their distributions will be shown. Two sample sub-images, the Leaves.0011 and the Bark.0000, are from Vistex database as in Fig. 5.7. Their extracted features, which are the parameters μ and ρ of wrapped Cauchy distributions (Fig. 5.8), are shown in Table 5.2. With the same subband, two different images have two different relative phase pdfs with different parameters, e.g. at subband 4, the Leaves.0011 has the center at $\mu_4 = -2.08$, while the Bark.0000 has the center at $\mu_4 = -1.60$. The relative phase pdf of the Bark.0000 is more peaky with the peak of 0.15 than that of the Leaves.0011 with the peak of 0.10 at subband 4. It is clear that the distinction between two different types of textures can be captured and recognized by relative phase pdfs. The proposed feature provides statistical information of relative phase for each subband and can be used to discriminate two textures. Therefore, the wrapped Cauchy model of relative phase pdf will be used as an additional feature for texture image retrieval in this section.

Table 5.2. Some examples of the proposed features using statistics of relative phase (wrapped Cauchy parameters μ and ρ) shown in Fig. 5.8.

Image	μ_4	μ_5	μ_6	ρ_4	ρ_5	ρ_6
Leaves.0011	-2.08	-0.05	1.89	0.55	0.61	0.58
Bark.0000	-1.60	-0.02	1.57	0.65	0.69	0.66

5.5.3 Texture Feature Extraction

Each image in the database is applied to three following decompositions: the dual-tree complex wavelet transform, the PDTDFB and the modified curvelet transform. The dual-tree CWT and the curvelet transform are applied with four scales and six orientations per scale, while the other has three scales of eight orientations. The mean and standard deviation of the absolute values of the coefficients are calculated as in [7] [39]. These features are denoted as magnitude (MAG).

In the traditional energy-based approach, only magnitude of complex coefficients are used to creature feature vector. In this chapter, we propose a new approach to create a feature based magnitude as well as phase which we name MAG-RP. First, the relative phase matrix of each complex subband in the complex wavelet domain are computed as in (6.14), and circular mean and standard deviation of this relative phase matrix will be estimated by (5.5) and (5.7) to form the relative phase feature. After that, we combine the MAG feature and the relative phase feature to create the MAG-RP feature.

In the statistical approach, two parameters of the GGD model for real coefficients in each subband will be estimated by [89]. In our new approach, we will create a feature based real part model as well as imaginary part model through GGD and relative phase models von Misses and wrapped Cauchy which we name GGD-VM and GGD-WC. For each relative phase matrix, two parameters μ and ν of the von Mises model are estimated using maximum-likelihood as shown in Appendix A.1, and two parameters μ and ρ of the wrapped Cauchy model are also estimated by maximum likelihood estimator obtained by a recursive algorithm as in Appendix A.2. These parameters are used to form the von Mises and wrapped Cauchy feature vectors.

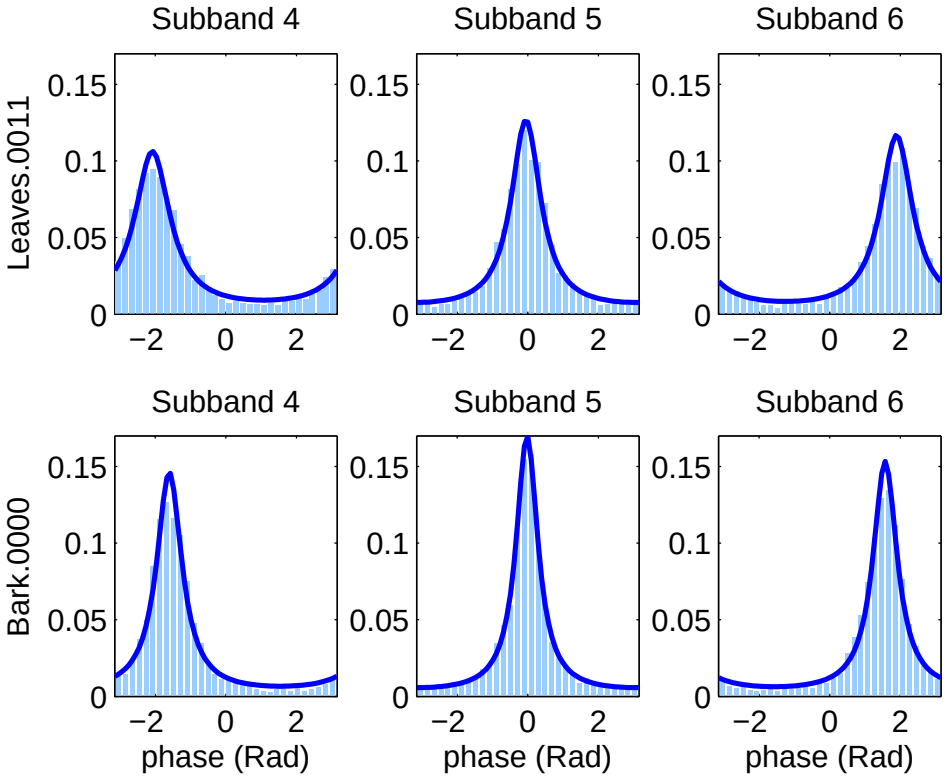


Figure 5.8. The wrapped Cauchy distributions fitted to three empirical histograms of relative phase at three subbands $4, 5$ and 6 with the size of 64×64 from two sub-images shown in Fig. 5.7.

In the first experiment with results shown in Tables 5.4, 5.5 and 5.6, we only use all subbands at the finest scale. The dual-tree CWT and the curvelet transform have six subbands. Each subband is extracted by two parameters. Therefore, the length of MAG, GGD feature is $6 = 12$, and the length of relative phase, von Mises and wrapped Cauchy feature is also 12. The combination of these feature will create the length double.

In the second experiment with results shown in Table 5.7, we use the modified curvelet transform and create the feature vector with length 48. This proposed feature contains 24 GGD parameters of two finest scales, 12 wrapped Cauchy parameters of the finest scale, and 12 MAG means of two coarsest scales. Since the size of subimage is 64, the relative phase models are only accurate enough for subbands at the finest scale with the size of 32. Therefore, the relative phase models of the finest scale should be used to form feature vector for retrieval application. Similarly, the GGD models of two finest

scales will be estimated. The feature of remaining subbands from two coarsest scales will be extracted by computing the mean of absolute values.

5.5.4 Distance Measure and Query Processing

5.5.4.1 Distance measure between two MAG features

The query pattern can be any one of the texture patterns from the image databases. Let f_x and f_y be two MAG feature vectors obtained from one of three transforms. The distance between them is given by [7]

$$d = \sum_k \left(\left| \frac{m_k(x) - m_k(y)}{\alpha(m_k)} \right| + \left| \frac{\sigma_k(x) - \sigma_k(y)}{\alpha(\sigma_k)} \right| \right),$$

where $\alpha(m_k)$ and $\alpha(\sigma_k)$ are the standard deviations of $m_k(\cdot)$ and $\sigma_k(\cdot)$ of the entire database.

5.5.4.2 Distance measure between two relative phase features

Let f_x and f_y be two relative phase feature vectors. We propose the distance between them as follows

$$d = \sum_k \left| \frac{S_k(x) - S_k(y)}{\alpha(S_k)} \right| + \frac{1 - \cos(\bar{\theta}_k(x) - \bar{\theta}_k(y))}{\alpha(\bar{\theta}_k)},$$

where $\alpha(S_k)$ is the standard deviations of $S_k(\cdot)$ and $\alpha(\bar{\theta}_k)$ is the circular variance of $\bar{\theta}_k(\cdot)$ of the entire database.

Since circular mean $\bar{\theta}$ is circular data, we will need a discrimination rule based on a circular distance. For any two points on the unit circle (θ_i, θ_j), the circular distance is defined by $d_{ij} = 1 - \cos(\theta_i - \theta_j)$. It is non-negative, symmetric in its indices and is invariant under rotation. To normalize the circular mean features over the whole database, in stead of using d_{ij} divided by the standard deviation as in [7], we replace the standard deviation by the circular variance of $\bar{\theta}_k(\cdot)$ of the entire database. Another feature of relative phase is circular variance S which is not circular data, we will compute the L_1 distance divided by the standard deviation of $S_k(\cdot)$ of the entire database.

5.5.4.3 Similarity measurement between two models

Given the von Mises model, the relative phase pdf in each subband can be defined via two parameters μ and ν. The Kullback-Leibler divergence (KLD) or the relative entropy between two pdfs $p(\theta;\mu_1,\nu_1)$ and $p(\theta;\mu_2,\nu_2)$ is

$$D_{KL}(P||Q) = \int_{-\pi}^{\pi} p(\theta;\mu_1,\nu_1)\log\frac{p(\theta;\mu_1,\nu_1)}{p(\theta;\mu_2,\nu_2)}d\theta. \tag{5.11}$$

Substitute (5.8) into (6.24) and after some manipulations we obtained a closed form for the KLD between the two von Mises pdfs as

$$D_{KL} = \log\frac{I_0(\nu_2)}{I_0(\nu_1)} + \frac{I_1(\nu_1)-I_1(-\nu_1)}{2I_0(\nu_1)}(\nu_1-\nu_2\cos(\mu_2-\mu_1)), \tag{5.12}$$

where I_1 denotes the modified Bessel function of the first kind and order one which can be defined by $I_1(\nu)=\frac{1}{\pi}\int_0^{\pi} e^{\nu\cos\theta}\cos\theta d\theta$. Therefore, the similarity measurement between two complex wavelet subbands can be calculated very efficiently using the model parameters. The overall similarity distance between two images is the sum of the KLDs given in (5.12) between corresponding pairs of all wavelet subbands as follows

$$d(f_x,f_y) = \sum_k D_{KL}(p(.;\mu_k(x),\nu_k(x))||p(.;\mu_k(y),\nu_k(y))), \tag{5.13}$$

In stead of using the von Mises model, we can use the wrapped Cauchy for modeling the RPs in each subband via two parameters μ and ρ. The Kullback-Leibler divergence (KLD) between two pdfs $p(\theta;\mu_1,\rho_1)$ and $p(\theta;\mu_2,\rho_2)$ is also given by (6.24), where von Mises density is replaced by wrapped Cauchy density and the overall similarity distance between two images is similar to (6.24). However, a closed form for the KLD between the two wrapped Cauchy pdfs is under investigation, and a numerical method with 128 bins is applied to estimate this KLD.

5.5.4.4 Query processing

For each query image, N nearest neighbors are selected, and the number of these textures belonging to the same class as the query texture, except for itself, is counted.

Table 5.3. Feature vector length, feature extraction time and similarity measurement time of query image.

	MAG	MAG-RP	GGD	GGD-VM	GGD-WC	GGD-WC
Feature type	m,σ	m,σ,$\bar{\theta}$,S	α, β	α, β, μ, ν	α, β, μ, ρ	$\alpha, \beta, \mu, \nu, m$
Feature length	12	24	12	24	24	48
Feature time	0.076 s	0.118 s	0.262 s	0.291 s	0.628 s	0.667 s
Similarity time	0.058 ms	0.562 ms	0.198 ms	0.769 ms	2.245 ms	2.314 ms

Table 5.4. Average retrieval accuracy of 40 texture images in the Vistex database using dual-tree complex wavelet with features extracted from the finest scale.

N	MAG	MAG-RP	Improve	GGD	GGD-VM	Improve	GGD-WC	Improve
1	89.38	96.25	7.69	92.34	96.41	4.40	96.56	4.57
3	84.69	93.07	9.90	88.96	93.28	4.86	94.53	6.26
5	80.66	89.47	10.93	85.19	89.75	5.36	90.91	6.71
7	77.30	86.36	11.72	81.58	86.96	6.59	87.99	7.85
10	71.92	81.61	13.47	76.89	82.30	7.03	83.69	8.84
15	63.12	72.71	15.18	67.72	72.16	6.55	73.60	8.69
20	69.04	77.96	12.91	73.52	77.64	5.60	78.97	7.41
30	75.70	83.68	10.54	80.68	83.45	3.43	84.30	4.49
40	80.35	86.82	8.05	84.78	87.06	2.69	88.04	3.85
50	83.83	89.11	6.30	88.20	89.92	1.95	90.68	2.81
60	86.39	90.74	5.04	90.67	92.15	1.63	92.54	2.07
			5.04% **→15.18%**			**1.63%** **→7.03%**		**2.01%** **→%8.88%**

This number (less than or equal to fifteen) divided by fifteen if $N > 15$ and divided by N if $N < 15$ is defined as the retrieval rate. The performance of the entire class is obtained by averaging this rate over sixteen members which belong to the same class of texture. The average of all classes is the overall performance of the transform.

5.5.5 Experimental Results

In the first experiment, the average retrieval rates of new approach based on both magnitude and phase can significantly improve up $8.69\% - 23.71\%$ when the top 15 texture images that are nearest to the search texture are considered ($N = 15$) shown in Tables 5.4, 5.5 and 5.6. These Tables summarize the overall retrieval rates when we use only all subbands from the finest scale of three different complex directional wavelet transforms. The feature vectors are detailed in Table 5.3. The MAG-RP is consistently better

Table 5.5. Average retrieval accuracy of 40 texture images in the Vistex database using complex directional filter bank with features extracted from the finest scale.

N	MAG	MAG-RP	Improve	GGD	GGD-VM	Improve	GGD-WC	Improve
1	89.06	94.69	6.32	90.31	96.41	6.75	95.47	5.71
3	85.00	92.66	9.01	87.50	94.11	7.56	94.64	8.15
5	80.88	90.50	11.90	83.38	91.94	10.27	91.81	10.12
7	76.61	87.68	14.45	79.93	89.49	11.95	89.84	12.40
10	71.39	83.84	17.44	74.33	85.53	15.07	85.58	15.14
15	62.80	75.74	20.60	65.47	76.66	17.09	77.35	18.15
20	68.21	81.16	18.98	70.94	81.80	15.32	82.39	16.14
30	74.84	86.34	15.37	77.57	87.17	12.37	87.06	12.23
40	79.22	89.33	12.77	82.00	90.49	10.35	89.97	9.72
50	82.44	91.48	10.97	85.38	92.64	8.50	92.41	8.24
60	84.95	92.95	9.42	87.64	94.10	7.38	93.94	7.19
			6.32% **→20.60%**			**6.75%** **→17.09%**		**5.71%** **→18.15%**

than the MAG, and the GGD-WM and GGD-WC are consistently better than the GGD as shown in Tables 5.4, 5.5 and 5.6. We observe that when using phase the retrieval rates are significantly higher than without using phase. The improvement occurs at any N from 1 to 60 and this agrees with all of three complex wavelet transforms. The dual-tree CWT yields the improvement rate of $2.01\% - 15.18$ %, while the PDTDFB gains $5.71\% - 20.60\%$ and the modified curvelet gains $5.56\% - 23.71\%$. It is clear that this improvement is caused by the presence of the relative phase in the new feature vector. The relative phase information in the proposed approach can definitely support the existing methods including energy-based method as well as the GGD-based method.

When number of texture images selected N increases with ($N < 15$), the performance decreases but the improvement rate will increase. On the contrary for the cases of $N > 15$, the performance increases but the improvement rate will decrease. Therefore the improvement rate will be maximum at $N = 15$ in most cases. In the second experiment with the same database, 40 VisTex textures, the performance of our proposed method is compared with the performances of different existing methods including energy based feature [39], GGD-based feature [89], WD-HMM based feature [34]. Simulation results in Table 5.7 show that our proposed method outperforms the others. The feature vectors

Table 5.6. Average retrieval accuracy of 40 texture images in the Vistex database using modified curvelet transform with features extracted from the finest scale.

N	MAG	MAG-RP	Improve	GGD	GGD-VM	Improve	GGD-WC	Improve
1	86.72	95.63	10.27	87.81	96.72	10.14	96.72	10.14
3	82.60	94.32	14.19	85.73	95.31	11.18	94.95	10.75
5	78.78	91.56	16.22	81.91	93.13	13.70	92.91	13.43
7	75.07	89.33	19.00	78.24	90.78	16.03	90.29	15.41
10	70.27	85.31	21.41	73.45	86.89	18.29	86.38	17.59
15	62.20	76.95	23.71	64.60	78.86	22.07	78.41	21.36
20	67.98	81.95	20.55	70.86	83.40	17.68	82.91	16.99
30	74.55	86.00	15.36	78.14	87.66	12.19	87.39	11.84
40	79.02	88.52	12.02	82.97	90.39	8.94	90.02	8.50
50	82.64	90.45	9.45	85.89	92.13	7.27	91.68	6.74
60	85.35	91.82	7.58	88.10	93.40	6.01	93.00	5.56
			7.58% **→23.71%**			**6.01%** **→22.07%**		**5.56%** **→21.36%**

of all previous methods are extracted from the magnitudes or the real values of wavelet subbands, while our method exploits the phase of complex coefficients and extracts image information in both magnitude and phase. The incorporated phase information which is absolutely complementary to magnitude is the reason why the new feature can achieve better performance than the others.

5.5.6 Computational Complexity

The proposed texture retrieval method is implemented in MATLAB R2006a. With an Intel core 2 CPU 2.13 GHz machine, the feature vector time and the similarity measurement time of different features are shown in Table 5.3.

The time to extract the proposed feature GGD-WC with the length of 48 from one image is about 0.667 second (s). It includes 0.039 s for decomposing image using the modified curvelet transform (4 scales and 6 subbands each scale), 0.259 s for estimating 24 GGD parameters of two finest scales, 0.003 s for computing 12 MAG features of two coarsest scales and 0.366 s for estimating 12 wrapped Cauchy parameters of the finest scales. It take more time to extract the GGD-WC feature vector than the other features as shown in Table 5.3. This can be a weak aspect of the proposed method. However, the

time to extract a feature vector from one image is still much less than 1 s. Therefore, this proposed feature can be suitable for real-time applications.

With the same feature vector length, the similarity measurement time between two images using the GGD-VM which is 0.769 ms is much less than 2.245 ms when using the GGD-WC. The reason is that we derived the closed form of distance between the two von Mises pdfs as in (5.12) while a closed form for the KLD between the two wrapped Cauchy pdfs is developing and we have to use a numerical method with 128 bins to estimate this KLD. Therefore, instead of using the wrapped Cauchy, we can use von Mises for a fast feature extraction and searching time. Although the wrapped Cauchy is better than von Mises in fitting model for relative phase of complex wavelet coefficients, the performance of the von Mises feature is still comparable with the performance of the wrapped Cauchy feature in texture image retrieval application as shown in Tables 5.4, 5.5 and 5.6.

In this experiment, with the database of 640 images, it takes about 1.5 s to retrieve 15 images which are similar to the query image. This searching time is still acceptable. However if the database is huge, the numerical method for estimating KLD between the two wrapped Cauchy pdfs is a disadvantage of the searching time of the proposed method. The fast estimation algorithm for KLD between the two wrapped Cauchy pdfs is the future work.

5.6 Conclusion

A new approach which exploits the probabilistic properties from the phase information of two-dimensional complex wavelet coefficients for the image modeling is studied. The property and the statistics of relative phase for the image modeling are investigated. We conclude that the von Mises fits very well only with the relative phase pdfs which are in Gaussian shapes. The wrapped Cauchy is proposed for the relative phase pdf which is in peaky and heavy-tailed shape. The parameters of the von Mises and wrapped Cauchy pdf are estimated by maximum-likelihood estimators. We demonstrate that the wrapped Cauchy captures the peaky and heavy-tailed behaviors of relative phases. The simula-

Table 5.7. Average retrieval rates over the whole database for various methods when $N = 15$

Methods	Feature length	Rate
Standard DWT: L1 + L2	18	64.83 %
GGD [89]	18	75.73 %
scalar WD-HMM [34]	33	76.51 %
vector WD-HMM [34]	41	80.05 %
DT-CWT + DT-RCWT [39]	80	81.16 %
CDFB-RP (Chapter 2)	48	82.86 %
proposed method: GGD-WC	48	**85.64 %**

tion results show that for most of the test images, the wrapped Cauchy model is more accurate than the von Mises, when images are decomposed by different complex wavelet transforms including the dual-tree complex wavelet (DTCWT), the pyramidal dual-tree directional filter bank (PDTDFB) and the modified curvelet. The statistics of relative phases are then applied to texture image retrieval application. The new approach based feature in which phase information is incorporated yields a higher retrieval accuracy. The new feature exploits the phase information which is rarely used before. The incorporated phase information which is absolutely complementary to magnitude is the reason why the new feature can achieve better performance than the others. The proposed relative phase approach would be applied to any complex wavelet transform and would be a promising approach for other applications in image processing.

CHAPTER 6
PROBABILITY DISTRIBUTION OF THE RELATIVE PHASE FOR IMAGE MODELING IN COMPLEX WAVELET DOMAIN

6.1 Introduction

Many applications in image processing such as image compression, restoration, synthesis, segmentation and retrieval can benefit from a statistical model to characterize the image in the transform domain. A clean, precise probability model which can sufficiently describe typical images becomes essential. In this chapter, a new model for relative phase of the complex directional wavelet coefficients is proposed for image modeling, and its application in the texture image retrieval is presented [90, 91].

There have been several works on the statistics of decomposition coefficients of the wavelet transform [14][16]. The wavelet coefficients within a subband are often assumed to be independent and identically distributed. With this assumption, the wavelet coefficients can be modeled by the marginal model whose distribution is a two-parameter generalized Gaussian density (GGD) [14][73][74]. The GGD is a suitable distribution for the peaky and heavy-tailed non-Gaussian statistic of typical image wavelet decomposition. It is applied to image denoising as a prior density and outperforms the Gaussian model [75][73][74]. Although the GGD wavelet marginal model is more powerful than the Gaussian model, it does not take into account the dependencies between different subbands as well as between a coefficient and its neighboring coefficients in the same subband.

A number of researchers have successfully developed joint statistical models in the wavelet domain [18][20][22][54]. A bivariate probability density function has been proposed to model the statistical dependencies between a wavelet coefficient and its parent [54]. The wavelet coefficients within each local neighboring are characterized by Gaussian scale mixture model (GSM), which can captured the behavior of the marginal distribution of wavelet coefficients and the correlation in their local amplitudes [22]. The hidden Markov model

with a two-state hidden multiplier variable was introduced in [18][20] to capture the key features of the joint probability density of the wavelet coefficients. A continuous multiplier variable has been proposed to governed the local variance [76][77][22][46]. In [22][17], the authors developed a model for the neighborhoods of oriented pyramid coefficients based on a Gaussian scale mixture (GSM) which is the product of a Gaussian random vector and an independent hidden random scalar multiplier. This model can account for both marginal and pairwise joint distributions of wavelet coefficients. The GSM estimator is more complex than the Gaussian or the wavelet marginal estimators but it provides significantly higher performances in terms of both mean square error and visual quality in image denoising [17].

In most of the above statistical models, only the real part or the magnitude of the coefficients is modeled and used for image processing applications. One of the earliest works that point out the importance of phase information is [40], with the famous example where the main image structure is reconstructed by using only the phase of the Fourier coefficients. The phase holds crucial information about image structures and features. The higher-order Fourier statistics have been then applied to examine the phase structure in natural images [92]. The image features such as edges and shadows are determined by analyzing the phase of the harmonic components [41] or computing the phase congruency [41][42][93]. At the points of isolated even and odd symmetric features such as lines and step edges, the arrival phases of all Fourier harmonics are identical [41][94]. The phase congruency matrix provides a quantitative measure of the significance of the edge at each pixel of the image, and yields high quality in edge detection [42]. Some previous works have used the Gabor phase in image classification applications such as the iris and palmprint identification [43][44] and the face recognition [45]. These methods are based on the quadrant bit coding (QBC) extracted from the complex Gabor coefficients. Each pixel in each subband image will be encoded to two bits according to the quadrant in which the Gabor phase angle lies. The feature vector is created from these coding bits over all subband images and the classification is achieved by the Hamming distance.

Some other applications exploit the local phase information across scales of the complex wavelet such as the description of texture images [46], the detection of blurred images [47], object recognition [48] and the face recognition [50]. In [46] the local phase behavior is captured by the cross-correlation of complex coefficients of bands at adjacent scales (fine-scale and coarse-scale), and this statistical measurement distinguishes edges from lines and helps in representing gradients due to shading and lighting effects. In [47], the local phase coherence relationship across scale and space has been suggested and the phases of the finest scale coefficients can be well predicted from those of the coarser scale coefficients. The disruption of this local phase is an effective measure for blur detection [47]. This across scale relationships are also captured using the modified product of coefficients at adjacent scales and has been used in [48][86]. Another investigation of local phase in the same orientation and the same scale is based on the dual-tree complex wavelet transform [49] and the complex directional filter bank (CDFB) [24]. The feature orientations are determined by the phase difference between the adjacent coefficients in six fixed directional subbands [49] and in 2^n directions [2].

In this chapter we propose a statistical model for the phase difference of two neighboring complex wavelet coefficients called relative phase (RP) in the complex wavelet domain. Parameters of the probability density function for the relative phase (RP-PDF) are estimated by maximizing the likelihood of the data under the model. In simulations, we show that the RP-PDF fits well with behaviors of the relative phase from various real images. The RP-PDF model is compared with other circular distributions including von Mises and wrapped Cauchy. We also test the RP-PDF model for different complex wavelet transforms including dual-tree complex wavelets (DTCWT), complex direction filter banks (CDFB) and our modified curvelet. The experimental results show that the RP-PDF model for relative phase is more accurate than the others. The RP-PDF model is then applied to obtain a new image feature for texture image retrieval application. In simulations, we show that the proposed model further improves the classification rate.

The remainder of this chapter is organized as follows. The proposed relative phase model is derived in Section 6.2 with the assumption of the Gaussian model for real coefficients. This RP-PDF model is proved to be also true with the GSM model in Section 6.3. In Section 6.4, we present an algorithm maximizing the likelihood of the data under the model to estimate the parameters of the RP-PDF. Section 6.5 details the comparison of the RP-PDF with other circular distributions in different complex wavelet transforms. Finally, Section 6.6 presents the texture image retrieval application and experimental results.

6.2 Complex Gaussian Model and Phase Distribution

Our objective is to find a statistical model which is able to accurately capture the phase information in the complex wavelet domain, and is also able to sufficiently describe the natural images. In this section, with the assumption that the distribution of the real and imaginary coefficients are Gaussian, we study the probability distribution of complex coefficients and the joint distribution of two neighboring phases. From the study results, the definition of the relative phase will be given. We then derive the probability density function for the relative phase within a particular subband.

6.2.1 Complex Gaussian Distribution

By construction of the complex wavelet, each pair of corresponding filters have the Hilbert transform relationship [11, 9]. Therefore, the equivalent directional complex filter for each subband has a one-sided frequency support as illustrated in Fig. 5.1(b), and the outputs of the complex filters are the complex coefficients. We assume that the real and imaginary coefficients in each subband are normally distributed. This assumption will be relaxed to cover a much broader class of distributions in the next section.

Let $\mathbf{z} = \left(z_1, z_2\right)^T$ be a complex random vector, where z_1 represents the reference coefficient, z_2 represents the neighboring coefficient and $\mathbf{z} = \mathbf{x} + j\mathbf{y}$. Hence, $\mathbf{x} = \left(x_1, x_2\right)^T$ and $\mathbf{y} = \left(y_1, y_2\right)^T$ are two real random vectors normally distributed with a joint density functions $p(\mathbf{x}, \mathbf{y})$. Since the complex directional filters are bandpass, it is clear that $E[z_n] =$

$E[x_n] + jE[y_n] = \mathbf{0}$ and $E[\mathbf{z}\mathbf{z}^T] \approx \mathbf{0}$, where $E[.]$ is the expectation operator and $n = 1, 2$. Furthermore, $\mathbf{C_z} = E[\mathbf{z}\mathbf{z}^H]$ is defined as the complex covariance matrix. In these expressions, the superscript T denotes transposition, the superscript H denotes complex conjugate transposition, and $j = \sqrt{-1}$.

By definition, $\mathbf{C_z}$ is positive definite and Hermitian symmetric, hence, its inverse exists. Then $p(\mathbf{x}, \mathbf{y})$ may be written as a function of $\mathbf{z}$ such as $p(\mathbf{x}, \mathbf{y}) \equiv p(\mathbf{z})$, where $p(\mathbf{z})$ is a real-valued function of the complex vector $\mathbf{z}$. The density function $p(\mathbf{z})$ is commonly referred to as the complex Gaussian density function and can be written as [55][56]

$$p(\mathbf{z}) = \frac{\exp\left(-\mathbf{z}^H \mathbf{C_z^{-1}} \mathbf{z}\right)}{\pi^2 \det(\mathbf{C_z})}. \tag{6.1}$$

6.2.2 Marginal Distribution of Phases

In stead of decomposing $\mathbf{z}$ into Cartesian coordinates, we can write it in terms of polar coordinates. For the one-D case, let $\mathbf{z} = z_1 = r_1 e^{j\theta_1}$. In each subband, the complex wavelet coefficients have zero mean, i.e. $E[z_1] = 0$, and the covariance $\mathbf{C_z} = \psi_{\mathbf{11}}$. Then we can have the joint distribution of $\mathbf{r}$ and $\mathbf{\Theta}$, $p(\mathbf{r}, \mathbf{\Theta}) = r_1 p(\mathbf{z})$ as

$$p(r_1, \theta_1) = \frac{r_1}{\pi \psi_{11}} \exp(-\frac{r_1^2}{\psi_{11}}). \tag{6.2}$$

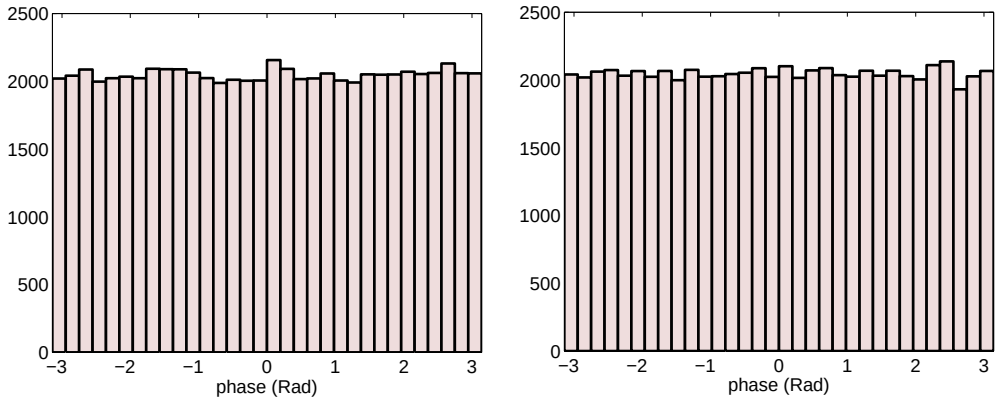
Hence the marginal distribution of the phases can be given by

$$\begin{aligned} p(\theta_1) &= \int_0^\infty \tfrac{r_1}{\pi\psi_{11}} \exp(-\tfrac{r_1^2}{\psi_{11}}) dr_1 \\ &= \tfrac{1}{2\pi}. \end{aligned} \tag{6.3}$$

The uniform distribution of phases as shown in Fig. 6.1 cannot inform us any information of the images. Therefore, the marginal distribution of phases should not be efficient for image processing applications.

6.2.3 Joint Distribution of Two Neighboring Phases

From the complex Gaussian model (6.1), we investigate the behaviors of the joint distribution of two neighboring phases in the complex wavelet domain. Let $\mathbf{z} = \left(z_1, z_2\right)^T =$



(a) (b)

Figure 6.1. Histogram of phases of complex coefficients at a particular wavelet subband. (a) Bark.0000, (b) Lena.

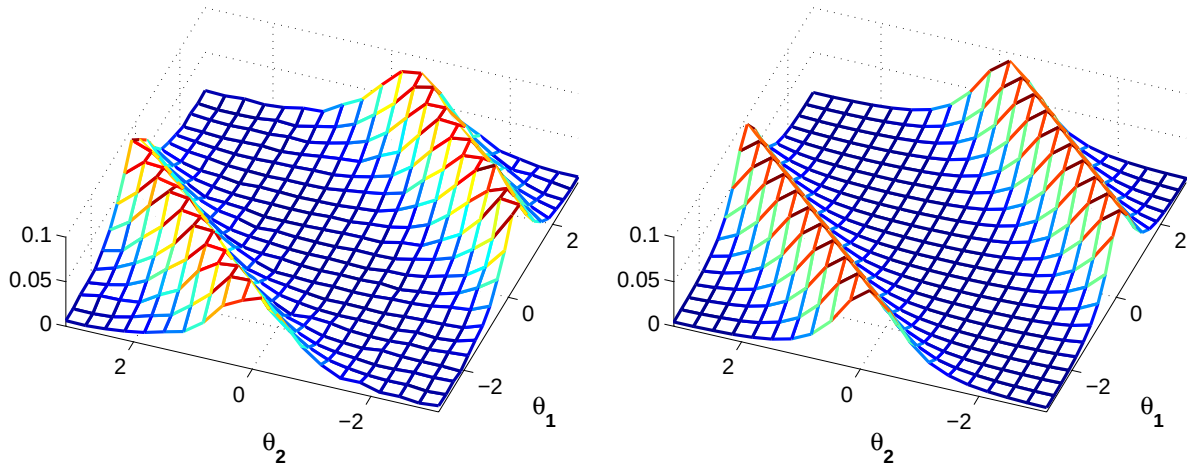
$\left(r_1 e^{j\theta_1}, r_2 e^{j\theta_2}\right)^T = \mathbf{r}e^{j\mathbf{\Theta}}$, where z_1 represents the reference coefficient, z_2 represents the neighboring coefficient and ($\mathbf{r}$, $\mathbf{\Theta}$) are the polar coordinates of $\mathbf{z}$. Hence, we can have the joint distribution $p(\mathbf{r}, \mathbf{\Theta}) = r_1 r_2 p_{\mathbf{z}}(\mathbf{z})$ [56] as

$$p(r_1, r_2, \theta_1, \theta_2) = r_1 r_2 \frac{\exp\left(-\mathbf{z}^H \mathbf{C}_{\mathbf{z}}^{-1} \mathbf{z}\right)}{\pi^2 \det(\mathbf{C}_{\mathbf{z}})}. \tag{6.4}$$

In each subband, the complex wavelet coefficients have zero mean and covariance $\mathbf{C_z} = E[\mathbf{z}\mathbf{z}^H] = \begin{bmatrix} \psi_{11} & \psi_{12} \\ \psi_{12}^* & \psi_{22} \end{bmatrix} = \mathbf{\Phi^{-1}}$. The matrix $\mathbf{C_z}$ is a positive definite Hermitian matrix, so its inverse $\mathbf{\Phi}$ exists. Let $\mathbf{\Phi} = \begin{bmatrix} \varphi_{11} & \varphi_{12} \\ \varphi_{12}^* & \varphi_{22} \end{bmatrix}$, where $\varphi_{12} = |\varphi_{12}| e^{j\mu}$, and the superscript $*$ denotes the complex conjugate. Then $\mathbf{z}^H \mathbf{C}_{\mathbf{z}}^{-1} \mathbf{z} = r_1^2 \varphi_{11} + r_2^2 \varphi_{22} + 2\mathbf{Re}[r_1 r_2 |\varphi_{12}| e^{j(\theta_1 - \theta_2 - \mu)}]$. Hence the joint distribution of two neighboring phases can be written as

$$\begin{aligned} p(\theta_1, \theta_2) &= \pi^{-2} \det(\mathbf{\Phi}) \int_0^\infty \int_0^\infty r_1 r_2 \exp(r_1^2 \varphi_{11} + r_2^2 \varphi_{22}) \\ &\quad . \exp(-2 r_1 r_2 |\varphi_{12}| \cos(\theta_1 - \theta_2 - \mu)) dr_1 dr_2, \\ &= \frac{1-\lambda^2}{4\pi^2 (1-c^2)} \left[1 - \frac{c \cos^{-1}(c)}{\sqrt{1-c^2}}\right], \end{aligned} \tag{6.5}$$

where $c = \lambda \cos(\theta_1 - \theta_2 - \mu + \pi)$, the correlation coefficient $\lambda = \frac{|\varphi_{12}|}{\sqrt{\varphi_{11}\varphi_{22}}} = \frac{|\psi_{12}|}{\sqrt{\psi_{11}\psi_{22}}}$, $\mu = \angle\varphi_{12} = \angle\psi_{12} + \pi$, and $\angle$ denotes the phase. We can see that the behaviors of the model (6.5) looks very similar to the empirical joint distribution of neighboring phases as in Fig. 6.2.



(a) (b)

Figure 6.2. The empirical joint distribution of two neighboring phases and the joint density function with the parameter values $\mu = 0.15, \lambda = 0.74$ at a particular complex wavelet subband for texture image 'Misc.0002'. (a) empirical distribution, and (b) model.

6.2.4 Distribution of Relative Phase

Theorem 6.2.1. *If the coefficients in a complex wavelet subband are characterized by a complex Gaussian, then the probability density function for the relative phase of the complex coefficients in this subband will be*

$$p(\theta) = \frac{1-\lambda^2}{2\pi(1-c^2)}\left[1 - \frac{c\cos^{-1}(c)}{\sqrt{1-c^2}}\right], \tag{6.6}$$

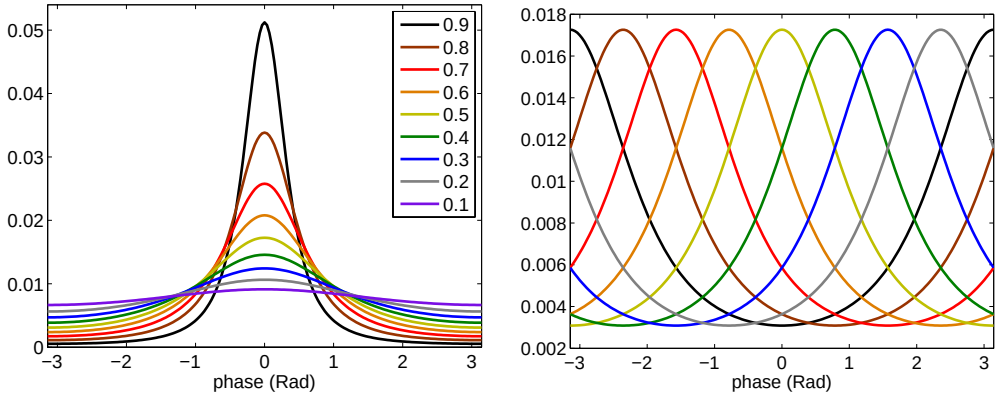
where $c = \lambda\cos(\theta - \mu + \pi)$, $-\pi \leq \theta, \mu \leq \pi$, *and* $0 \leq \lambda \leq 1$.

Proof. If the coefficients in a complex wavelet subband are characterized by a complex Gaussian, the joint distribution $p(\theta_1, \theta_2)$ is given in (6.5). The relative phase can be considered as $\theta = \theta_1 - \theta_2$, where θ_1 is the phase of the reference coefficient at location (i, j) and θ_2 is the phase of the nearest neighbor coefficient at location $(i, j+1)$ or $(i+1, j)$.

Therefore, the distribution of the relative phase θ is given by

$$\begin{aligned} p(\theta) &= \int_{-\pi}^{\pi} p_\Theta(\theta + \theta_2, \theta_2)d\theta_2 \\ &= \frac{1-\lambda^2}{2\pi(1-c^2)}\left[1 - \frac{c\cos^{-1}(c)}{\sqrt{1-c^2}}\right]. \end{aligned}$$

□



(a) (b)

Figure 6.3. Distributions of relative phase. (a) $\lambda = (0.1, 0.2, ..., 0.9)$ and $\mu = 0$, and (b) $\lambda = 0.5$ and $\mu = (-\pi, -3\pi/4, ..., 3\pi/4)$.

The density function of the relative phase θ (RP-PDF) in (6.6) with various values of λ and μ are depicted in Fig. 6.3. The RP-PDF distribution is unimodal with two parameters μ and λ, and is symmetrical about $\theta = \mu$. The parameter μ is the mean direction and the parameter λ is the correlation parameter. The larger the value of the correlation parameter λ, the denser the clustering around μ. Note that $p(\theta; \mu, \lambda)$ and $p(\theta + \pi; \mu, -\lambda)$ are the same distribution. For our model, we set the values of λ to be non-negative, and the range of θ is $[-\pi\ \pi]$. The maximum-likelihood (ML) estimator for RP-PDF distribution will be discussed in section 6.4.

6.3 Complex Gaussian Scale Mixture Model and Phase Distribution

In section 6.2, we proposed the density function in (6.6) for the relative phase within a particular subband with the assumption that the distribution of the real and imaginary coefficients are Gaussian. However, this assumption is often not realistic. A more widely acceptable model is when these coefficients are Gaussian scale mixture (GSM) distributed developed in [22], which is the product of a real Gaussian random vector and an independent hidden random scalar multiplier. The real wavelet coefficients are linked indirectly by their shared dependency on the hidden multiplier. Hence, the GSM model can describe the shape of real wavelet coefficient distributions and the correlation between neighbor co-

efficients. In this section, we will derive the density function of the relative phase with the assumption that the real coefficients and the imaginary coefficients are GSM distributed.

6.3.1 Gaussian Scale Mixture for Real Wavelet Coefficients

In [22][17], the coefficients within each local neighborhood around a reference coefficient of a pyramid subband are characterized by the GSM model. Suppose that the random vector $\mathbf{u}$ has a Gaussian distribution, the scalar real variable $\sqrt{v}$ has some distribution on $(0, \infty)$ with a density $p(v)$ $(v > 0)$, and $\mathbf{u}$ and v are independent. Let $\mathbf{x} \stackrel{\Delta}{=} \sqrt{v}\mathbf{u}$ be defined as the scale mixtures of Gaussian distribution with [95]

$$p(\mathbf{x}) = \int p(\mathbf{x}|v)p(v)dv = \int \frac{\exp\left(\frac{-\mathbf{x}^T(v\mathbf{C_u})^{-1}\mathbf{x}}{2}\right)}{(2\pi)^{\frac{N}{2}}(\det(v\mathbf{C_u}))^{1/2}}p(v)dv,$$

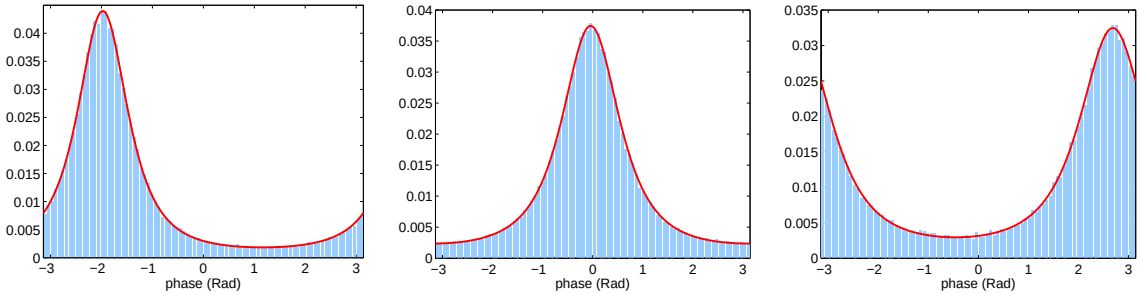
where $\mathbf{C_u} = E[\mathbf{u}\mathbf{u}^T]$ is the covariance matrix of $\mathbf{u} = \left(u_1, u_2, ...u_N\right)^T$, and N is the dimensionality of $\mathbf{u}$ and $\mathbf{x}$.

The conditional density (on v) of $\mathbf{x}$ is Gaussian, and the variable v is known as the multiplier. In general, the neighborhood may include coefficients from other subbands, as well as from the same subbands. The probability density of the multiplier $p(v)$ can be found by using the maximum log likelihood approach for estimating a nonparametric $p(v)$ from an observed set of M neighboring vectors [17]

$$\widehat{p}(v) = \arg\max_{p(v)} \sum_{m=1}^{M} \log \int_0^\infty p(\mathbf{x}_m|v)p(v)dv. \tag{6.7}$$

6.3.2 Complex Gaussian Scale Mixture for Complex Coefficients

The complex directional filters produce real coefficients $\mathbf{x}$ and imaginary coefficients $\mathbf{y}$ which are characterized by a GSM model: $\mathbf{x} \stackrel{\Delta}{=} \sqrt{v}\mathbf{u_r}$ and $\mathbf{y} \stackrel{\Delta}{=} \sqrt{v}\mathbf{u_i}$. Then $\mathbf{z} \stackrel{\Delta}{=} \sqrt{v}\mathbf{u}$, where $\mathbf{z} = \mathbf{x} + j\mathbf{y}$ and $\mathbf{u} = \mathbf{u_r} + j\mathbf{u_i}$. Since the distribution of the real part $\mathbf{x}$ is GSM, $p(\mathbf{x}|v)$ is Gaussian. Similarly for the imaginary part $\mathbf{y}$, $p(\mathbf{y}|v)$ is also Gaussian. Thus from



(a) (b) (c)

Figure 6.4. Distribution of relative phase fitted to the empirical histograms at a particular finest complex wavelet subband. (a) Fabric.0017, $(\mu, \lambda) = (-1.98, 0.77)$ and $\Delta H/H = 0.00018$; (b) Stone.0004, $(\mu, \lambda) = (-0.04, 0.71)$ and $\Delta H/H = 0.00009$; and (c) Brick0004, $(\mu, \lambda) = (2.68, 0.64)$ and $\Delta H/H = 0.00023$.

section 6.2.1, the distribution of complex wavelet coefficients $\mathbf{z}$ will be complex Gaussian when it is conditioned on v as follows

$$p(\mathbf{z}|v) = \frac{\exp\left(-\mathbf{z}^H \mathbf{C}_{\mathbf{z}|v}^{-1} \mathbf{z}\right)}{(\pi)^N \det(\mathbf{C}_{\mathbf{z}|v})}, \tag{6.8}$$

where the covariance matrix $\mathbf{C}_{\mathbf{z}|v} = v\mathbf{C}_{\mathbf{u}}$, $\mathbf{C}_{\mathbf{u}} = E[\mathbf{u}\mathbf{u}^H]$ is complex covariance matrix of $\mathbf{u}$, and N is the dimensionality of $\mathbf{z}$ and $\mathbf{u}$.

The distribution of the vector $\mathbf{u}$ is the complex Gaussian and the scalar real variable $\sqrt{v}$ has some distribution on $(0, \infty)$ with a density $p(v)$ $(v > 0)$. We refer $\mathbf{z} \stackrel{\Delta}{=} \sqrt{v}\mathbf{u}$ as the scale mixtures of complex Gaussian distribution (CGSM) [23] as follows

$$p_{\mathbf{z}}(\mathbf{z}) = \int \frac{\exp\left(-\mathbf{z}^H (v\mathbf{C}_{\mathbf{u}})^{-1} \mathbf{z}\right)}{(\pi)^N \det(v\mathbf{C}_{\mathbf{u}})} p(v) dv. \tag{6.9}$$

Similar to the GSM model, the probability density of the multiplier v can be found as in (6.7).

6.3.3 Joint Distribution of Two Neighboring Phases

We have proved that the coefficients $\mathbf{z} = \mathbf{x} + j\mathbf{y}$ within each local neighborhood around a reference coefficient of a complex subband are characterized by a CGSM model. Now we consider the case of $N = 2$ with $\mathbf{z} = (z_1, z_2)^T$ and

$$p(r_1, r_2, \theta_1, \theta_2|v) = r_1 r_2 \frac{\exp\left(-\mathbf{z}^H \mathbf{C}_{\mathbf{z}|v}^{-1} \mathbf{z}\right)}{(\pi)^2 \det(\mathbf{C}_{\mathbf{z}|v})}, \tag{6.10}$$

where $z_n = r_n e^{j\theta_n}$, $n = 1, 2$, and covariance $\mathbf{C}_{\mathbf{z}|v} = E[\mathbf{z}\mathbf{z}^H|v] = v\mathbf{C}_\mathbf{u} = v \begin{bmatrix} \psi_{11} & \psi_{12} \\ \psi_{12}^* & \psi_{22} \end{bmatrix} = \mathbf{\Phi^{-1}}$.

Hence the joint density of neighboring phases when conditioned on v can be written as

$$\begin{aligned} p(\theta_1, \theta_2|v) &= \pi^{-2} \det(\mathbf{\Phi}) \int_0^\infty \int_0^\infty r_1 r_2 \exp(r_1^2 \varphi_{11} + r_2^2 \varphi_{22}) \\ &\quad . \exp(-2r_1 r_2 |\varphi_{12}| \cos(\theta_1 - \theta_2 - \mu)) dr_1 dr_2, \\ &= \frac{1-\lambda^2}{4\pi^2(1-c^2)} \left[1 - \frac{c \cos^{-1}(c)}{\sqrt{1-c^2}}\right], \end{aligned} \tag{6.11}$$

where $c = \lambda \cos(\theta_1 - \theta_2 - \mu + \pi)$, $\lambda = \frac{|\varphi_{12}|}{\sqrt{\varphi_{11}\varphi_{22}}} = \frac{|v\psi_{12}|}{\sqrt{v^2\psi_{11}\psi_{22}}}$ and $\mu = \angle\varphi_{12} = \angle v\psi_{12} + \pi$. Since v is real and nonnegative, λ and μ are independent from v. Hence

$$\begin{aligned} p(\theta_1, \theta_2) &= \int p(\theta_1, \theta_2|v) p(v) dv, \\ &= \frac{1-\lambda^2}{4\pi^2(1-c^2)} \left[1 - \frac{c \cos^{-1}(c)}{\sqrt{1-c^2}}\right]. \end{aligned} \tag{6.12}$$

6.3.4 Distribution of Relative Phase

Theorem 6.3.1. *If the coefficients in a complex wavelet subband are characterized by the scale mixtures of complex Gaussian model (CGSM), then the probability density function for relative phase of the complex coefficients in this subband will be*

$$p(\theta) = \frac{1-\lambda^2}{2\pi(1-c^2)} \left[1 - \frac{c \cos^{-1}(c)}{\sqrt{1-c^2}}\right], \tag{6.13}$$

where $c = \lambda \cos(\theta - \mu + \pi)$, $-\pi \leq \theta, \mu \leq \pi$ *and* $0 \leq \lambda \leq 1$.

Proof. When the coefficients of a complex subband are characterized by a complex Gaussian or a CGSM model, the joint distribution of neighboring phases in (6.5) and in (6.12) are identical and independent from v. So the distribution of the relative phase shown in (6.6) is also true with the assumption of CGSM model. □

The RP-PDF distribution in (6.6) fits well with the distribution of the RPs in complex subbands. Fig. 6.4 shows an empirical histogram of RP in a particular complex wavelet subband for three different images, along with the best fitting of the RP-PDF

Table 6.1. Average relative entropy of RP-PDF model and histogram, as a fraction of the total entropy of the histogram ($\Delta H/H$) in curvelet domain.

Texture (Vistex)	Subbands	Subband Size	Bins	$\Delta H/H$
640 subimages	3840	64×64	32	0.0021
40 subimages	240	128×128	64	0.0011
40 images	240	256×256	128	0.0006
Lena	6	256×256	128	0.0007
Barbara	6	256×256	128	0.0026
Boat	6	256×256	128	0.0015
Fingerprint	6	256×256	128	0.0003
Peppers	6	128×128	64	0.0015
House	6	128×128	64	0.0027
Cameraman	6	128×128	64	0.0019

distribution. Fitting was performed by maximizing the likelihood function of the relative phase samples within a subband (See in section 6.4).

We also show the relative entropy ΔH (KLD) between the histogram and the model divided by the histogram entropy H in Table 6.1. It is clear that RP-PDF fits very well in the 40 Vistex textures with the size of 256×256, the Lena and Fingerprint images with the average of $\Delta H/H = 0.0007, 0.0008$ and 0.0003 respectively. For other images in Table 6.1, the performance of the fitted model is still acceptable with $\Delta H/H < 0.0027$. It is noted that the images in Fig. 6.4 and Table 6.1 are decomposed by the curvelet transform which is described in section 5.2.

6.3.5 An Extension of Definition for Relative Phase

In this section, we extend the definition of relative phase. The relative phase is the phase difference of two complex wavelet coefficients within a local neighboring in the same subband or in two different subbands. The previous definition in Section 6.2.4 is a special case when two complex coefficients are adjacent. In the general case, the relative phase can be given by

$$\theta(i,j) \quad = \quad \angle z(i,j) - \angle z(i+d_i, j+d_j), \tag{6.14}$$

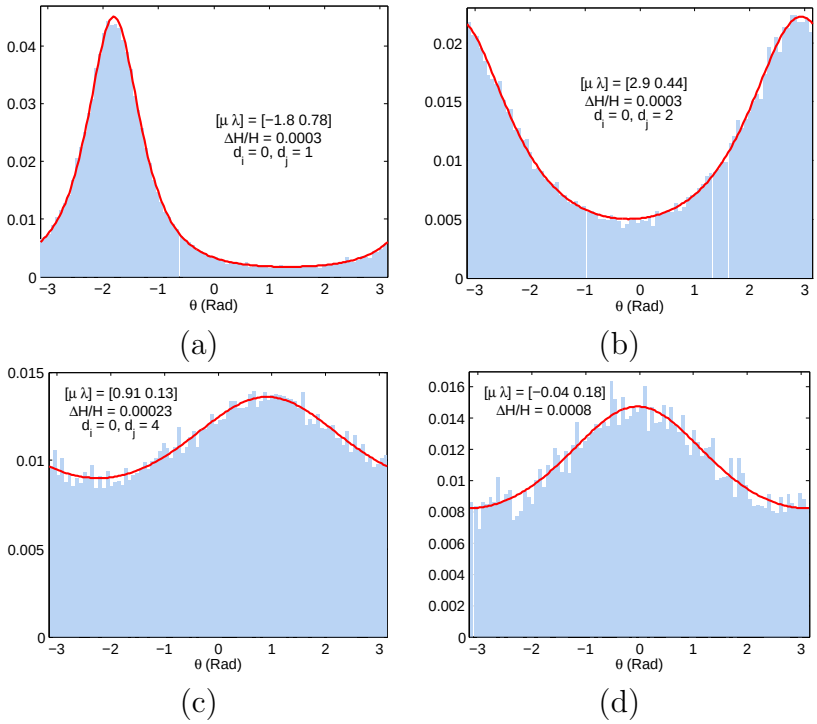


Figure 6.5. Relative phase pdf fitted to the empirical histograms at a particular complex wavelet subband of fingerprint image. In each plot, the estimated parameter values, and the relative entropy $\Delta H/H$ are shown. (a) RP (d = 1), (b) RP (d = 2), (c) RP (d = 4) and (d) RP (parent & children).

where $z(i,j)$ is the coefficient at position (i,j), and d_i, d_j are the distances between two coefficients in row and in column, respectively, e.g., $d_i = 4, d_j = 0$, or $d_i = 2, d_j = 8$.

In equation (6.12), if θ_1 is the phase of a reference phase at location (i,j), and θ_2 is the phase of a neighboring coefficient at location $(i+d_i, j+d_j)$, then the relative phase, $\theta = \theta_1 - \theta_2$, has the probability density function as in (6.13). Some histograms of relative phase in the same subband and fitted models corresponding to various values of distance d_j are shown in Fig. 6.5(a)-(c). When the distance d_j increases, the correlation parameter λ decreases, and vice versa. This conclusion is consistent with the correlation between two coefficients. If the distance d_j is smaller, the correlation between them is higher. Therefore λ is larger and the relative phases clustering around μ is denser. The distribution in Fig. 6.5(c) with $d_j = 4$ is flatter than the distributions with $d_j = 1$ and $d_j = 2$ in Fig. 6.5(a) and (b), respectively. We also show the histogram of relative phase for parent and children coefficients in two scales at the same direction and fitted model in

Fig. 6.5(d). The RP-PDF distribution in (6.13) fits well with the distribution of the RPs with different distances d_j in the same subband as well as in two different subbands.

6.4 Maximum Likelihood Estimator For Distribution of Relative Phase

In this section we describe how to estimate the parameters of RP distribution using the maximum-likelihood estimator (ML). Let $\theta_1, \theta_2, ..., \theta_n$ be a set of observations from a RP-PDF distribution with parameters μ and λ, then $\theta_1, \theta_2, ..., \theta_n$ are i.i.d with pdf

$$p(\theta) = \frac{1-\lambda^2}{2\pi(1-c^2)} \left[1 - \frac{c\cos^{-1}(c)}{\sqrt{1-c^2}}\right], \tag{6.15}$$

where $c = \lambda\cos(\theta - \mu + \pi)$.

The likelihood function is given by

$$L(\mu, \lambda|\theta_1, \theta_2, ..., \theta_n) = \prod_{i=1}^{n} p(\theta_i; \mu, \lambda),$$

and its logarithm,

$$l(\mu, \lambda|\theta_1, \theta_2, ..., \theta_n) = \log L, \tag{6.16}$$

where μ and λ are parameters to be estimated as follows

$$[\hat{\mu}\ \hat{\lambda}] = \arg\max_{[\mu\ \lambda]} \sum_{i=1}^{n} \log p(\theta_i; \mu, \lambda).$$

Differentiating (6.16) with respect to μ and λ, and equating to zero, we obtain the likelihood equations

$$f(\mu) = \frac{\partial l(\mu, \lambda|\theta_1, \theta_2, ..., \theta_n)}{\partial \mu} = 0, \tag{6.17}$$

$$g(\lambda) = \frac{\partial l(\mu, \lambda|\theta_1, \theta_2, ..., \theta_n)}{\partial \lambda} = 0. \tag{6.18}$$

These equations can be solved numerically to find the parameters μ and λ. However, it should be noted that the parameter μ can be also estimated by the mean direction

$$\hat{\mu} = \arctan \frac{\sum_{i=1}^{n} \sin(\theta_i)}{\sum_{i=1}^{n} \cos(\theta_i)}, \tag{6.19}$$

where arctan is the four-quadrant inverse tangent. Therefore, to simplify the estimation problem, we propose using mean direction to estimate μ and using the Newton Raphson iterative method to find solution for $g(\lambda) = 0$ with $\mu = \hat{\mu}$.

Substitute $\hat{\mu}$ into (6.18), the Newton iteration can be stated as

$$\lambda^{[k+1]} = \lambda^{[k]} - \frac{g(\lambda^{[k])}}{g'(\lambda^{[k]})}. \tag{6.20}$$

We derive $g(\lambda)$ and $g'(\lambda)$ in Appendix B.1. They are given by

$$\begin{aligned} g(\lambda) = \; & \textstyle\sum_{i=1}^{n} h(\lambda,-1) + h(\lambda,+1) \\ & -1.5\left[h(\lambda,-x_i) + h(\lambda,x_i)\right] + \frac{b'(\lambda,x_i)}{b(\lambda,x_i)}, \end{aligned} \tag{6.21}$$

$$\begin{aligned} g'(\lambda) = \; & \textstyle\sum_{i=1}^{n} -\left[h^2(\lambda,-1) + h^2(\lambda,+1)\right] \\ & +1.5\left[h^2(\lambda,-x_i) + h^2(\lambda,x_i)\right] + \frac{b"b'-b'^2}{b^2(\lambda,x_i)}, \end{aligned} \tag{6.22}$$

where

$$x_i = \cos(\theta_i - \hat{\mu} + \pi), \quad h(\lambda, x_i) = \frac{x_i}{1+\lambda x_i},$$

$$b(\lambda, x_i) = \sqrt{1-\lambda^2 x_i^2} - \lambda x_i \cos^{-1}(\lambda x_i),$$

$$b'(\lambda, x_i) = -x_i \cos^{-1}(\lambda x_i), \text{ and } \; b"(\lambda, x_i) = \frac{x_i^2}{\sqrt{1-\lambda^2 x_i^2}}.$$

We propose using the correlation coefficient as a good initial value for the root of $g(\lambda)$ as follows

$$\lambda^{[0]} = \frac{|\psi_{12}|}{\sqrt{\psi_{11}\psi_{22}}}, \tag{6.23}$$

where covariance of complex wavelet coefficients in a subband is $\mathbf{C_z} = \begin{bmatrix} \psi_{11} & \psi_{12} \\ \psi_{12}^* & \psi_{22} \end{bmatrix}$.

The Newton-Raphson algorithm can be given step-by-step as follows.

(a) `Initialize` $\lambda^{[0]}$ `using (6.23).`

(b) `Calculate` $g(\lambda^{[k]})$ `using (6.21).`

(c) `Calculate` $g'(\lambda^{[k]})$ `using (6.22).`

(d) `Calculate` $\lambda^{[k+1]}$ `using (6.20).`

(e) `Find the difference between` $\lambda^{[k+1]}$ `and` $\lambda^{[k]}$, $\epsilon = \lambda^{[k+1]} - \lambda^{[k]}$

(f) `If` ϵ `is small, go to step 7. Otherwise,` $k = k+1$, `and go to step 2.`

(g) $\hat{\lambda} = \lambda^{[k+1]}$.

Table 6.2. Average relative entropy of model and histogram ($\Delta H/H$) for the von Mises, wrapped Cauchy and RP-PDF distributions in dual-tree complex wavelet domain.

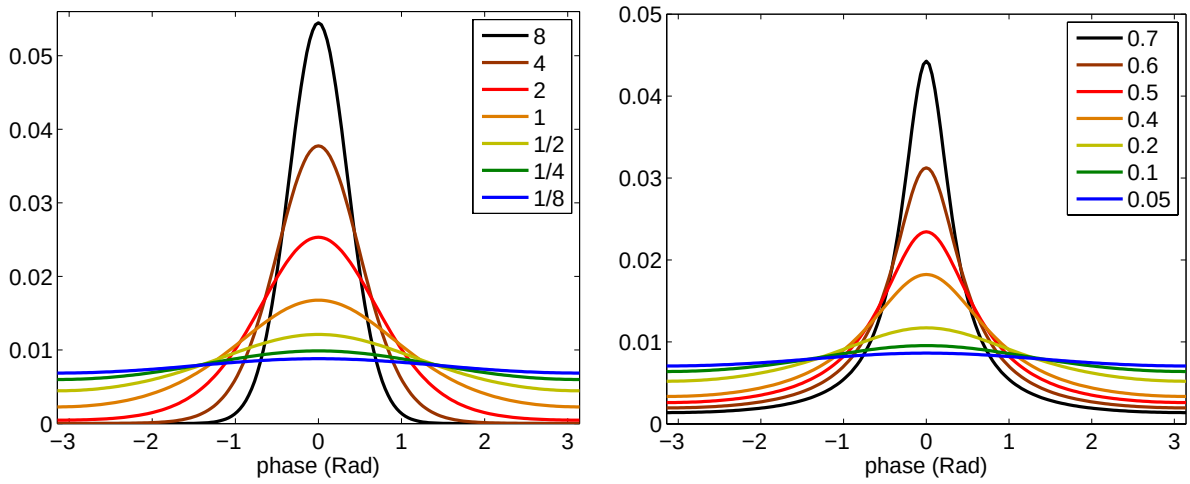
Texture (Vistex)	Von Mises	Wrapped Cauchy	RP-PDF
640 subimages	0.0042	0.0027	**0.0024**
40 subimages	0.0028	0.0015	**0.0014**
40 images	0.0020	0.0010	**0.0008**
Lena	**0.0006**	0.0011	0.0008
Barbara	**0.0013**	0.0017	**0.0013**
Boat	**0.0020**	0.0025	0.0023
Fingerprint	0.0014	0.0007	**0.0003**
Peppers	**0.0013**	0.0020	0.0016
House	0.0024	0.0018	**0.0017**
Cameraman	**0.0019**	0.0022	0.0020

Table 6.3. Average relative entropy of model and histogram ($\Delta H/H$) for the von Mises, wrapped Cauchy and RP-PDF distributions in complex directional filter bank domain.

Texture (Vistex)	Von Mises	Wrapped Cauchy	RP-PDF
640 subimages	0.0139	0.0065	**0.0033**
40 subimages	0.0110	0.0042	**0.0016**
40 images	0.0086	0.0029	**0.0009**
Lena	0.0024	0.0022	**0.0006**
Barbara	0.0026	0.0042	**0.0023**
Boat	0.0044	0.0031	**0.0013**
Fingerprint	0.0098	0.0029	**0.0004**
Peppers	0.0045	0.0028	**0.0020**
House	**0.0043**	0.0073	0.0046
Cameraman	0.0033	0.0025	**0.0012**

Table 6.4. Average relative entropy of model and histogram ($\Delta H/H$) for the von Mises, wrapped Cauchy and RP-PDF distributions in modified curvelet domain.

Texture (Vistex)	Von Mises	Wrapped Cauchy	RP-PDF
640 subimages	0.0103	0.0045	**0.0021**
40 subimages	0.0078	0.0031	**0.0011**
40 images	0.0060	0.0024	**0.0006**
Lena	0.0015	0.0015	**0.0007**
Barbara	**0.0023**	0.0047	0.0026
Boat	0.0041	0.0019	**0.0015**
Fingerprint	0.0092	0.0029	**0.0003**
Peppers	0.0029	0.0023	**0.0015**
House	**0.0019**	0.0052	0.0027
Cameraman	0.0043	0.0027	**0.0019**



(a) (b)

Figure 6.6. Other circular distributions: (a) Von Mises with $\nu = (1/8, 1/4, ...8)$ and $\mu = 0$, and (b) Wrapped Cauchy with $\rho = (0.05, 0.1, ...0.7)$ and $\mu = 0$.

With the initial value $\lambda^{[0]}$ as in (6.23), our ML estimator converges with a few number of iterations. In a practical implementation, we fit 3840 RP-PDFs using the ML estimator. The simulation results showed that the average number of iterations is around 5, and it takes about 0.03 (s) to estimate two parameters for one RP-PDF.

6.5 Comparison with Other Circular Distributions

In this section, the RP-PDF model is compared with other circular distributions including the von Mises [70] and the wrapped Cauchy [71] (see Chapter 5). We also test the RP-PDF model for different complex wavelet transforms including dual-tree complex wavelets (DTCWT) [8], complex directional filter banks (CDFB) [2] and the modified curvelet (See 5.2).

6.5.1 Comparison Results

The von Mises, wrapped Cauchy and RP-PDF distributions fit well with the marginal distribution of the RPs at a subband of the Lena image shown in the first row of Fig. 6.7 with $\Delta H/H = 0.0009, 0.0008$ and 0.0002, respectively. However, for a subband of the Boat image, the von Mises distribution cannot capture the peaky and heavy-tailed distri-

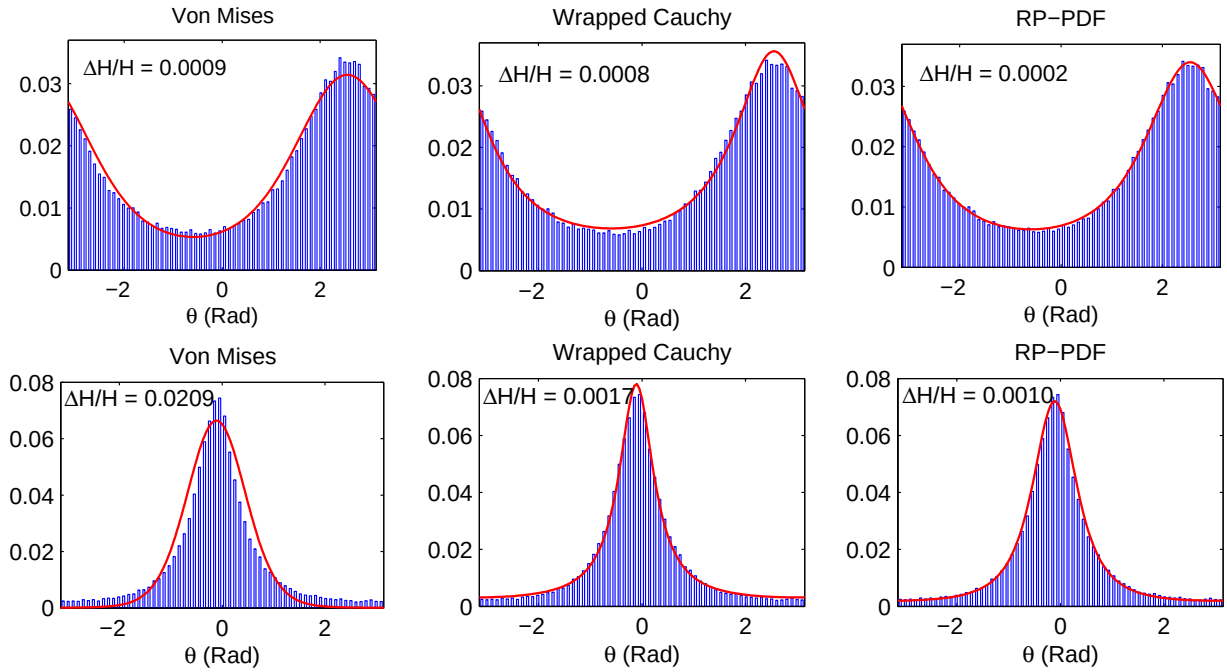
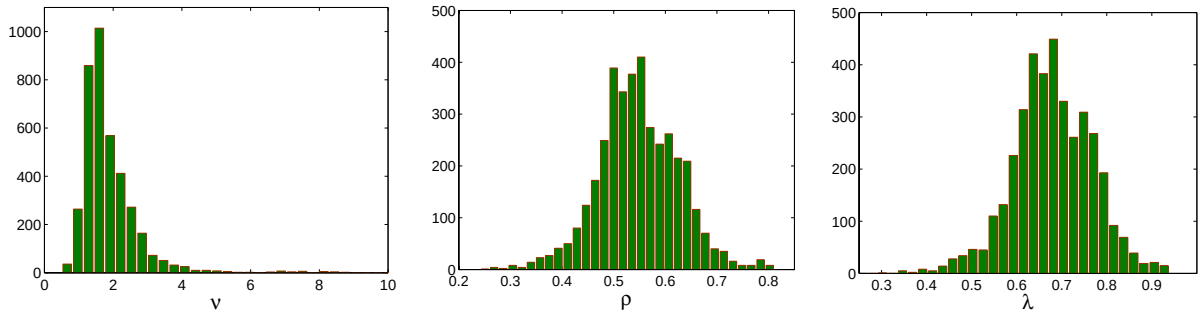


Figure 6.7. Circular distributions fitted to the empirical histograms of relative phase at a particular finest complex wavelet subband. In each plot,the relative entropy $\Delta H/H$ are shown. The first row is the Lena image, and the second one is the Boat image.

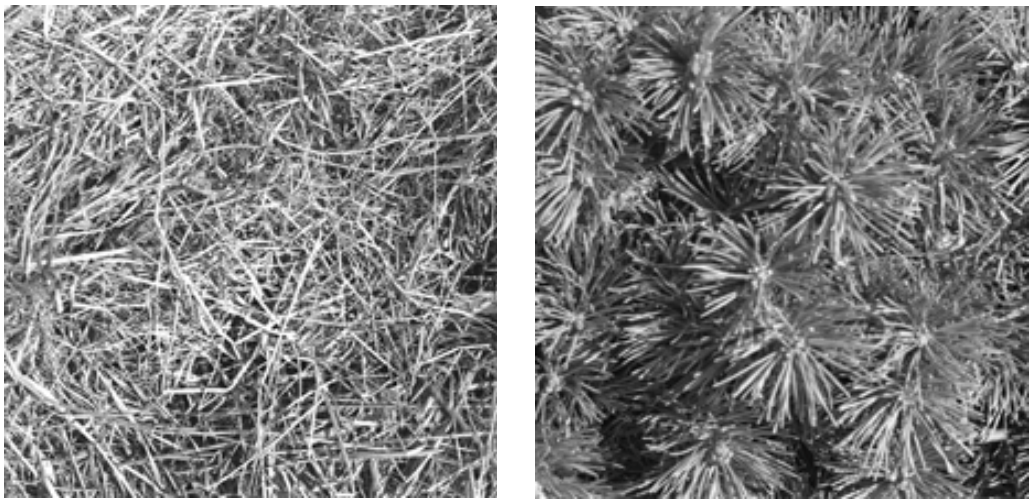
bution as shown in the second row of Fig. 6.7 with $\Delta H/H = 0.0209$. While the RP-PDF model precisely describes the histogram of relative phase with $\Delta H/H = 0.001$. For both examples, the wrapped Cauchy distribution is rather accurate to capture the histogram of relative phases. However when compared with the RP-PDF model, the WC model is less precise. Fitting of the RP-PDF distribution was performed by maximizing a likelihood function proposed in Section 6.4.

We also evaluate the fitted RP-PDF model for many images including texture images as well as natural images, which are decomposed by various complex wavelet transforms such as the DTDWT [8], CDFB [2] and our modified curvelet. The data for evaluation and their information such as the number of subbands, the size of subband and the number of bins are the same as in Table 6.1. The experimental results in Tables 6.2, 6.3 and 6.4 show that for most of the tested images, the RP-PDF model is much more accurate than the von Mises and wrapped Cauchy. For only several images, the accuracy of the fitted fRP model approximates the von Mises.



(a) (b) (c)

Figure 6.8. Histogram of estimated values for the parameters of 3840 finest subbands of size 64×64 from 640 texture images of size 128×128. (a) ν (von Mises), (b) ρ (Wrapped Cauchy) and (c) λ (RP-PDF) .



(a) (b)

Figure 6.9. Two sub-images with the size of 128×128 from VisTex data. (a) Grass.0001 and (b) Leaves.0010.

6.5.2 Range of Estimated Parameter in RP-PDF Model

It is of interest to know the common range for the values of λ in RP-PDF model for texture images as well as of ν in VM model and of ρ in WC model. Fig. 6.8 shows the histograms of the estimated parameters for 3840 finest subbands of size 64×64 from 640 texture images of size 128×128 described in Section 6.6.3. The curvelet transform is used to decompose texture images. For VM model, most of values of estimated ν parameter are from around 1 to 3. The ρ values of the WC model are from 0.45 to 0.65, while the λ values of the RP-PDF model are from 0.55 to 0.8.

Table 6.5. Some examples of the proposed features using statistics of relative phase (RP-PDF parameters μ and λ) shown in Fig. 6.10.

Image	μ_4	μ_5	μ_6	λ_4	λ_5	λ_6
Grass.0001	-2.13	0.01	2.20	0.67	0.72	0.65
Leaves.0010	-1.66	0.07	1.68	0.80	0.80	0.81

6.6 Application to Texture Image Retrieval

In this section, the relative phase distribution model is applied to texture image retrieval. A comparison of various features such as energy feature [7], GGD-based feature [89], RP feature [24], our RP-PDF based feature, and the combination of these features using the curvelet transform in texture retrieval is presented. We also include two other multiresolution directional decompositions in feature extraction, namely Gabor decomposition [7] and CDFB [2]. The latter is very similar to the curvelet transform in terms of directionality of the filters and one-sided supports in the frequency domain.

6.6.1 New Texture Image Feature Using RP-PDF Model

In this image retrieval application, we propose using statistics of relative phase to extract feature for texture images. This feature extraction method is a statistical approach. When compared with the von Mises and the wrapped Cauchy, the RP-PDF is much better in fitting the relative phase pdf for all 640 sub-images from Vistex data in Tables 6.2, 6.3 and 6.4. Therefore, two parameters of the RP-PDF model for relative phase pdfs in each subband will be estimated.

Some examples of both images and their extracted features plus their distributions will be shown. Two sample sub-images, the Grass.0001 and the Leaves.0010, are from Vistex database as in Fig. 6.9. Their extracted features, which are the parameters μ and λ of the RP-PDF (Fig. 6.10), are shown in Table 6.5. With the same subband, two different images have two different relative phase pdfs with different parameters, e.g. at subband 4, the Grass.0001 has the center at $\mu_4 = -2.13$, while the Leaves.0010 has the center at $\mu_4 = -1.66$. The relative phase pdf of the Leaves.0010 is more peaky with the peak of

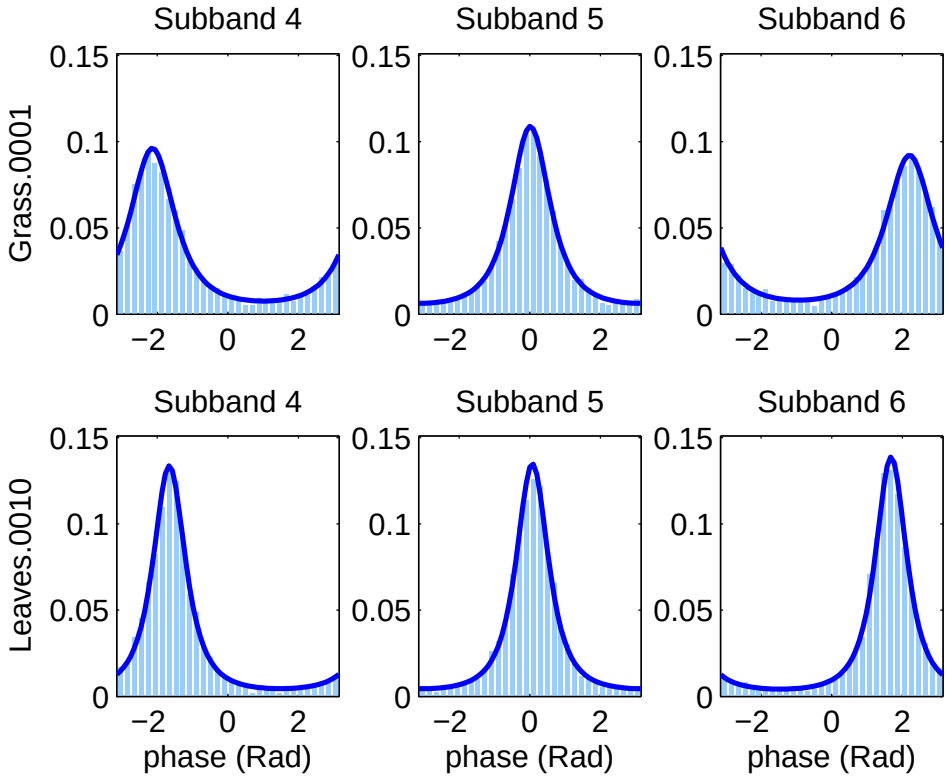


Figure 6.10. The RP-PDF distributions fitted to three empirical histograms of relative phase at three subbands $4, 5$ and 6 with the size of 64×64 from two sub-images shown in Fig. 6.9.

0.15 than that of the Leaves.0011 with the peak of 0.10 at subband 4. It is clear that the distinction between two different types of textures can be captured and recognized by relative phase pdfs.

The proposed feature provides statistical information of relative phase for each subband and can be used to discriminate two textures. Therefore, the wrapped Cauchy model of relative phase pdf will be used as an additional feature for texture image retrieval in this section.

6.6.2 Texture Feature Extraction

Each image in the database is decomposed by the following three decompositions: the curvelet, the 2-D Gabor transform and the CDFB. The Gabor wavelet and the curvelet are applied with four scales and six orientations per scale, while the CDFB has three scales

of eight orientations. For each subband, the mean and standard deviation of the absolute values of the coefficients are calculated as in [7]. The RP feature which includes circular means and standard deviations of the relative phase are computed as in [24] and the GGD features are estimated as in [89]. To construct the RP-PDF based feature vector, the RP matrix of each complex subband in the curvelet domain is created as in (6.14). For each RP matrix, the two parameters μ and λ of the RP-PDF model are estimated by fitting the RP histogram and the RP-PDF density function (6.6). These parameters are used to form the RP-PDF model based feature vector.

For the first experiments (results shown in Table 6.6), the feature vectors are formed from the six subbands at the finest scale of the curvelet transform. The length of the MAG, GGD, RP, RP-PDF feature vectors is twelve because each subband is represented by two parameters. Since the magnitude and phase are combined, the length of the MAG-RP or GGD-RP-PDF feature vector is twenty four. The weighting is needed when we combined the GGD and the RP-PDF feature vector.

In the second experiment (results shown in Table 6.7), in order to obtain a feature vector which has the same dimension as that of the Gabor [7] and the CDFB-RP [24], the GGD-RP-PDF feature vector is formed by twelve features of the RP-PDF model from the finest scale, twenty-four features of the GGD model from the two finest scales, and twelve means of the magnitudes of the curvelet coefficients from the two coarsest scales. For each kind of features, the weighting is needed to get the best performance.

6.6.3 Texture Image Database and Feature Database

We select 40 image textures from the VisTex databases used in [89, 34] for our experiments. Each of these 512×512 images is divided into sixteen 128×128 non-overlapping sub-images, thus creating a database of 640 texture samples. Each original image is treated as a single class and therefore there are 16 samples from each of the 40 classes. To reduce the intensity correlation, all images are normalized to have zero mean and unit variance. For each image in the database, the curvelet transform is applied. The

RP matrix of each subband is created as in (6.14), and their corresponding feature vectors are computed.

6.6.4 Distance Measure and Query Processing

The query pattern can be any one of the texture patterns from the image database. The distance between two magnitude feature vectors and two RP feature vectors are computed as in [24] and the distance between two GGD feature vectors is computed as in [89]. The distance between two RP-PDF feature vectors f_x and f_y is given by

$$d(f_x, f_y) = \sum_k D_{KL}(p(.; \mu_k(x), \nu_k(x))||p(.; \mu_k(y), \nu_k(y))),$$

where k is the index of the subbands and the Kullback-Leibler divergence D_{KL} between two PDFs $p(\theta; \mu_1, \lambda_1)$ and $p(\theta; \mu_2, \lambda_2)$ is defined as

$$D_{KL}(P_1||P_2) = \int_{-\pi}^{\pi} p(\theta; \mu_1, \lambda_1) \log \frac{p(\theta; \mu_1, \lambda_1)}{p(\theta; \mu_2, \lambda_2)} d\theta. \tag{6.24}$$

A closed form for the KLD between the two RP-PDFs is under investigation, and a numerical method with 128 bins is applied to estimate this KLD.

For each query image, N nearest neighbors are selected, and the number of these textures belonging to the same class as the query texture, except for itself, is counted. This number (less than or equal to fifteen) divided by fifteen is defined as the retrieval rate. The performance of the entire class is obtained by averaging this rate over the sixteen members which belong to the same class of texture. The average of all classes is the overall performance of the method.

6.6.5 Experimental Results

Table 6.6 summarizes the overall retrieval rates using the curvelet transform with various features extracted from the finest subbands. If only the top 15 texture images that are nearest to the query texture are considered, and only 12 features are used, the RP-PDF feature and the GGD feature give the best overall retrieval performances of 67.68 % and 64.64 %, while the magnitude and RP features are at 62.07 %, and 60.29 %,

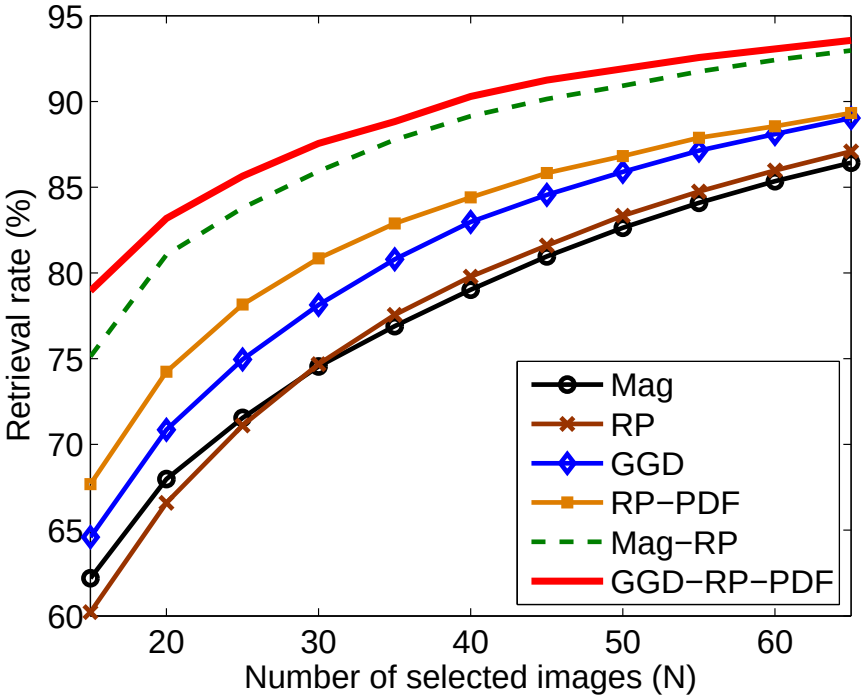


Figure 6.11. Average retrieval rate according to the number of top images considered when the database is 40 VisTex textures. The curvelet transform with various features extracted at the finest scale are used.

respectively. Fig. 6.11 shows the overall performances for the case of N from 15 to 65. It is clear that the feature vector based on the RP-PDF model is consistently better than the magnitude feature and the RP feature. This confirms that the behavior of the RPs is captured accurately by the RP-PDF distribution even with a small number of samples. When the magnitude and phase information are combined, the overall retrieval accuracy of the GGD-RP-PDF feature is also higher than the MAG-RP feature which is formed by the magnitude and the RP features (m_c and σ_c) as proposed in [24] for the case of $N = 15$ as shown in Table 6.6.

In the second experiment, we compare our proposed GGD-RP-PDF feature using the modified discrete curvelet transform with the Gabor and CDFB. In this experiment, all twenty four subbands are used to form feature vectors. If only the top 15 texture images nearest to the query texture are considered, the GGD-RP-PDF gives the best overall retrieval performance of 85.82 % as shown in Table 6.7. The CDFB-RP [24] and

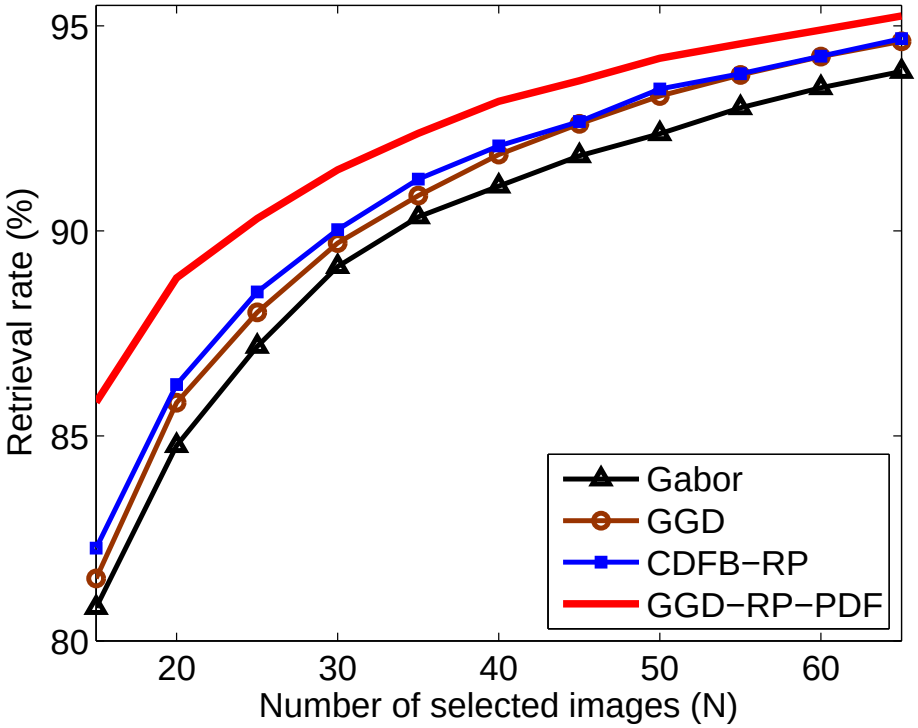


Figure 6.12. Average retrieval rate according to the number of top images considered when various features are used.

Table 6.6. Average retrieval accuracy of 40 VisTex texture images using curvelet transform with various features extracted from six subbands at the finest scale. (MAG denotes the magnitude feature.)

Feature	MAG	GGD	RP	RPPDF	MAG-RP	GGD-RP
Type	m,σ	α,β	m_c,σ_c	μ,λ	m,σ m_c,σ_c	α, β, μ,λ
Length	12	12	12	12	24	24
Feature extraction time	0.076 s	0.262 s	0.081 s	0.218 s	0.118 s	0.441 s
Similarity measurement time (between two images)	0.058 ms	0.198 ms	0.504 ms	2.579 ms	0.562 ms	2.777 ms
N = 15	62.07	64.64	60.29	67.68	74.68	**78.95**

Table 6.7. Average retrieval accuracy of 40 VisTex texture images using various features.

	Gabor	**Cur-GGD**	**CDFB-RP**	**GGD-RP-PDF**
Feature type	m, σ	α, β m, σ	m, σ m_c, σ_c	α, β, μ, λ, m
Feature length	48	48	48	48
N = 15	80.81	81.52	82.26	**85.82**

Table 6.8. Average retrieval rates over the whole database for various existing methods when $N = 15$

Methods	Feature length	Rate
Standard DWT: L1 + L2	18	64.83 %
GGD [89]	18	75.73 %
scalar WD-HMM [34]	33	76.51 %
vector WD-HMM [34]	41	80.05 %
DT-CWT + DT-RCWT [39]	80	81.16 %
CDFB-RP (Chapter 2)	48	82.86 %
GGD-WC(Chapter 5): GGD-WC	48	**85.64 %**
proposed method: GGD-RP-PDF	48	**85.82 %**

the Cur-GGD, which is the GGD feature [89] with the curvelet transform, are at 82.26 % and 81.52 %, while the magnitude based Gabor [7] is at 80.81 %. Fig. 6.12 shows that the overall performances of the GGD-RP-PDF model is consistently better than the others. It is clear that the information gained from the RP-PDF phase model raises the performance of the GGD-RP-PDF significantly higher than those of other features.

The performance of our proposed method is also compared with the performances of different existing methods including energy based feature [39], GGD-based feature [89], WD-HMM based feature [34]. Simulation results in Table 6.8 show that our proposed method outperforms the others. The feature vectors of all previous methods are extracted from the magnitudes or the real values of wavelet subbands, while our method exploits the phase of complex coefficients and extracts image information in both magnitude and phase. The incorporated phase information which is absolutely complementary to magnitude is the reason why the new feature can achieve better performance than the others.

6.6.6 Computational Complexity

The proposed texture retrieval method is implemented in MATLAB R2006a. With an Intel core 2 CPU 2.13 GHz machine, the feature vector time and the similarity measurement time of different features are shown in Table 6.6. The time to extract the proposed feature GGD-RP-PDF with the length of 48 from one image is about 0.515 second (s).

It includes 0.039 s for decomposing image using the modified curvelet transform (4 scales and 6 subbands each scale), 0.259 s for estimating 24 GGD parameters of two finest scales, 0.003 s for computing 12 MAG features of two coarsest scales and 0.214 s for estimating 12 RP-PDF parameters of the finest scales. It take more time to extract the GGD-RP-PDF feature vector than the other features as shown in Table 6.6. This can be a weak aspect of the proposed method. However, the time to extract a feature vector from one image is still much less than 1 (s). Therefore, this proposed feature can be suitable for real-time applications.

With the same feature vector length, the similarity measurement time between two images using the GGD which is 0.198 ms is much less than 2.627 ms when using the RP-PDF. The reason is that there is a closed form of distance between the two GGDs as in [89] while a closed form for the KLD between the two RP-PDFs is developing and we have to use a numerical method with 128 bins to estimate this KLD. In our experiment, with the database of 640 images, it takes about 1.68 s to retrieve 15 images which are similar to the query image. This searching time is still acceptable. However if the database is huge, the numerical method for estimating KLD between the two RP-PDFs is a disadvantage of the searching time of the proposed method. The fast estimation algorithm for KLD between the two RP-PDFs is the future work.

6.7 Conclusion

A new probability density is proposed for modeling the relative phase distribution in complex wavelet domain with the assumptions of Gaussian model as well as Gaussian scale mixture model for real coefficients. The parameters of the new model are estimated using the maximum-likelihood estimator. The RP-PDF distribution captures the behaviors of RPs from various real images including texture images as well as natural images in the complex wavelet domain quite accurately. The experimental results show that the new model is more accurate than the von Mises and Wrapped Cauchy models with various complex wavelet transforms including the modified curvelet, the DTCWT, as well as the

CDFB. Moreover, a new image feature based on the RP-PDF model is proposed for the texture image retrieval application. The new feature captures nicely the directional information from the texture images because higher image retrieval accuracy is achieved by using the new feature instead of using only the magnitude [7] or only the GGD model of real coefficients [89]. In addition to the GGD model of real coefficients [89], the RP-PDF based phase information is incorporated to further improve the performance. The RP-PDF is a promising image model. It exploits the phase information in complex wavelet domain which is rarely used before. Therefore the RP-PDF should be useful in other image processing applications.

CHAPTER 7

SUMMARY AND FUTURE WORKS

7.1 Summary

A new image feature, which we called CDFB-RP, is proposed for feature extraction in Chapter 2. The feature is calculated based on the CDFB decomposition, which has several attractive properties such as multiscale, multi-directional and shiftablity. Moreover, unlike other directional decompositions that require the frequency domain implementation, the decomposition by a separable FB structure. By combining of the magnitude and phase information of the CDFB coefficients, the CDFB-RP feature is used in texture image retrieval. Compared to other directional transforms including the 2-D Gabor wavelet, the contourlet, the steerable pyramid and the CDFB, the CDFB-RP yields best overall performance in classification rate, while keeping the complexity relatively low. The overcomplete ratio of the CDFB-RP is bounded by 8/3 which is much less than those of the Gabor wavelet ($KS = 24$) and the steerable pyramid ($4K/3 = 32/3$), and is only twice that of the contourlet (4/3). Fast decomposition structure and low redundancy make CDFB-RP more efficient in searching and browsing texture images.

The modified version of the PDTDFB has been also proposed for image denoising in Chapter 3. The shiftable properties of this approximately tight-frame decomposition benefits the image denoising application. In comparison to the existing transforms including the wavelet, the contourlet, the PDTDFB yields the best image denoising performance with the thresholding method. Although the overcomplete ratio of the PDTDFB is much lower than this of steerable pyramid, by combining the CGSM model with BLS estimator, PDTDFB could achieve the denoised image quality comparable to steerable pyramid with the BLS-GSM algorithm. In Chapter 4, a new feature extraction method is also proposed for texture segmentation. The approach bases on incorporating the phase infor-

mation obtained from complex filter banks. The complex directional filter bank (CDFB) is used to decompose a texture image in order to provide complex subband coefficients. The local mean direction, extracted from the phases of the coefficients, is defined as additional features for classification and segmentation. Simulation results show that the CDFB phase information is complementary to the magnitude. Lower classification error rates are achieved. Performance of the proposed method is also compared with other complex filter banks including the Gabor transform and the dual-tree complex wavelet.

We develop a new approach which exploits the probabilistic properties from the phase information of two-dimensional complex wavelet coefficients for the image modeling in Chapter 5. Definition, property and statistics of relative phase of the complex coefficients are studied in detail. We proposed the von Mises and the wrapped Cauchy for the probability density function (pdf) of the RP in the complex wavelet domain. The maximum-likelihood method is used to estimate two parameters of the von Mises and the wrapped Cauchy. We demonstrate that the von Mises and the wrapped Cauchy fit well with behaviors of the RP from various real images including texture images as well as natural images. The von Mises and the wrapped Cauchy models are compared, and the experimental results show that for most of the tested images, the wrapped Cauchy model is more accurate than the von Mises, when images are decomposed by different complex wavelet transforms including the dual-tree complex wavelet (DTCWT), the complex direction filter banks (CDFB) and the modified curvelet. Moreover, the statistical models of the relative phase are applied to obtain a new feature for texture image retrieval application. In stead of using only the real or magnitude coefficients, the new approach uses a feature in which phase information is incorporated yielding a higher retrieval accuracy. The relative phase information which is complementary to magnitude would be a promising approach in image processing.

With the assumption of the Gaussian image model as well as the Gaussian scale mixture (GSM) in real wavelet domain, the marginal and joint distributions for the phases of the complex wavelet coefficients are studied in detail. From these hypotheses, we then

derive the probability density function of the relative phase (RP-PDF) in the complex wavelet domain in Chapter 6. We also propose the maximum-likelihood method to estimate two RP-PDF parameters. We demonstrate that the RP-PDF fits well with behaviors of the relative phase from various real images including texture images as well as natural images. The RP-PDF model is compared with other circular distributions including the von Mises and the wrapped Cauchy. The experimental results, in which the real images are decomposed by various complex wavelet such as the dual-tree complex wavelet (DTCWT), the complex directional filter banks (CDFB) and the modified curvelet, show that the RP-PDF model for relative phase is more accurate than the others. Therefore, we propose the RP-PDF for modeling images, especially texture images. Moreover, the RP-PDF model is applied to obtain a new image feature for the texture image retrieval application. In addition to the feature only based on generalized Gaussian density (GGD) for real coefficients, the RP-PDF based feature is incorporated to yield a higher retrieval accuracy.

7.2 Future Works

In this thesis, we derived the probability density function for the relative phase (RP-PDF) in the complex wavelet domain. The RP-PDF is a promising image model. It exploits the phase information in the complex wavelet domain which is rarely used before. Therefore the RP-PDF should be useful in other image processing applications. In the future, we will investigate more applications of the RP-PDF model in bio-informatics and biomedical signals/images.

In Chapter 3, we make a simplifying assumption that $\mathbf{w}$ is a zero-mean complex Gaussian vector. In the general case, $\mathbf{w}$ may or may not be a complex Gaussian vector. Hence, the estimation of noise model is necessary. The probability density of the multiplier, $p_z(z) \propto 1/z$ as shown in [17] is applied to the experiments in this chapter. We can estimate this density. With the true noise model and more accurate CGSM model, the performance of denoised image may be improved. Therefore, estimating the pdf of the multiplier in the CGSM model and the model of noise is the future work.

In Chapters 6, if the database is huge, the numerical method for estimating KLD between the two RP-PDFs is a disadvantage of the searching time of the proposed method. Hence, the fast estimation algorithm for KLD between the two RP-PDFs should be investigated.

APPENDIX A

THE VON MISES AND WRAPPED CAUCHY DISTRIBUTIONS

A.1 Maximum Likelihood Estimation for Von Mises Distribution

Let $\theta_1, \theta_2, ..., \theta_n$ be a set of observations from a von Mises distribution with parameters μ and ν, then $\theta_1, \theta_2, ..., \theta_n$ are i.i.d with pdf

$$p(\theta; \mu, \nu) = \frac{1}{2\pi I_0(\nu)} e^{\nu \cos(\theta - \mu)}, -\pi \leq \theta < \pi.$$

The likelihood function is given by

$$\begin{aligned} L(\mu, \nu | \theta_1, \theta_2, ..., \theta_n) &= \prod_{i=1}^{n} p(\theta_i; \mu, \nu) \\ &= [2\pi I_0(\nu)]^{-n} e^{\sum_{i=1}^{n} \nu \cos(\theta_i - \mu)}, \end{aligned}$$

and its logarithm,

$$l = \log_e L = -n \log_e (2\pi I_0(\nu)) + \nu \sum_{i=1}^{n} \cos(\theta_i - \mu). \tag{A.1}$$

Differentiating (A.1) with respect to μ and ν, and equating to zero, we obtain the likelihood equations

$$\frac{\partial l}{\partial \mu} = \sum_{i=1}^{n} \sin(\theta_i - \mu) = 0, \tag{A.2}$$

$$\frac{\partial l}{\partial \nu} = -nA(\nu) + \sum_{i=1}^{n} \cos(\theta_i - \mu) = 0, \tag{A.3}$$

where $A(\nu) = \frac{I_1(\nu)}{I_0(\nu)}$ and $I_1(\nu) = \frac{dI_0(\nu)}{d\nu}$, the modified Bessel function of order 1.

From Equation (A.2), we have

$$\hat{\mu} = \arctan \frac{\sum_{i=1}^{n} \sin(\theta_i)}{\sum_{i=1}^{n} \cos(\theta_i)}, \tag{A.4}$$

where arctan is the four-quadrant inverse tangent.

From Equation (A.3), the maximum likelihood estimate $\hat{\nu}$ of ν is the solution of

$$A(\nu) = \frac{\sum_{i=1}^{n} \cos(\theta_i - \hat{\mu})}{n} = \bar{R}. \tag{A.5}$$

That is

$$\hat{\nu} = A^{-1}(\bar{R}). \tag{A.6}$$

The solution of (A.6) can only be obtained numerically. The approximate solutions to (A.6) can be obtained as follows

$$\hat{\nu} = \begin{cases} 2\bar{R} + \bar{R}^3 + \frac{5}{6}\bar{R}^5 & : 0 \leq \bar{R} < 0.53, \\ -0.4 + 1.39\bar{R} + \frac{0.43}{(1-\bar{R})} & : 0.53 \leq \bar{R} < 0.85, \\ \frac{1}{(\bar{R}^3 - 4\bar{R}^2 + 3\bar{R})} & : \bar{R} \geq 0.85. \end{cases} \tag{A.7}$$

A.2 Wrapped Cauchy Distribution and Maximum Likelihood Estimation

The wrapped Cauchy (WC) distribution is obtained by wrapping the Cauchy distribution on the real line with density $f(x)$ around the circle, where

$$f(x) = \frac{\sigma}{\pi(\sigma^2 + (x-\mu)^2)}, \qquad -\infty < x < \infty. \tag{A.8}$$

It has the probability density function [88]

$$\begin{aligned} p(\theta) &= \textstyle\sum_{i=-\infty}^{\infty} f(\theta + 2\pi i), \\ &= \tfrac{1}{2\pi}(1 + 2\textstyle\sum_{k=1}^{\infty} \rho^k \cos k(\theta - \mu)), \end{aligned} \tag{A.9}$$

where $\rho = e^{-\sigma}$. By considering the real parts of the geometric series $\sum_{k=1}^{\infty} a^k = \frac{1}{1-a}$ with $a = \rho e^{-j(\theta-\mu)}$, the distribution (A.9) reduces to

$$p(\theta) = \frac{1}{2\pi} \frac{1-\rho^2}{1 + \rho^2 - 2\rho\cos(\theta - \mu)}. \qquad -\pi \leq \theta \leq \pi, \tag{A.10}$$

The location parameter $\pi \leq \mu < \pi$ and scale parameter $0 \leq \rho < 1$ of the WC distribution can be estimated by maximum likelihood. Set

$$\mu_1 = \frac{2\rho\cos\mu}{1+\rho^2}, \quad \mu_2 = \frac{2\rho\sin\mu}{1+\rho^2}, \quad c_3 = \frac{1}{\sqrt{1-\mu_1^2-\mu_2^2}},$$

and

$$\eta_1 = c_3\mu_1, \quad \eta_2 = c_3\mu_2, \text{ and } c_3 = \sqrt{1+\eta_1^2+\eta_2^2}.$$

The WC distribution becomes

$$p(\theta, \eta_1, \eta_2) = \frac{1}{2\pi(\sqrt{1+\eta_1^2+\eta_2^2} - \eta_1\cos\theta - \eta_2\sin\theta)}.$$

Differentiating the $\log_e$ likelihood function with respect to η_1 and η_2 leads to the likelihood equations

$$\frac{1}{c_3}\sum_{i=1}^{n} w_i(\cos\theta_i - \mu_1) = 0, \quad \text{and} \quad \frac{1}{c_3}\sum_{i=1}^{n} w_i(\sin\theta_i - \mu_2) = 0,$$

where $i = 1, 2, ..., n$, and

$$w_i = \frac{1}{1 - \mu_1 \cos\theta_i - \mu_2 \sin\theta_i}. \tag{A.11}$$

These equations can be written as

$$\mu_1 = \frac{\sum_{i=1}^{n} w_i \cos\theta_i}{\sum_{i=1}^{n} w_i}, \quad \text{and} \quad \mu_2 = \frac{\sum_{i=1}^{n} w_i \sin\theta_i}{\sum_{i=1}^{n} w_i}, \tag{A.12}$$

The iterative re-weighting algorithm for maximum likelihood estimator can be given step-by-step as follows.

(a) Initialize $\mu_1^{[0]}$ and $\mu_2^{[0]}$ with $\mu_1^{[0]}+\mu_2^{[0]} < 1$, and calculate $w^{[0]}$ using (A.11).

(b) Given $\mu_1^{[k]}$, $\mu_2^{[k]}$ and $w^{[k]}$ at iteration k, calculate $\mu_1^{[k+1]}$ and $\mu_2^{[k+1]}$ using (A.12).

(c) Repeat step 2 until the algorithm converges, giving $\hat{\mu}_1$ and $\hat{\mu}_2$

(d) Calculate$\hat{\mu}$ and $\hat{\rho}$ by

$$\hat{\mu} = \arctan\frac{\hat{\mu}_2}{\hat{\mu}_1}, \tag{A.13}$$

and

$$\hat{\rho} = \frac{1 - \sqrt{1 - \hat{\mu}_1^2 - \hat{\mu}_2^2}}{\sqrt{\hat{\mu}_1^2 + \hat{\mu}_2^2}}, \tag{A.14}$$

where arctan is the four-quadrant inverse tangent.

APPENDIX B

MAXIMUM LIKELIHOOD ESTIMATION FOR THE RP-PDF

B.1 Derivation of $g(\lambda)$ (6.21) and $g'(\lambda)$ (6.22)

Let $x_i = \cos(\theta_i - \hat{\mu} + \pi)$. From (6.16), we have

$$\begin{aligned} l = \log L = & \sum_{i=1}^{n} \log \tfrac{1-\lambda^2}{2\pi} - 1.5 \log(1-\lambda^2 x_i^2) \\ & + \log[\sqrt{1-\lambda^2 x_i^2} - \lambda x_i \cos^{-1}(\lambda x_i)], \end{aligned}$$

$$\begin{aligned} g(\lambda) = \tfrac{\partial l}{\partial \lambda} = & \sum_{i=1}^{n} \tfrac{1}{1+\lambda} - \tfrac{1}{1-\lambda} - \tfrac{1.5x_i}{1+\lambda x_i} + \tfrac{1.5x_i}{1-\lambda x_i} \\ & + \tfrac{b'(\lambda, x_i)}{b(\lambda, x_i)}, \end{aligned} \tag{B.1}$$

$$\text{where} \quad b(\lambda, x_i) = \sqrt{1-\lambda^2 x_i^2} - \lambda x_i \cos^{-1}(\lambda x_i),$$

$$\text{and} \quad b'(\lambda, x_i) = -x_i \cos^{-1}(\lambda x_i).$$

$$\begin{aligned} g'(\lambda) = & \sum_{i=1}^{n} -\tfrac{1}{(1+\lambda)^2} - \tfrac{1}{(1-\lambda)^2} + \tfrac{1.5x_i^2}{(1+\lambda x_i)^2} + \tfrac{1.5x_i^2}{(1-\lambda x_i)^2} \\ & + \tfrac{b"(\lambda, x_i) b(\lambda, x_i) - b'^2(\lambda, x_i)}{b^2(\lambda, x_i)}, \end{aligned} \tag{B.2}$$

$$\text{where} \quad b"(\lambda, x_i) = \frac{x_i^2}{\sqrt{1-\lambda^2 x_i^2}}.$$

$$\text{Let} \quad h(\lambda, x_i) = \frac{x_i}{1+\lambda x_i}. \tag{B.3}$$

Substituting (B.3) into the (B.1) and (B.2) gives (6.21) and (6.22).

APPENDIX C

ABBREVIATION LIST

2-D	Two Dimensions.
BiShrink	Bivariate Shrinkage.
BLS	Bayes Least Squares.
CDFB	Complex Directional Filter Bank.
CGSM	Complex Gaussian Scale Mixture.
CPU	Central Processing Unit .
CWT	Complex Wavelet Transfrom.
DFB	Directional Filter Bank.
DFT	Discrete Fourier Transform.
DTCWT	Dual-Tree Complex Wavelet.
DWT	Discrete Wavelet Transform.
FB	Filter Bank.
FDCT	Fast Discrete Curvelet Transform.
FS	Full Steerabe Pyramid.
GGD	Generalized Gaussian Density.
GSM	Gaussian Scale Mixtures.
HMM	Hidden Markov Model.
HMT	Hidden Markov Tree.
ICP	Inter-Coefficient Product.
KLD	Kullback-Leibler Divergence .
LMD	Local Mean Direction.
MAG	Magnitude .
MPEG	Moving Picture Experts Group.
MSE	Mean Squared Error.
pdf	Probability Density Function.
PDTDFB	Pyramidal Dual-Tree Directional Filter Bank.

PSNR	Peak Signal-To-Noise Ratio.
QMF	Quadrature Mirror Filter Bank.
RP-PDF	Probability Density Function of Relative Phase.
SLP	SameLevel Product .
SSIM	Structural Similarity.
UDCT	Uniform Discrete Curvelet Transform.
VM	Von Mises.
WC	Wrapped Cauchy.

REFERENCES

[1] R. H. Bamberger and M. J. Smith, "A filter bank for the directional decomposition of images: theory and design," *IEEE Transactions on Signal Processing*, vol. 40, no. 7, pp. 882–893, Apr 1992.

[2] T. T. Nguyen and S. Oraintara, "The shiftable complex directional pyramid, part 1: Theoretical aspects," *Accepted for publication in the IEEE Transactions on Signal Processing*, 2008.

[3] I. W. Selesnick, R. G. Baraniuk, and N. C. Kingsbury, "The dual-tree complex wavelet transform," *IEEE Signal Processing Magazine*, vol. 22, no. 6, pp. 123–151, Nov 2005.

[4] T. Chen and P. Vaidyanathan, "Considerations in multidimensional filter bank design," in *Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS'93)*, May 1993, pp. 643–645.

[5] M. N. Do and M. Vetterli, "The contourlet transform: An efficient directional multiresolution image representation," *IEEE Transactions on Image Processing*, vol. 14, no. 12, pp. 2107–2116, Dec 2005.

[6] E. P. Simoncelli, W. T. Freeman, E. H. Adelson, and D. J. Heeger, "Shiftable multiscale transform," *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 587–607, Mar 1992.

[7] B. Manjunath and W. Ma, "Texture features for browsing and retrieval of image data," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 8, pp. 837–842, Aug 1996.

[8] N. Kingsbury, "Image processing with complex wavelets," *Philosophical Transactions on Royal Society London A*, vol. 357, no. 1760, pp. 2543–2560, Sept 1999.

[9] ——, "Complex wavelets for shift invariant analysis and filtering of signals," *Journal of Applied and Computaional Harmonic Analysis*, vol. 10, no. 3, pp. 234–253, May 2001.

[10] Y. P. Lin and P. P. Vaidyanathan, "Theory and design of two-dimensional filter banks: a review," *Multidimensional Systems and Signal Processing*, vol. 7, no. 3-4, pp. 263–330, Oct 1996.

[11] T. T. Nguyen and S. Oraintara, "A shift-invariant multiscale multidirection image decomposition," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'06)*, vol. 2, May 2006, pp. II–153–II–156.

[12] T. T. Nguyen and H. Chauris, "Uniform discrete curvelet transform," *Submitted for publication in the IEEE Transactions on Image Processing*, 2008.

[13] E. Candès, L. Demanet, D. Donoho, and L. Ying, "Fast discrete curvelet transforms," *Multiscale Modeling Simulation*, vol. 5, no. 3, pp. 861–899, Mar 2006.

[14] S. Mallat, *A wavelet tour of signal processing.* San Diego: Academic Press, 1998.

[15] ——, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674–693, July 1989.

[16] E. P. Simoncelli, *Handbook of video and image processing*, 2nd ed. Academic Press, 2005, ch. 4, pp. 431–441.

[17] J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli, "Image denoising using scale mixtures of Gaussians in the wavelet domain," *IEEE Transactions on Image Processing*, vol. 12, no. 11, Nov 2003.

[18] M. S. Course, R. D. Nowak, and R. G. Baraniuk, "Wavelet-based signal processing using hidden Markov models," *IEEE Transactions on Signal Processing. (Special Issue on Wavelets and Filter banks)*, vol. 46, no. 4, pp. 886–902, Apr 1998.

[19] E. P. Simoncelli, "Modeling the joint statistics of images in the wavelet domain," in *Proceedings of SPIE 44th Annual Meeting*, 1999, pp. 188–195.

[20] J. K. Romberg, H. Choi, and R. G. Baraniuk, "Bayesian tree-structured image modeling using wavelet-domain hidden Markov models," *IEEE Transactions on Image Processing*, vol. 10, no. 7, pp. 1056–1068, Jul 2001.

[21] M. J. Wainwright, E. P. Simoncelli, and A. S. Willsky, "Random cascades on wavelet trees and their use in modeling and analysing natural images," *Journal of Applied and Computational Harmonic Analysis*, vol. 11, pp. 89–123, Jul 2001.

[22] M. J. Wainwright and E. P. Simoncelli, "Scale mixtures of Gaussians and the statistics of natural images," *Advances in Neural Information Processing Systems*, vol. 12, pp. 855–861, 2000.

[23] A. Vo, T. T. Nguyen, and S. Oraintara, "Image denoising using shiftable directional pyramid and scale mixtures of complex Gaussians," in *Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS'07)*, May 2007, pp. 4000–4003.

[24] A. Vo, S. Oraintara, and T. T. Nguyen, "Using phase and magnitude information of the complex directional filter bank for texture image retrieval," in *Proceedings of IEEE International Conference on Image Processing (ICIP'07)*, vol. 4, Sept. 2007, pp. IV–61–IV–64.

[25] T. Randen and J. H. Husøy, "Filtering for texture classification: A comparative study," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 4, pp. 291–310, Apr 1999.

[26] A. Lain and J. Fan, "Texture classification by wavelet packet signatures," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 11, pp. 1186–1191, Nov 1993.

[27] T. Chang and C.-C. J. Kuo, "Texture analysis and classsification with tree-structured wavelet transform," *IEEE Transactions on Image Processing*, vol. 2, no. 4, pp. 429–441, Oct 1993.

[28] J. G. Daugman, "Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 7, pp. 1169–1179, Jul 1988.

[29] M. Porat and Y. Y. Zeevi, "The generalized Gabor scheme of image representation in biological and machine vision," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 10, no. 4, pp. 452–468, Jul 1988.

[30] A. C. Bovik, M. Clark, and W. S. Geisler, "Multichannel texture analysis using localized spatial filters," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 1, pp. 55–73, Jan 1990.

[31] A. K. Jain and F. Farrokhnia, "Unsupervised texture segmentation using Gabor filters," *Pattern Recognition*, vol. 24, no. 12, pp. 1167–1186, Dec 1991.

[32] B. Manjunath, J.-R. Ohm, V. V. Vasudevan, and A. Yamada, "Color and texture descriptors," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 11, no. 6, pp. 703–715, Jun 2001.

[33] H. Greenspan, S. Belongie, R. Goodman, and P. Perona, "Rotation invariant texture recognition using a steerable pyramid," in *Proceedings of the 12th IAPR International Conference on Pattern Recognition (IAPR'94)*, vol. 2, Oct 1994, pp. 162–167.

[34] M. N. Do and M. Vetterli, "Rotation invariant texture characterization and retrieval using steerable wavelet-domain hidden Markov models," *IEEE Transactions on Multimedia*, vol. 4, no. 4, pp. 517–527, Dec 2002.

[35] P. S. Hong, L. M. Kaplan, and M. J. Smith, "A comparison of the octave-band directional filter bank and Gabor filters for texture classification," in *Proceedings of IEEE International Conference on Image Processing (ICIP'04)*, vol. 3, Oct 2004, pp. 1541–1544.

[36] K. Cheng, N. Law, and W. Siu, "Structured and random texture patterns characterization using multiscale directional filter bank," in *Proceedings of 2004 International Symposium on Intelligent Multimedia, Video and Speech Processing (ISIMP'04)*, Oct 2004, pp. 707–710.

[37] Z. He and M. Bystrom, "Color texture retrieval through contourlet-based hidden Markov model," in *Proceedings of IEEE International Conference on Image Processing (ICIP'05)*, vol. 1, Sept 2005, pp. 513–516.

[38] P. de Rivaz and N. Kingsbury, "Complex wavelet features for fast texture image retrieval," in *Proceedings of IEEE International Conference on Image Processing (ICIP'99)*, vol. 1, Oct 1999, pp. 109–113.

[39] M. Kokare, P. K. Biswas, and B. N. Chatterji, "Texture image retrieval using new rotated complex wavelet filters," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 35, no. 6, pp. 1168–1178, Dec 2005.

[40] A. V. Oppenheim and J. S. Lim, "The importance of phase in signals," in *Proceedings of the IEEE*, vol. 69, May 1981, pp. 529–541.

[41] M. C. Morrone and R. A. Owens, "Feature detection from local energy," *Pattern Recognition Letters*, vol. 6, no. 5, pp. 303–313, Dec 1987.

[42] P. Kovesi, "Image features from phase congruency," *Videre: A Journal of Computer Vision Research*, vol. 1, no. 3, pp. 2–26, Summer 1999.

[43] J. G. Daugman, "High confidence visual recognition of persons by a test of statistical independence," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 11, pp. 1148–1161, Nov 1993.

[44] D. Zhang, W.-K. Kong, J. You, and M. Wong, "Online palmprint identification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 9, pp. 1041–1050, Sept 2003.

[45] B. Zhang, S. Shan, X. Chen, and W. Gao, "Histogram of Gabor phase patterns (HGPP): A novel object representation approach for face recognition," *IEEE Transactions on Image Processing*, vol. 16, no. 1, pp. 57–68, Jan 2007.

[46] J. Portilla and E. P. Simoncelli, "A parametric texture model based on joint statistics of complex wavelet coefficients," *International Journal of Computer Vision*, vol. 40, no. 1, pp. 49–71, Oct 2000.

[47] Z. Wang and E. P. Simoncelli, "Local phase coherence and the perception of blur," in *Advances in Neural Information Processing Systems*, vol. 16. MIT Press, May 2004, pp. 786–792.

[48] R. Anderson, N. Kingsbury, and J. Fauqueur, "Coarse-level object recognition using interlevel products of complex wavelets," in *Proceedings of IEEE International Conference on Image Processing (ICIP'05)*, vol. 1, Sept 2005, pp. 3066–3077.

[49] J. F. Ryan Anderson, Nick Kingsbury, "Determining multiscale image feature angles from complex wavelet phases," in *Proceedings of International Conference on Image Analysis and Recognition (ICIAR)*, Sept 2005, pp. 490–498.

[50] X. Zhang and Y. Jia, "Face recognition with local steerable phase feature," *Pattern Recognition Letters*, vol. 27, no. 16, pp. 1927–1933, Dec 2006.

[51] A. Vo, T. T. Nguyen, and S. Oraintara, "Texture image retrieval using complex directional filter bank," in *Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS'06)*, May 2006, pp. 5495–5498.

[52] M. N. Do and M. Vetterli, "Contourlets," in *Beyond Wavelets*, G. V. Welland, Ed. Amsterdam, The Netherlands: Academic Press, 2003, ch. 4, pp. 83–105.

[53] P. Brodatz, *A Photographic Album for Artists and Designers*. Newyork: Dover, 1966.

[54] L. Sendur and I. W. Selesnick, "Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency," *IEEE Transactions on Signal Processing*, vol. 50, no. 11, pp. 2744–2756, Nov 2002.

[55] N. Woodman, "Statiscal analysis based on a certain multivariate complex Gaussian distribution," *The Annals of Mathematical Statistics*, vol. 34, no. 1, Mar 1963.

[56] K. Miller, "Complex Gaussian processes," *SIAM Review*, vol. 11, no. 4, Oct 1969.

[57] H. Teicher, "On the mixture of distributions," *The Annals of Mathematical Statistics*, vol. 31, no. 1, Mar 1960.

[58] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, Apr 2004.

[59] T. Randen and J. H. Husøy, "Multichannel filtering for image texture segmentation," *Optical Engineering*, vol. 33, no. 8, pp. 2617–2625, Aug 1994.

[60] D. Dunn, W. E. Higgins, and J. Wakeley, "Texture segmentation using 2-D Gabor elementary functions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 2, pp. 130–149, Feb 1994.

[61] D. Dunn and W. E. Higgins, "Optimal Gabor filters for texture segmentation," *IEEE Transactions on Image Processing*, vol. 4, no. 7, pp. 947–964, Jul 1995.

[62] M. Unser, "Texture classification and segmentation using wavelet frames," *IEEE Transactions on Image Processing*, vol. 4, no. 11, pp. 1549–1560, Nov 1995.

[63] S. Krishnamachari and R. Chellappa, "Multiresolution Gauss-Markov random field models for texture segmentation," *IEEE Transactions on Image Processing*, vol. 6, no. 2, pp. 251–267, Feb 1997.

[64] H. Choi and R. G. Baraniuk, "Multiscale image segmentation using wavelet-domain hidden Markov models," *IEEE Transactions on Image Processing*, vol. 10, no. 9, pp. 1309–1321, Sept 2001.

[65] Y. Xia, D. Feng, and R. Zhao, "Adaptive segmentation of textured images by using the coupled Markov random field model," *IEEE Transactions on Image Processing*, vol. 15, no. 11, pp. 3559–3566, Nov 2006.

[66] D. Charalampidis and T. Kasparis, "Wavelet-based rotational invariant roughness features for texture classification and segmentation," *IEEE Transactions on Image Processing*, vol. 11, no. 8, pp. 825–837, Aug 2002.

[67] K.-H. Liang and T. Tjahjadi, "Adaptive scale fixing for multiscale texture segmentation," *IEEE Transactions on Image Processing*, vol. 15, no. 1, pp. 249–256, Jan 2006.

[68] X. Liu and D. Wang, "Image and texture segmentation using local spectral histograms," *IEEE Transactions on Image Processing*, vol. 15, no. 10, pp. 3066–3077, Oct 2006.

[69] A. Vo and S. Oraintara, "Using phase and magnitude information of the complex directional filter bank for texture segmentation," in *Proceedings of The 16th European Signal Processing Conference (EUSIPCO'08)*, Aug 2008.

[70] A. Vo, S. Oraintara, and T. T. Nguyen, "Statistical image modeling using von Mises distribution in the complex directional wavelet domain," in *Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS'08)*, May 2008, pp. 2885–2888.

[71] A. Vo and S. Oraintara, "On the distributions of relative phase of complex wavelet coefficients," *Submitted to the 2009 IEEE International Symposium on Circuits and Systems (ISCAS09)*.

[72] ——, "A study of relative phase in complex wavelet domain: Property, statistics and application," *Submitted to Signal Processing: Image Communication*, 2008.

[73] P. Moulin and J. Liu, "Analysis of multiresolution image denoising schemes using generalized Gaussian and complexity priors," *IEEE Transactions on Information Theory*, vol. 45, no. 3, pp. 909–919, Apr 1999.

[74] E. P. Simoncelli, "Bayesian denoising of visual images in the wavelet domain," in *Bayesian Inference in Wavelet Based Models*, P. Müller and B. Vidakovic, Eds. Springer-Verlag, New York, 1999, ch. 18, pp. 291–308.

[75] H. A. Chipman, E. D. Kolaczyk, and R. E. Mcculloch, "Adaptive Bayesian wavelet shrinkage," *Journal of the American Statistical Association*, vol. 92, pp. 1413–1421, Dec 1997.

[76] S. LoPresto, K. Ramchandran, and M. T. Orchard, "Wavelet image coding based on a new generalized Gaussian mixture model," in *Data Compression*, Snowbird, Utah, Mar 1997.

[77] K. Mihak, I. Kozintev, K. Ramchandran, and P. Moulin, "Low-complexity image denoising based on statistical modeling of wavelet coefficients," *IEEE Signal Processing Letters*, vol. 6, no. 12, pp. 300–303, Dec 1999.

[78] H. Choi, J. K. Romberg, R. G. Baraniuk, and N. Kingbury, "Hidden Markov tree modeling of complex wavelet transforms," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'00)*, vol. 1, Jun 2000, pp. 133–136.

[79] J. K. Romberg, H. Choi, R. G. Baraniuk, and N. Kingbury, "Multiscale classification using complex wavelets and hidden Markov tree models," in *Proceedings of IEEE International Conference on Image Processing (ICIP'00)*, vol. 2, Sept 2000, pp. 371–374.

[80] C. W. Shaffrey, N. G. Kingsbury, and I. H. Jermyn, "Unsupervised image segmentation via Markov trees and complex wavelets," in *Proceedings of IEEE International Conference on Image Processing (ICIP'02)*, vol. 3, Jun 2002, pp. 801–804.

[81] F. Yan, L. Cheng, and S. Peng, "Dual-tree complex wavelet hidden Markov tree model for image denoising," *Electronics Letters*, vol. 43, pp. 973–975, Aug 2007.

[82] S. Rahman, M. Ahmad, and M. Swamy, "Statistics of 2d dt-cwt coefficients for a Gaussian distributed signal," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, Accepted for publication.

[83] J. K. Romberg, H. Choi, and R. G. Baraniuk, "Multiscale edge grammars for complex wavelet transforms," in *Proceedings of IEEE International Conference on Image Processing (ICIP'01)*, vol. 1, Oct 2001, pp. 614–617.

[84] G. Hua and M. T. Orchard, "Image inpainting based on geometrical modeling of complexwavelet coefficients," in *Proceedings of IEEE International Conference on Image Processing (ICIP'07)*, vol. 1, Oct 2007, pp. 553–556.

[85] ——, "Image reconstruction from the phase or magnitude of its complex wavelet transform," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'08)*, Apr 2008, pp. 3261–3264.

[86] M. Miller and N. Kingsbury, "Statistical image modelling using interscale phase relationships of complex wavelet coefficients," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'06)*, vol. 2, May 2006, pp. II–789–II–792.

[87] K. V. Mardia and P. E. Jupp, *Directional Statistics*. John Wiley and Sons, 2000.

[88] S. Jammalamadaka and A. SenGupta, *Topics in circular statistics*. World Scientific, 2001.

[89] M. N. Do and M. Vetterli, "Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance," *IEEE Transactions on Image Processing*, vol. 11, no. 2, pp. 146–158, Dec 2002.

[90] A. Vo, S. Oraintara, and T. T. Nguyen, "Statistical image modeling using distribution of relative phases in the complex wavelet domain," in *Proceedings of The 16th European Signal Processing Conference (EUSIPCO'08)*, Aug 2008.

[91] ——, "Probability distribution of the relative phase for image modeling in complex wavelet domain," *Submitted to IEEE Transactions on Signal Processing*, 2008.

[92] M. G. A. Thomson, "Visual coding and the phase structure of natural scenes," *Network: Computation in Neural Systems*, vol. 10, no. 2, p. 123132, May 1999.

[93] P. Kovesi, "Phase congruency: a low-level image invariant," *Psychologycal Research*, vol. 64, no. 2, pp. 136–48, 2000.

[94] M. C. Morrone and D. C. Burr, "Feature detection in human vision: A phase-dependent energy model," *Proceedings of the Royal Society of London. Series B, Biological Sciences.*, vol. 235, pp. 221–245, Dec 1988.

[95] D. Andrews and C. Mallows, "Scale mixtures of normal distributions," *Journal of the Royal Statistical Society*, vol. 36, no. 1, pp. 99–102, 1974.

BIOGRAPHICAL STATEMENT

An Phuoc Nhu Vo received the B.S and the M.S degrees in Electrical Engineering from HCM City University of Technology, Hochiminh city, Vietnam, in 1997 and 2000, respectively. She worked in the HCMC University of Technology as a lecturer in the Department of Electrical Engineering from 1997 to 2004. She has been a PhD student in the Department of Electrical Engineering at University of Texas at Arlington since Summer 2004. Her research interests are in digital signal/image processing, wavelets and filter banks with applications in image analysis and modeling, biomedical imaging, bioinformatics, communications and control systems. She is a student member of the IEEE.