

2000

Texture representation using wavelet filterbanks

Nam-Deuk Kim
Iowa State University

Follow this and additional works at: <https://lib.dr.iastate.edu/rtd>

 Part of the [Computer Sciences Commons](#), and the [Electrical and Electronics Commons](#)

Recommended Citation

Kim, Nam-Deuk, "Texture representation using wavelet filterbanks " (2000). *Retrospective Theses and Dissertations*. 12338.
<https://lib.dr.iastate.edu/rtd/12338>

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

**Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600**

UMI[®]

Texture representation using wavelet filterbanks

by

Nam-Deuk Kim

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of
DOCTOR OF PHILOSOPHY

Major: Electrical Engineering (Communications and Signal Processing)

Major Professor: Satish Udpa

Iowa State University

Ames, Iowa

2000

Copyright © Nam-Deuk Kim, 2000. All rights reserved.

UMI Number: 9990463

UMI[®]

UMI Microform 9990463

Copyright 2001 by Bell & Howell Information and Learning Company.

All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.

Bell & Howell Information and Learning Company

300 North Zeeb Road

P.O. Box 1346

Ann Arbor, MI 48106-1346

Graduate College
Iowa State University

This is to certify that the Doctoral dissertation of
Nam-Deuk Kim
has met the dissertation requirements of Iowa State University

Signature was redacted for privacy.

Committee Member

Signature was redacted for privacy.

Major Professor

Signature was redacted for privacy.

For the Major Program

Signature was redacted for privacy.

For the Graduate College

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	xiii
CHAPTER 1 INTRODUCTION	1
1.1 What is Texture?	1
1.2 Applications of Texture Analysis	2
1.3 Issues in Texture Analysis	4
1.4 Scope of this dissertation	5
1.5 Outline of this dissertation	6
CHAPTER 2 RELATED WORK ON TEXTURE	9
2.1 Statistical Approach	9
2.2 Model-based Approach	10
2.2.1 Gaussian Markov Random Field	11
2.3 Filtering-based Approaches	14
2.3.1 Fourier Transform	14
2.3.2 Gabor Filters	16
2.3.3 Wavelet Transform-based Methods	19
2.4 Texture Classification Performance Measure	19
CHAPTER 3 WAVELETS AND TEXTURE ANALYSIS	22
3.1 The Continuous Wavelet Transform	22
3.2 Discrete Wavelet Transform	23
3.3 Multiresolution Decomposition Theory	25
3.4 Quadrature Mirror Filter (QMF)	27
3.5 Extension to Two-Dimensional Wavelet Transforms	29

3.6	Stationary Wavelet Transform	30
3.7	Wavelet Packets	31
3.8	Tree-Structured Wavelet Transform (TSWT)	33
3.9	Summary	34
CHAPTER 4 ORIENTED TWO-DIMENSIONAL WAVELET FILTERS .		35
4.1	2-D Wavelet transforms using 2-D Filters	35
4.2	Rotated Wavelet Filters	36
4.3	Dual Wavelet Filter Bank	40
4.4	Summary	41
CHAPTER 5 WAVELET-BASED TEXTURE CLASSIFICATION		42
5.1	Wavelet-based Texture Features	42
5.2	Experimental Implementation	45
5.2.1	Experiment I and II	45
5.2.2	Experiment III	47
5.2.3	Bayes Distance Classifier	48
5.3	Results and Discussions	50
5.3.1	Experiment I	50
5.3.2	Experiment II	52
5.3.3	Experiment III	53
5.4	Summary	54
CHAPTER 6 VIEWPOINT-INVARIANT TEXTURE CLASSIFICATION		55
6.1	Introduction	55
6.2	Grayscale Transform-Invariance	57
6.3	Subband Energy Distribution	58
6.4	Rotation-invariant feature extraction	61
6.4.1	Scale Energy-Based Features	61
6.4.2	Directionality Feature	63
6.5	Experimental Setup	64

6.6	Feature Analysis	65
6.6.1	Scale Energy-based Features	65
6.6.2	Directionality Measure	68
6.6.3	Feature Selection	71
6.7	Classification Results	74
6.7.1	Experiment I	74
6.7.2	Experiment II	77
6.8	Summary	78
CHAPTER 7 UNSUPERVISED TEXTURE SEGMENTATION USING		
FEATURE DIMENSION REDUCTION		80
7.1	Introduction	80
7.2	Unsupervised Texture Segmentation Algorithm in Wavelet Domain	82
7.2.1	Feature Extraction	82
7.2.2	Selection of parameter σ	84
7.2.3	Channel Selection using Histogram Analysis	85
7.2.4	Feature Clustering	87
7.2.5	Texture Segmentation	87
7.3	Summary	89
CHAPTER 8 TEXTURE ANALYSIS FOR TENDON INJURY EVALU-		
ATION		91
8.1	Introduction	91
8.2	Tendon Image Acquisition and Clinical Evaluation	93
8.2.1	Lesion Segmentation for Cross-sectional Images	96
8.2.2	Ultrasonic Fiber Measurement for Longitudinal Images	96
8.3	Results	97
8.4	Summary	99
CHAPTER 9 CONCLUSIONS		101
9.1	Major Contributions	101

9.2 Major Findings	103
9.3 Suggestions for Further Research	103
BIBLIOGRAPHY	105

LIST OF TABLES

Table 5.1	Percent of correct classification results obtained in experiment I as a function of the number of decomposition scale levels I	51
Table 5.2	Percent of correct classification results obtained in experiment II as a function of the number of decomposition scale levels.	52
Table 5.3	Percent of correct classification obtained in experiment III (thirty VisTex images) as a function of the number of decomposition scale levels.	53
Table 6.1	Nine feature combinations studied in the experiment.	74
Table 6.2	Classification performance of viewpoint-invariant texture features for thirteen Brodatz textures using two different energy measures.	75
Table 6.3	Classification performance for each texture class using ten selected features.	76
Table 6.4	Classification performance of features with rotation-invariance only.	78
Table 6.5	Performance of the proposed feature set for 30 VisTex texture classification.	78
Table 8.1	Percentage of damaged SDFT and the fiber alignment measure for the ROIs in fig. 8.2. The numbers in parentheses are the subjective grades.	98
Table 8.2	Results of fiber alignment grading for longitudinal image analysis. *standard deviation	99

LIST OF FIGURES

Figure 1.1	Examples of images with texture content. (a) natural scene image with trees, sky, and clouds. (b) some of the texture images from well-known Brodatz texture album and MeasTex database.	2
Figure 1.2	Image formation from texture surface. θ =rotation, ϕ =view angle, δ =scale, λ =illumination.	6
Figure 2.1	Neighbor sets for the first- through fourth-order GMRF models.	12
Figure 2.2	Texture images generated by using the fourth-order GMRF model. Lower figures are power spectrum densities used to obtain corresponding upper texture images.	14
Figure 2.3	Examples of power spectrum of a relatively fine texture (a) and a coarse texture (b). The difference in coarseness of the original images (upper row) is apparent in their power spectrums (lower row).	15
Figure 2.4	Variation of Gabor filters due to changes in center frequency u_0 . The other parameters are fixed at $\theta = 0^\circ$, $\alpha_x = 0.2$, $\alpha_y = 0.1$	17
Figure 2.5	Variation of Gabor filters due to rotation angle θ . The other parameters are fixed at $u_0 = 2$, $\alpha_x = 0.2$, $\alpha_y = 0.2$	18
Figure 3.1	Examples of Daubechies wavelet family representing low to high frequency (from left to right). (a) $a = 2^2$. (b) $a = 2^1$. (c) $a = 2^0$	24
Figure 3.2	Schematic representation of the ladder of subspaces in wavelet transforms.	26
Figure 3.3	A structure of the discrete wavelet transform using Quadrature Mirror Filters.	28

Figure 3.4	An example of a two level DWT of the well-known ‘Lena’ image (a). The letters in (b) indicate the corresponding frequency components of the detail images.	30
Figure 3.5	Partitions of frequency domain resulting from a 2-level decomposition using DWT. The letters correspond to the ones in Fig. 3.4(b)	31
Figure 3.6	A block diagram of stationary wavelet transforms using Quadrature Mirror Filters. (a) 1-D SWT: (b) 2-D SWT, $I = 2$	32
Figure 3.7	Schematic representation of the subspaces in the wavelet packet decomposition.	33
Figure 3.8	Example of a possible four-level TSWT decomposition. Subimages with high energy are further decomposed.	34
Figure 4.1	Angle reference for the 2-D wavelet transforms used in this study. . . .	37
Figure 4.2	Design of rotated wavelet filters (filter length $L = 4$). (a) Rotated 2-D filter. Empty cells are zeros: (b) The filter coefficient values for four 2-D RWFs.	38
Figure 4.3	Frequency partitioning for the stationary wavelet filters (a) and rotated filters (b).	39
Figure 4.4	Demonstration of the one-level decomposition using SWF and RWF for a natural texture image “D34”. (a) Original image: (b) Outputs of SWF: (c) Outputs of RWF. For (b) and (c), from the top left clockwise, filtered image using H_0 , H_1 , H_2 , and H_3	40
Figure 4.5	A diagram of the dual wavelet filterbank with scale level three. The upper bank represents a SWFB and the lower represents a RWFB. The output of H_0^s filter is further decomposed into eight subbands at the next level.	41

Figure 5.1	Brodatz texture data used in experiment I. First row: D3, D4, D6, D9, second row: D11, D16, D19, D21, third row: D24, D28, D29, D32, fourth row: D34, D52, D53, D55, fifth row: D57, D65, D68, D74, sixth row: D77, D78, D82, D84, seventh row: D92, D93, D95, D103.	46
Figure 5.2	Visually similar pattern textures used in experiment II: D4, D29, D32, D57, D77.	47
Figure 5.3	VisTex images for experiment III: from left to right: first row: Bark0, Bark4, Bark6, Bark8, Bark9; second row: Brick1, Brick4, Brick5, Fabric0, Fabric4; third row: Fabric7, Fabric9, Fabric11, Fabric13, Fabric16; fourth row: Fabric17, Fabric18, Food0, Food2, Food5; fifth row: Food8, Grass1, Sand0, Stone4, Tile1; sixth row: Tile3, Tile7, Water6, Wood1, Wood2.	49
Figure 6.1	Energy distribution at the SWFB outputs $k = 4, 5, 6$ with the "brick" image as an input texture.	59
Figure 6.2	Subband energy distribution from a SWFB for selected texture images that have different directionality. Multiple lines in each plot indicate seven rotation angles: 0, 30, 60, 90, 120, 150, and 200 degrees.	60
Figure 6.3	Different view of Fig. 6.2 for three directional textures. From top row to bottom, scale $I = 1, 2,$ and $3,$ respectively.	61
Figure 6.4	Frequency coverage for scale sum energy feature, F_{SSE} for (a) SWFB and (b) RWFB.	62
Figure 6.5	Visualization of SEF feature.	63
Figure 6.6	Textures used in the experiment. Each texture was digitized at seven different rotation angles: 0, 30, 60, 90, 120, 150, and 200 degrees. The size displayed here is 256×256	66
Figure 6.7	Examples of rotated texture image "straw". Sixteen non-overlapping samples of size 128×128 are shown.	67
Figure 6.8	Overall diagram of rotation-invariant texture classification procedure. .	68

Figure 6.9	Scale sum energy features from the output of (a) SWFB and (b) RWFB. The F_{SSE}^3 was amplified two times for better visualization.	69
Figure 6.10	Scale energy ratio features from the output of (a) SWFB and (b) RWFB.	70
Figure 6.11	Correlation plot of F_{RSEER}^2 feature with thirteen texture classes.	71
Figure 6.12	Normalized distance D for both filter banks and the summation of the two for directionality measurement (F_{DM}^3) of “wood” texture.	72
Figure 6.13	Correlation of F_{DM}^i with thirteen texture classes, where i denotes the scale level. (From top to bottom) F_{DM}^1 ; F_{DM}^2 ; F_{DM}^3 . Note the variations across the scale level.	73
Figure 6.14	Example of tree-pruning results obtained using four perceptual texture features.	77
Figure 7.1	(From left to right) (a1-a2) A 128×128 artificial image and region labels; (b1-b3) channel outputs, $r_n(x, y)$, $n = 1, 3$, and 5 , respectively; (c1-c3) image histograms of (b).	82
Figure 7.2	(From left to right) (a1-a3) Feature images $e_n(x, y)$ that correspond to $r_n(x, y)$ in Fig. 7.1, where $n = 1, 3$, and 5 , respectively; (b1-b3) image histograms of (a) with $\sigma = 2.5$	83
Figure 7.3	Block diagram of an image processing system for wavelet-based texture segmentation.	85
Figure 7.4	Illustration of the effect of σ for Gaussian function on the feature image and its histogram. The first row represents $e_2(x, y)$ with increasing σ (from left to right), and the second row represents the corresponding histograms.	86
Figure 7.5	Segmentation results of an artificial image using $e_n(x, y)$: (a) $n = 1, 2$, and 3 ; (b) $n = \text{all}$; (c) $n = 1$	88

- Figure 7.6 Segmentation experiment for a natural texture image. (a) original image (256×256); (b) segmentation result using selected channels ($n = 1, 2, 4, 7$); (c) segmentation result using the all channels; (d) histograms of $e_n(x, y)$ for the selected channels used for (b). (d1) $n = 1$. $\sigma = 2.5$; (d2) $n = 2$, $\sigma = 2$; (d3) $n = 4$, $\sigma = 2.5$; (d4) $n = 7$, $\sigma = 3$ 90
- Figure 8.1 Sagittal view of the metacarpus. A. Frozen section. B. Schematic of frozen section showing zones of scanning. DD = deep digital flexor tendon; SD = superficial digital flexor tendon. 94
- Figure 8.2 Typical ultrasound images of equine leg showing SDFT and DDFT. Transducer is at top edge of images. Actual image size is 4 cm wide and 7 cm deep (309×426 pixels). First row: Cross-sectional images from zone 1A through 3B. The subjective grades are 1, 1, 3, 3, 2, and 2, respectively. Second row: Corresponding longitudinal view images. The fiber alignment grades are 1, 1, 3, 3, 3, and 2. The upper regions of interest (ROI) represent the SDFT and the lower ROIs represent the DDFT. The transducer is at the top edge of the images. 95
- Figure 8.3 Injury evaluation procedure for two types of B-scan tendon images. . . 96
- Figure 8.4 Lesion segmentation results. (a1-a6) The SDFTs from cross-sectional images in Fig. 8.2; (b1-b6) the corresponding results indicating the percentage of the injured SDFT. 98

ACKNOWLEDGEMENTS

I thank Lord for being my light and giving me a strength and wisdom whenever I needed. Without the guidance from my Lord, I would not be what I am now.

This work would not have been possible without help and support from my major professor Dr. Satish Udpa who has strong enthusiasm, engineering insight, and creative ideas. I would like to thank Dr. Viren Amin at the center for NDE for his support and guidance during my Masters program and thereafter. I am also thankful to Dr. Fritz Keinert, who has shared his expertise in wavelet theory for my research: Dr. Lalita Udpa, who has provided me with new ideas and support while working in the same group: Dr. Julie Dikerson, Dr. Paul Sacks, and Dr. William Lord for their helpful comments. Without their help and support, I would never have finished this study.

I would also like to thank Dr. R. Chandramouli at Stevens Institute of Technology for introducing me to multimedia and related issues which have strongly interested me. It has been a great pleasure to work with friends in the Material Assessment Research Group. The friendship and the time we shared will be remembered for a long time. The times with my best friends in the Korean community will also be in my memory for a long time.

I cannot thank my parents enough for their endless love and support. Their lives have taught me to understand the value of hard work and sharing. I am also in great debt to my brothers and sisters for their love and support throughout my life. Finally, I thank my wife, Jungyun, for her understanding and support, and my two sons, Hyunrae and Wonrae for giving me smiles and big hugs.

CHAPTER 1 INTRODUCTION

The ability to represent texture is the single most important step in the development of systems for measuring the similarity of textures and segmenting images on the basis of differences in textures.

1.1 What is Texture?

A formal definition of texture does not exist, although texture is an intuitive concept. It is said in [29] that texture has been extremely refractory to precise definition. Bovik *et al.* [2] suggest that an exact definition of texture either as a surface property or as an image property has never been adequately formulated. Despite this lack of a universally agreed definition, textures are often described as homogeneous patterns of spatial arrangements of pixels that regional intensity or color alone cannot sufficiently describe. They may consist of the structured and/or random placement of elements, but also may be without fundamental subunits. For example, Fig. 1.1 illustrates a variety of real-world images where texture contributes considerably to the “look and feel” of the image.

Some researchers describe texture in terms of the human visual system: that textures do not have uniform intensity, but are nonetheless perceived as homogeneous regions by a human observer. For example, Chaudhuri *et al.* [6] notes that :

Texture regions give different interpretations at different distances and at different degrees of visual attention. At a standard distance with normal attention, it gives the notion of macroregularity that is characteristic of the particular texture. When viewed closely and attentively, homogeneous regions and edges, sometimes constituting texels, are noticeable.

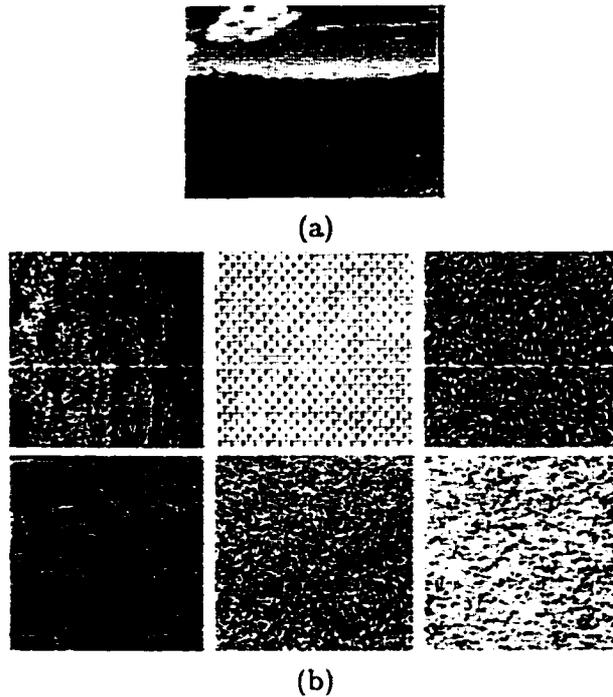


Figure 1.1 Examples of images with texture content. (a) natural scene image with trees, sky, and clouds. (b) some of the texture images from well-known Brodatz texture album and MeasTex database.

However, a definition based on human acuity poses problems when used as the theoretical basis for a quantitative texture analysis algorithm. Faugeras and Pratt [23] note that the basic pattern and repetition frequency of a texture sample could be perceptually invisible, although quantitatively present.

Most natural surfaces exhibit texture, for example, clouds, trees, bricks, grass, and water images can be described as textures. Therefore a texture recognition system will be a natural part of many computer vision systems. Application domains of texture analysis are described in the following section.

1.2 Applications of Texture Analysis

There exists a wide range of the successful applications of texture recognition systems. An overview article by Tuceryan and Jain [79] contains more information about applications of texture analysis. We discuss a few examples briefly in this section.

Remote Sensing Remote sensing is the measurement of object properties far from the object. Remote sensing techniques include satellite photography, sonar surveys, seismic surveys, etc. A number of texture recognition approaches have been applied to images obtained by remote sensing techniques for the purpose of terrain classification [82], cloud classification [43], sea floor mapping [49], and seismic pattern recognition [64].

Medical Imaging This field has benefited significantly from texture analysis techniques. Ultrasound, x-ray, tomography, and Magnetic Resonance Image (MRI) images are routinely analyzed with respect to their texture properties in medical diagnosis applications. Applications include classification of ultrasound liver(normal/abnormal) images [85], lesion detection in mammographic images for early cancer detection [27], ultrasound beef quality grading [37], and abdominal organ segmentation on MRI [39].

Document Analysis Most of the information available is either recorded on paper or in the form of photographs and videos. This information needs to be digitized into images and the text converted to ASCII for storage, retrieval, and manipulation to build a digital library. The text regions of printed documents are perceived as texture when viewed from a sufficient distance. Texture segmentation techniques have been successfully applied for discriminating text regions from other objects such as figures and plots [86, 33]. The extracted text regions are then subsequently analyzed using optical character recognition (OCR) systems for character recognition.

Content based image retrieval Digital media and, in particular, images and video data, present many challenging problems in storage, access, and distribution. Large collections of multimedia documents can be found in diverse application domains such as the entertainment industry, education, medical imaging, and geographic information systems. However, locating an image (or video frame) is not straightforward. Content based image search schemes have been an evolving research area in computer vision for the past few years. Since some content

types, such as grass, trees, leaves, hair, sky, water, and brick wall have texture as an important characteristic, texture analysis methods have been successfully applied to content based image retrieval systems [50, 88, 56, 72].

Digital image watermarking With the rapid growth of Internet technologies and availability of multimedia computing facilities, the enforcement of multimedia copyright protection becomes an important issue. Digital watermarking is viewed as an effective way to deter content users from illegal distribution [13]. One approach to achieve this goal in digital image watermarking involves adding an invisible information (called watermark) to a 'host image' to mark ownership of the image. The watermark must be imperceptible (transparent) yet must be unaffected (robust) by signal processing methods that attempt to remove the embedded watermark. However, these are conflicting requirements where a suitable trade-off needs to be made. By applying texture segmentation algorithms, content based watermarking techniques can be developed [34]. It is shown that by embedding different amounts of key information depending on the texture of a region improves both the robustness and perceived quality of the watermarked image [34].

1.3 Issues in Texture Analysis

As discussed above, texture recognition systems find a large number of applications. While a considerable research has been carried out in the field of texture analysis, the problems related to texture processing have still only been partially solved and active research is continuing. In this section, we address a few important issues in texture analysis area.

First, there are no widely used performance metrics for texture analysis algorithms, as evidenced by the lack of quantitative comparisons in literature. It is obvious that obtaining a "good" feature set is the most important step for texture classification and segmentation. However, the bases for comparison which involves choice of texture databases, size of input samples, preprocessing technique, classification algorithms, feature selection method, and amount of training and testing data set. These are generally not consistent in each experiment. A

different database is employed even among the comparative studies, making the problem more complicated.

Second, perceptual meanings of texture features are not clearly provided, which implies either the characteristics of the features are not accurately measured or the features do not have a physical connection to texture properties, such as “coarseness”, “directionality”, and “regularity”. This needs to be taken into account during the feature extraction procedure.

The last issue involves view-point invariant properties of texture features. Textures in real-world images are very often distorted by artifacts such as those resulting from non-uniform lighting (illumination), shading, or different view angles. This makes the problem of texture analysis more difficult. The variables associated with image formation in the case of real world images is illustrated in Fig. 1.2. It is very difficult to obtain texture features that are independent of all possible variables simultaneously. Until recently, successful studies on this important issue have been limited.

1.4 Scope of this dissertation

The main focus of this dissertation is on wavelet-based approaches for efficient texture classification and segmentation. To achieve this goal, the dissertation addresses the following issues.

1. **Oriented wavelet filter design** Conventional wavelet transforms result in four subimages with a mixture of the two orientations (horizontal/vertical). However, it has limited freedom with regard to orientation tuning for image analysis. This problem can be alleviated by using oriented wavelet filters by providing information that complements information generated via standard wavelet filters. In this study, a new oriented wavelet filter set capable of obtaining better characterization of oriented textures is described.
2. **Viewpoint invariant texture representation** We address one of the most important and difficult issues in this study. We assume that the view angle is perpendicular to the texture surface ($\phi = 90^\circ$) and the scale is restricted ($\delta = \text{constant}$). However, the

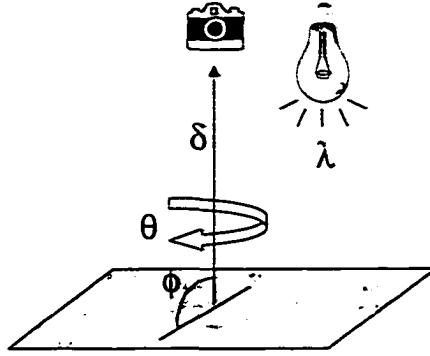


Figure 1.2 Image formation from texture surface. θ =rotation, ϕ =view angle, δ =scale, λ =illumination.

rotation (θ) is not restricted, and we explore texture representations that approximate the rotation invariance since it is one of the important view-point invariance problems. We also do not restrict the illumination (λ) parameter and obtain illumination-invariance using rotation-invariant texture features.

3. **Perceptually meaningful texture feature extraction** As the importance of digital libraries are increasing, a feature set that is capable of representing texture properties is also becoming increasingly important. We take this issue into account during the feature extraction process.
4. **Feature dimensionality reduction for efficient texture segmentation** An important issue in the design of a multichannel filtering system for texture segmentation is the selection of useful channels. Channels that contain weak discriminatory information do not improve the overall segmentation performance. An efficient channel selection method will be proposed for addressing the unsupervised texture segmentation problem.

1.5 Outline of this dissertation

This dissertation can be divided into three parts; the introductory part (Chapter 1, 2, and 3), the methodology part (Chapter 4, 6, and 7), and the application part (Chapter 5, 6, 8).

Note that Chapter 6 includes both methodology and application. An overview of the chapters is given in this section.

Chapter 1: Introduction to the problems and a brief overview of applications

Chapter 2: A summary of popular texture analysis methods is given. In addition, a survey covering a broad range of prior work relative to each method is provided.

Chapter 3: A description of a standard wavelet transform is given. Furthermore, wavelet frames, wavelet packet, and tree-structured wavelet transform used for characterizing texture are discussed. Literature survey on the previous work relating to wavelet and texture analysis is also given.

Chapter 4: Two oriented wavelet transforms are described and a new wavelet filterbank is proposed for efficient characterization of texture images. A combination of the new filterbank and the standard wavelet filterbank constitutes a dual wavelet filterbank (DWFB) with which texture can be more efficiently characterized.

Chapter 5: The effectiveness of the filterbank proposed in Chapter 4 is demonstrated for a texture classification problem using two different texture databases which include a variety of texture patterns.

Chapter 6: A set of new DWFB-based features is proposed for rotation-invariance and grayscale transform-invariance problems. Experimental evidence showing effectiveness of the feature set are provided.

Chapter 7: A feature reduction method for a multichannel-based segmentation approach is developed. The feature vector dimensions are reduced by discarding channels that contain weak discriminatory power.

Chapter 8: An application of wavelet analysis to the tendon injury evaluation of racehorse is demonstrated. The segmentation algorithm developed in Chapter 7 is applied to localize the damaged tendon area. The effectiveness of wavelet-based energy features for tendon fiber alignment is also demonstrated.

Chapter 9: A few concluding remarks are presented. A summary of the findings and contributions of this study, followed by a list of areas where further investigation may lead to improvement in the proposed texture analysis algorithms is included.

CHAPTER 2 RELATED WORK ON TEXTURE

A variety of methods for texture feature extraction have been proposed during the last few decades. A good overview of this subject is found in [79]. Texture analysis, in general, can be classified into three categories: statistical, structural, and spectral. However, we prefer to use the terms ‘statistical’, ‘model-based’, and ‘filtering-based’, since many popular and successful approaches belongs to one of these categories.

This chapter summarizes concepts underlying a few popular texture analysis methods in each category, and discusses issues of comparative performance among algorithms. Popular methods include co-occurrence matrix, Markov Random Field model, Fourier transform, Gabor filter set, and the wavelet filterbank method.

2.1 Statistical Approach

Statistical approaches rely on the assumption that texture can be defined by local statistical properties of pixel gray levels. One of the simplest approaches for describing textures is to use moments of the gray-level histogram of a texture image [26]. In 1973, Haralick *et al.* [29] proposed a set of fourteen features from a gray-level cooccurrence matrix (GLCM) whose elements represent estimates of the probability of transitions from one level to another in a given direction (θ) at a given interpixel distance (δ). The probability measure is defined by $Pr(x) = \{C_{ij}(\delta, \theta)\}$, where the GLCM C_{ij} is defined by

$$C_{ij} = P_{ij} / \sum_{i,j=1}^M P_{ij},$$

P_{ij} represents the number of occurrences of gray level g_i and g_j , and M is the total possible number of gray level pairs.

The size of \mathbf{C} is determined by the number of distinct gray levels in the input image. However, to reduce the computational demands for the GLCM feature extraction, the image intensities need to be quantized into fewer gray levels (typically from 256 to 16 or 32 levels). This reduces the size of GLCMs and causes a dramatic decrease in computation time.

Once \mathbf{C} is computed for pre-determined δ and θ (normally 0° , 45° , 90° , and 135°), the problem is to analyze the matrix \mathbf{C} to categorize the texture of the region over which it is computed. Energy, entropy, correlation, local homogeneity, and inertia are a commonly used subset of the original 14 features introduced by Haralick *et al.* [29]. These are defined as

$$\text{energy} = \sum_{i,j} C_{ij}^2 \quad (2.1)$$

$$\text{entropy} = -\sum_{i,j} C_{ij} \log C_{ij} \quad (2.2)$$

$$\text{correlation} = \sum_{i,j} \frac{(i - \mu_x)(j - \mu_y)C_{ij}}{\sigma_x \sigma_y} \quad (2.3)$$

$$\text{local homogeneity} = \sum_{i,j} \frac{1}{1 + (i - j)^2} C_{ij} \quad (2.4)$$

$$\text{inertia} = \sum_{i,j} (i - j)^2 C_{ij}, \quad (2.5)$$

where μ_x and σ_x are the mean and standard deviation of the row sums of the matrix \mathbf{C} , and μ_y and σ_y are the corresponding statistics of the column sums.

The GLCM approach has been successfully applied to various applications, such as ultrasound liver image classification [85], terrain classification [82], and texture image retrieval [47]. It is also shown that a more accurate classification can be obtained by using a set of interpixel distances [12].

The drawbacks of the GLCM approach include issues such as deciding the number of quantization levels and selecting optimal interpixel distance (δ) and orientations (θ) that can be obtained by conducting a few experiments, *a priori*.

2.2 Model-based Approach

In the model-based approach, a texture is assumed to be a realization of a stochastic process that can be described by a few parameters. Hence, texture analysis is viewed as a parameter

estimation problem: given a texture image, estimate the parameters of the assumed random process. The estimated parameters can serve as features for texture classification [14, 7, 35, 36] and segmentation [19, 55, 18] problems. In addition, a synthetic texture image can be created by sampling the process. However, the difficulty with this approach is that many natural textures do not conform to the restrictions of a particular model.

Cross and Jain [14] demonstrate one of the first successful texture models using a Markov Random Field (MRF) model. They show how to estimate model parameters and demonstrate that the MRF models fit some of the textures reasonably well. Since their work, a growing number of researchers have used MRF models for texture analysis. Other model-based schemes proposed to date include fractal models [62] and Wold models [50].

2.2.1 Gaussian Markov Random Field

Let \mathbf{X} denote the random field defined on Ω and Γ denote the set of all possible configurations of \mathbf{X} on Ω . \mathbf{X} is a Markov Random Field on Ω with respect to the neighborhood η of site r if the following conditions are satisfied.

1. *Markov Property* $P(\mathbf{X} = \mathbf{x}) > 0$ for all $\mathbf{x} \in \Gamma$.
2. *Positivity condition* $P(x_s | x_r, r \in \Omega, r \neq s) = P(x_s | x_r, r \in \eta_s)$.
3. *Homogeneous property* $P(x_s | x_r, r \in \eta_s)$ is the same for all sites r .

The condition (1) states the probability of \mathbf{X} having a certain value at a particular pixel, given the values of \mathbf{X} elsewhere, is dependent only on the values of \mathbf{X} in the neighborhood of the pixel. The condition (2) states that all combinations of pixel values are possible.

A Gaussian Markov Random Field (GMRF) model is a special case of MRF, where it is assumed that texture intensity values can be modeled by a Gaussian Markov Random Field. GMRF models have been shown to perform well in both the classification and segmentation of textured images [7, 55]. In the GMRF model, the texture $y(m, n)$ at pixel (m, n) is modeled by the following equation.

$$y(m, n) = \sum_{k, l \in N_s} \theta_{k, l} y(m - k, n - l) + e(m, n) \quad (2.6)$$

		4	3	4		
	4	2	1	2	4	
	3	1	⊙	1	3	
	4	2	1	2	4	
		4	3	4		

Figure 2.1 Neighbor sets for the first- through fourth-order GMRF models.

where N_s is a neighborhood set and $e(m, n)$ is a stationary Gaussian noise sequence with zero mean and autocorrelation function given by

$$R_e(k, l) = \begin{cases} \sigma^2, & \text{if } (k, l) = (0, 0) \\ -\sigma^2 \theta_{k, l}, & \text{if } (k, l) \in N_s \\ 0, & \text{otherwise} \end{cases} \quad (2.7)$$

The definition of neighborhood N_s controls the characteristics of the model, and the order of the model is one if GMRF assumes that each pixel value depends only on those pixels immediately adjacent to it (i.e. $N_s = \{(0, 1), (0, -1), (1, 0), (-1, 0)\}$), while the second-order model assumes the diagonally adjacent pixels also affect its value, i.e. $N_s = \{(0, 1), (0, -1), (1, 0), (-1, 0), (1, 1), (1, -1), (-1, 1), (-1, -1)\}$. The neighbor sets for the first- to fourth-order GMRF models are shown in Fig. 2.1.

To further investigate the properties of the model, the Z-transform of Eq. (2.6) can be written by

$$Y(z_1, z_2) = \sum_{k, l \in N_s} \theta_{k, l} Y(z_1, z_2) z_1^{-k} z_2^{-l} + E(z_1, z_2). \quad (2.8)$$

Hence, the transfer function $H(z_1, z_2)$ is

$$H(z_1, z_2) = \frac{E(z_1, z_2)}{Y(z_1, z_2)} = \frac{1}{1 - \sum_{k, l \in N_s} \theta_{k, l} z_1^{-k} z_2^{-l}} = [A(z_1, z_2)]^{-1}. \quad (2.9)$$

Then, the PSD of GMRF is

$$S_y(z_1, z_2) = S_e(z_1, z_2) / [A(z_1, z_2) A(z_1^{-1}, z_2^{-1})]. \quad (2.10)$$

But, from the Eq. (2.7), the PSD of $e(m, n)$ is

$$S_e(z_1, z_2) = \sigma^2 \left[1 - \sum_{k,l \in \mathcal{N}_s} \theta_{k,l} z_1^{-k} z_2^{-l} \right] = \sigma^2 A(z_1, z_2). \quad (2.11)$$

Then, we have $S_y(z_1, z_2) = \sigma^2 / A(z_1^{-1}, z_2^{-1})$. Due to the symmetric structure of the neighborhood defined and the θ parameters, $A(z_1, z_2) = A(z_1^{-1}, z_2^{-1})$. Hence,

$$S_y(z_1, z_2) = \sigma^2 / A(z_1, z_2). \quad (2.12)$$

Evaluating $S_y(z_1, z_2)$ on the unit circles $z_1 = e^{j\omega_1}$, $z_2 = e^{j\omega_2}$ yields the power spectral density

$$S_y(\omega_1, \omega_2) = \frac{\sigma^2}{1 - \sum_{k,l \in \mathcal{N}_s} \theta_{k,l} \exp(-j(k\omega_1 + l\omega_2))}. \quad (2.13)$$

Thus, GMRF can be modeled by filtering a stationary Gaussian noise sequence $e(m, n)$ using the transfer function $1/A(z_1, z_2)$. From Eq. (2.13), it is also observed that GMRF can be completely characterized by the set of parameters $\{\underline{\theta}, \sigma^2\}$. An example of textures obtained using the GMRF model is shown in Fig. 2.2. The upper figures are created by using a fourth-order GMRF, and the lower plots correspond to power spectrum densities which are obtained from Eq. (2.13). The set of parameters $\{\underline{\theta}, \sigma^2\}$ used are in [7].

Given a texture, we wish to determine the parameters of a GMRF model which could be used for reconstructing the original texture. The two popular methods for estimation are the Least Square method and Maximum Likelihood Estimate method [36].

Although GMRF models have been applied successfully for texture analysis, a number of issues need to be addressed. For example, the model parameters do not uniquely represent a texture, the computational complexity is high for parameter estimation, and not all textures can be modeled as Gaussian random processes. It is found that GMRF model parameters are not superior to filtering-based texture analysis methods for texture classification applications [67, 60].

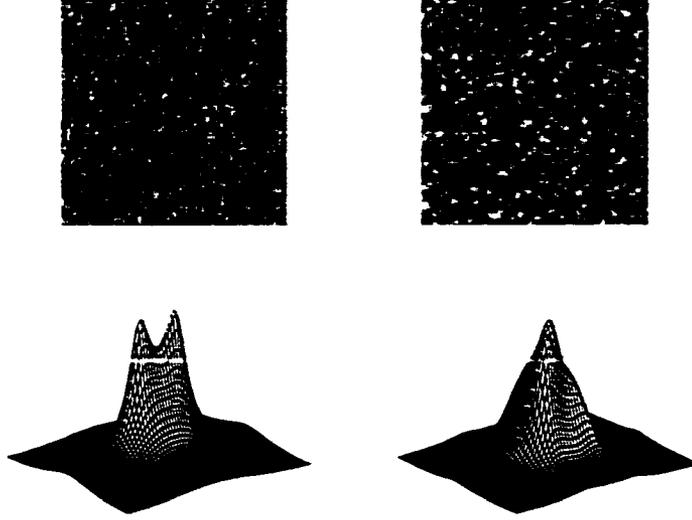


Figure 2.2 Texture images generated by using the fourth-order GMRF model. Lower figures are power spectrum densities used to obtain corresponding upper texture images.

2.3 Filtering-based Approaches

2.3.1 Fourier Transform

Fourier spectrum-based approaches have been one of the popular spectral approaches, since the Fourier spectrum is ideally suited for describing the directionality of periodic or almost periodic 2-D patterns in an image [26]. Such methods are used primarily to detect global periodicity in an image by identifying high-energy, narrow peaks in the spectrum.

The Fourier transform of an image $f(x, y)$ is defined by [26]

$$F(u, v) = \int \int_{-\infty}^{\infty} f(x, y) \cdot e^{-j2\pi(ux+vy)}, \quad (2.14)$$

and its corresponding discrete Fourier transform is given by

$$F(u, v) = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m, n) \cdot e^{-j2\pi(um+vn)}. \quad (2.15)$$

where M and N are the ranges along the two axes. The power spectrum is computed by $|F|^2 = FF^*$ (where $*$ denotes the complex conjugate).

It is known that the radial distribution of values in $|F|^2$ is sensitive to texture coarseness [82]. For example, a coarse texture would have high values of $|F|^2$ concentrated near the

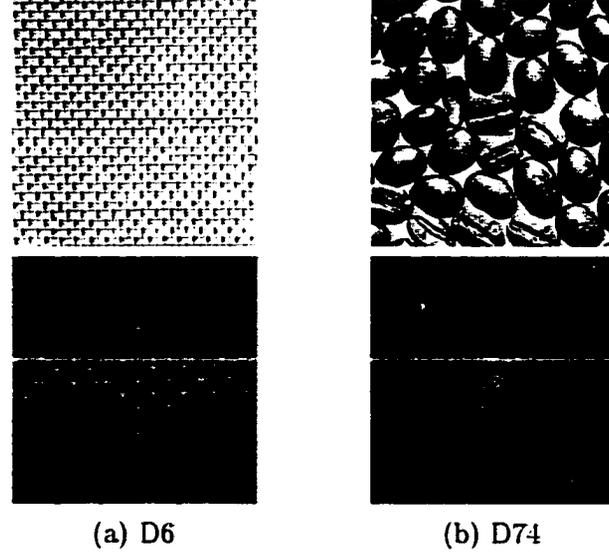


Figure 2.3 Examples of power spectrum of a relatively fine texture (a) and a coarse texture (b). The difference in coarseness of the original images (upper row) is apparent in their power spectrums (lower row).

origin. while in the case of image with a fine texture, the values of $|F|^2$ will be more spread out. This is illustrated in Fig. 2.3. Thus, the averages of $|F|^2$ taken over ring-shaped regions centered at the origin may be used as features. Such features are given by

$$\phi_r = \int_0^{2\pi} |F(r, \theta)|^2 d\theta. \quad (2.16)$$

where $F(r, \theta)$ is the $F(u, v)$ expressed in polar coordinates with $r = \sqrt{u^2 + v^2}$, and $\theta = \tan^{-1}(v/u)$.

Similarly, it is known that the angular distribution of values in $|F|^2$ represents the directionality of the texture in the image f . A texture with many texture primitives in a given direction θ would have high values of $|F|^2$ concentrated around the perpendicular direction $\theta + (\pi/2)$, while in a weakly directional texture, $|F|^2$ is small in the direction. Again, texture directionality can be measured by averages of $|F|^2$ taken over wedge-shaped regions centered at the origin,

$$\phi_\theta = \int_0^\infty |F(r, \theta)|^2 dr \quad (2.17)$$

for various values of θ , the wedge slope.

The standard form of texture features based on ring- and wedge-shaped samples of the discrete Fourier spectrum are of the form

$$\phi_{r_1, r_2} = \int_{r=r_1}^{r_2} |F(r, \theta)|^2 d\theta, \quad \phi_{\theta_1, \theta_2} = \int_{\theta=\theta_1}^{\theta_2} |F(r, \theta)|^2 dr, \quad (2.18)$$

respectively.

The two sets of features utilize the basic concepts of texture characteristics in the frequency domain and provide reasonable classification performance [82, 85].

2.3.2 Gabor Filters

Recently, multi-channel filtering methods have shown significant potential for texture classification and segmentation [2, 80, 40, 48, 32, 41]. The multi-channel filtering decomposes images into multiple oriented spatial-frequency (s - f) channels, maximizing the simultaneous localization of energy in both spatial and frequency domains. The approach to texture analysis is intuitively appealing because it allows us to exploit differences in dominant sizes and orientations of different textures. The notion of s - f analysis is also supported by experimental research on human and mammalian vision [1]. Gabor filters are one of s - f techniques and they measure the localized s - f information at specific scales and orientations.

A two-dimensional Gabor filter can be defined by [32]

$$h(x, y) = \exp \left\{ -\frac{1}{2} \left[\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right] \right\} \cos(2\pi u_0(x \cos \theta + y \sin \theta)), \quad (2.19)$$

where u_0 is called "center frequency", and is the frequency of a sinusoidal plane wave along the x -axis (i.e. the 0° orientation), and σ_x and σ_y are the space constants of the Gaussian envelope along the x and y coordinate system.

The frequency and orientation-selective properties of Gabor filter are more explicit in its frequency domain representation. The frequency domain representation of Eq. (2.19) is given by

$$H(u, v) = A \left(\exp \left\{ -\frac{1}{2} \left[\frac{(u - u_0)^2}{\sigma_u^2} + \frac{v^2}{\sigma_v^2} \right] \right\} + \exp \left\{ -\frac{1}{2} \left[\frac{(u + u_0)^2}{\sigma_u^2} + \frac{v^2}{\sigma_v^2} \right] \right\} \right), \quad (2.20)$$

where $\sigma_u = 1/2\pi\sigma_x$, $\sigma_v = 1/2\pi\sigma_y$, and $A = 2\pi\sigma_x\sigma_y$. Hence, by varying u_0, σ_x, σ_y , and θ parameter values, we can obtain a set of filters that are sensitive to frequency and orientations.

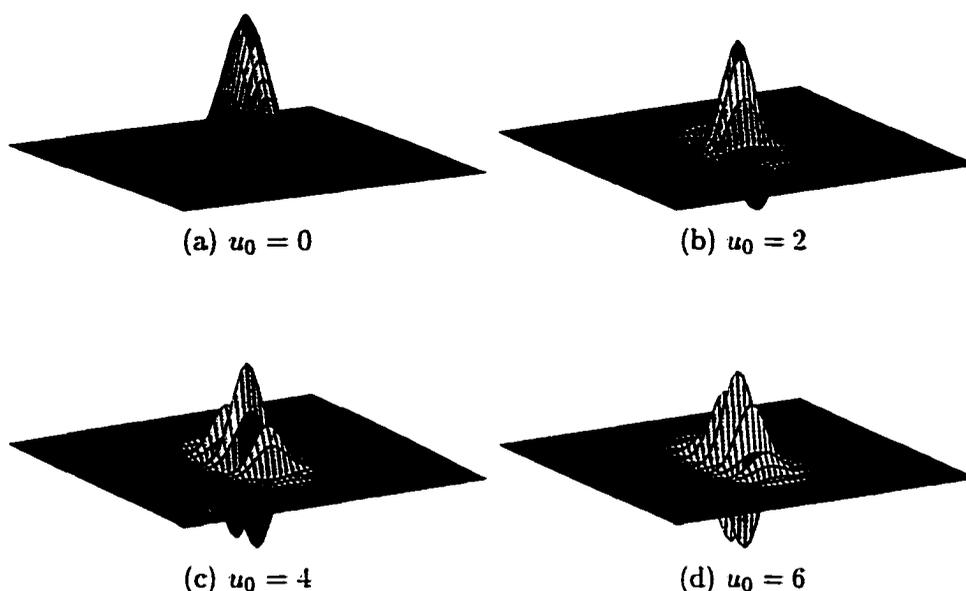


Figure 2.4 Variation of Gabor filters due to changes in center frequency u_0 . The other parameters are fixed at $\theta = 0^\circ$, $\alpha_x = 0.2$, $\alpha_y = 0.1$.

Variation in u_0 provides frequency-selective filters, while variation in θ provides orientation-selective filters. The impulse response of corresponding filters are shown in Fig. 2.4, and Fig. 2.5, respectively.

In multi-channel approaches, it is common to decompose a texture image using a set of Gabor filters and compute a feature set from the filtered output. Then, the features are classified using conventional pattern recognition techniques to determine the original texture. Hence, computing a “good” measure of the texture is crucial in texture classification. It is known that smaller bandwidths in the $s - f$ domain are desirable to obtain finer distinctions among different textures. However, the effective width of a filter in the spatial domain and its bandwidth in the frequency domain are inversely related due to the uncertainty principle [17]. Gabor filters achieve the theoretical lower bound of the uncertainty principle. They attain maximum joint resolution in $s-f$ and are bounded by the relations $\Delta x \cdot \Delta u \geq \frac{1}{4\pi}$ and $\Delta y \cdot \Delta v \geq \frac{1}{4\pi}$ where $(\Delta x, \Delta y)$ and $(\Delta u, \Delta v)$ are the resolution in space and frequency domains, respectively. These observations need to be taken account in the choice of filter parameters.

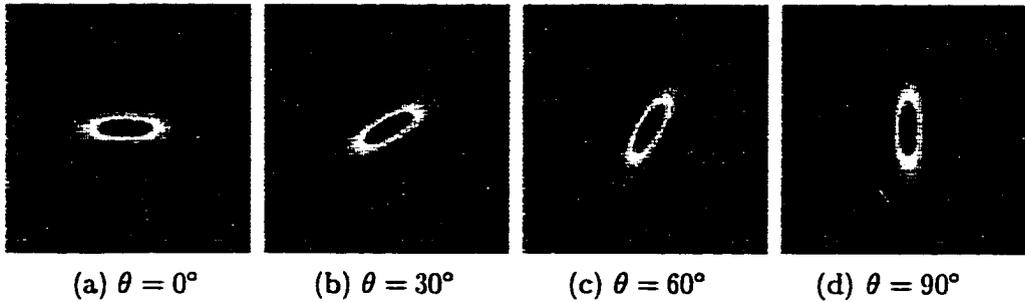


Figure 2.5 Variation of Gabor filters due to rotation angle θ . The other parameters are fixed at $u_0 = 2$, $\alpha_x = 0.2$, $\alpha_y = 0.2$.

Jain and Farrokhnia [32] present a procedure for choosing the filter parameters. They restrict the number of orientation parameters to four: 0° , 45° , 90° , and 145° for computational efficiency. It is also noted that the four θ s are sufficient for discriminating many textures. For u_0 , the following values are used:

$$1\sqrt{2}, 2\sqrt{2}, 4\sqrt{2}, \dots, (M/4)\sqrt{2}.$$

where M is a width of image and is a power of 2. These frequencies are 1 octave apart and they guarantee that the passband of the filter with the highest frequency $(M/4)\sqrt{2}$ falls inside the image array. Several experiments also show that simple cells in the visual cortex are also about 1 octave wide [65].

For an image of size 256×256 , a total of $4 \times \log_2(256/2) = 28$ can be used. This set of filters results in a nearly uniform coverage in the s - f domain. A decomposition obtained by the filter set is nearly orthogonal, since the amount of overlap between the filters is small. Even with the (nearly) uniform and orthogonal decomposition scheme, a selection of a subset of the filtered images is desirable since it is not necessarily true that all filter provides discriminatory information about the texture. A selection can be made based on the amount of energy in the filtered image [32].

It is found that the Gabor filter-based analysis is useful in texture classification [28, 67], segmentation [32, 66, 2, 20], and content-based image retrieval [56, 88] applications. However, there are drawbacks with using Gabor filters in practical applications. The selection of parameters for frequency, orientation and bandwidth for each filter is highly dependent on the

image. This process is nontrivial [2]. Otherwise, the complete Gabor expansion results in an impractical number of filters. In addition, the Gabor functions are not orthogonal and there is a trade-off between the redundancy and completeness in the design of the Gabor filter bank.

2.3.3 Wavelet Transform-based Methods

A more practical way to obtain the spatial-frequency analysis, without using Gabor functions, is through the use of a quadrature mirror filter (QMF) wavelet filterbank. The wavelet filterbank provides dyadic trade-offs in s - f resolution by spacing the filters at octave-band distances. Obtaining high joint resolution in both dimensions is critical in texture region extraction where the localization of the texture pattern is important. Additional information relating to this filterbank and related work of wavelet-based approaches on texture analysis is presented in greater detail in Chapter 3.

2.4 Texture Classification Performance Measure

Most papers describing texture classification algorithms usually give percentage classification accuracy results on sample textures. Although texture features are drawn from a widely used set of textures, such as the Brodatz textures, the selection of textures varies from paper to paper. It is also true that papers describing texture segmentation algorithms usually demonstrate the performance of the algorithm on arbitrarily selected texture images. This has the advantage that images can be chosen to ensure that the algorithm is presented in the best light.

The issue of bias has been addressed and comparative studies have been performed in [82, 12, 60, 61, 4, 68, 67]. For example, Weszka *et al.* [82] compare the accuracy of three different approaches (Fourier power spectrum, second-order gray level statistics, and first-order statistics of gray level difference) on terrain images. Connors and Harlow [12] conduct a theoretical analysis of the algorithm [82] which shows, for example, uniqueness of the features and the amount of information that could be captured by the texture features. The increasing sophistication of algorithms and lack of a formal definition of texture has made such a theoretic study

very difficult since then.

Ohanian and Dubes [60] provide additional comparative performance among four commonly used texture feature sets based on classification error. The feature sets include GLCM, GMRF parameters, Gabor filter energy parameters, and fractal parameters. Four classes of images were used in the study: fractal images, GMRF images, natural leather image, and painted surface images. Each class has four samples. The GLCM method has been found to be the best except for the fractal images. However, half of the image sets are artificially generated, and the number of images are too limited to draw a general conclusion.

Ojala *et al.* [61] provide another comparison of texture features using the same image set used in [60] and nine classes of Brodatz images. They use simple texture parameters based on the gray level difference method, Law's texture energy [44], center-symmetric covariance, and local binary patterns. The goal of their study is to find pairs of such features which provide complementary information about the texture so that uncorrelated texture information is obtained. The classification results obtained based on the complementary pairs are better than most of the results obtained with feature sets in [60].

Buf *et al.* [4] describe a benchmark test for comparing the accuracy of seven texture segmentation algorithms. Their benchmark measures the accuracy along the borders of segmentation results; it assumes that the classification of the texture in distinct regions is correct. However, their benchmark is based entirely on artificial textures. Additional comparative study on texture segmentation has been carried out by Reed and Buf [70].

Porter and Canagarajah [67] describes a rotation-invariant classification scheme for comparing three different feature sets on 16 Brodatz texture images. The feature sets include GMRF, Gabor filters, and wavelet filters. The wavelet-based approach is found to be the most accurate and the least sensitive to noise.

A comparative study for most major filtering approaches to texture segmentation is conducted by Randen and Husøy [68]. The methods include Law's mask, ring/wedge filters, dyadic Gabor filterbanks, wavelet transforms, wavelet packets and wavelet frames, quadrature mirror filters, discrete cosine transform, eigenfilters, optimized Gabor filters, linear predictors, and

optimized finite impulse response filters. It is found that various filtering approaches yield different results for different images. No single approach performs best or even very close to best for *all* images.

Most of the texture features described above lack a connection with the perceptual meaning of textures. It is therefore hard to describe textures using the feature parameters. In [50], Liu and Picard used the Wold transform to extract perceptual parameters for image retrieval applications. Basic texture properties such as regularity, fineness/coarseness, and direction were measured from the model. They claim that the Wold model offers a perceptually more satisfying pattern measure while exceeding the performance of other well-known pattern recognition methods including the TSWT [5] method.

Smith and Burns [76] propose a framework, MeasTex, for measuring the accuracy of texture classification algorithms. This framework provides a way of standardizing the practice for reporting algorithm results. MeasTex addresses the aforementioned deficiencies by: incorporating a large set of texture classification problems; being publicly available; having supporting software; being easily extendible.

There is a very important issue which is missing in all the papers describing comparative performance of feature sets: the conclusions made by the comparative studies may not be true. For example, it is not safe to draw such a statement: "one transform domain is superior than the other", unless the texture features obtained in a specific transform domain are the best features we could obtain from the domain. In addition, if a standard feature set is not available in each method as in the GLCM approach, the conclusion needs to be made very carefully. Thus, we should be concerned not only with the image transform method, but also with the method of feature extraction in each transform domain.

In this dissertation, the classification experiments employ the same data set used in recent literature to provide a partial comparison between the proposed approach and other successful approaches.

CHAPTER 3 WAVELETS AND TEXTURE ANALYSIS

Texture discrimination is one of the most difficult tasks among low-level computer vision problems [54]. Although textures are quickly preattentively discriminated by a human observer [52], appropriate models for textures do not exist. The perception of texture depends on local but not pointwise properties. However, a prescriptive procedure for selecting the neighborhood size over which textures can be analyzed is not available. This has motivated the use of a transform with which image properties are measured over domains of varying sizes. In parallel, psychophysics, and physiological experiments [16] have shown that multiscale/multiresolution transforms seem to appear in the visual cortex of mammals. Haar and Gabor decompositions, Gaussian and Laplacian pyramids are examples of multiscale decompositions. In the last decade, wavelet theory has emerged and now provides a solid and unified framework for multiscale image analysis, making it a preferred tool because of several conceptual and computational advantages [51].

This chapter focuses on that part of wavelet theory which is of interest to the computer vision community, in particular, for texture classification and segmentation. Variations of wavelet transforms used in the texture analysis literature are also described.

3.1 The Continuous Wavelet Transform

A wavelet is a function $\psi(t)$ whose Fourier spectrum $\hat{\psi}(f)$ satisfies the admissibility criterion give by

$$\int_{-\infty}^{+\infty} \frac{|\hat{\psi}(f)|^2}{|f|} df < \infty. \quad (3.1)$$

If $\psi(t)$ is sufficiently regular (which means that it decreases exponentially as t increases), this condition becomes

$$\hat{\psi}(0) = 0 \quad \Leftrightarrow \quad \int_{-\infty}^{+\infty} \psi(t) dt = 0 \quad (3.2)$$

This shows that a wavelet is a zero mean function and hence exhibits some oscillatory behavior.

A family of wavelet functions (a wavelet basis) is constructed from one single function (called the *mother wavelet*) by translation and scaling of the mother wavelet as follows:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad (3.3)$$

where a denotes a dilation parameter (which is inversely proportional with frequency) and b indicates a translation parameter. A classical example of such a family is the Daubechies wavelet family [15] of which three members of the family are represented in Fig. 3.1. The wavelet becomes narrower for smaller values of a and examines a small region of the signal to resolve time details more accurately. Likewise, for larger a , the wavelet is expanded and examines a larger region of the signal to resolve low frequency details more accurately. That is, the mother wavelet can be considered as a bandpass function centered around some center frequency and is translated in time to select the part of the signal to be analyzed. This feature of wavelet transform is very useful in the analysis of nonstationary waveforms.

The one-dimensional continuous wavelet transform (CWT) of a signal $f(t)$ is performed by projecting it onto the wavelet basis

$$Wf_a(b) = \int f(t) \psi_{a,b}^*(t) dt = \langle f, \psi_{a,b} \rangle, \quad a \neq 0. \quad (3.4)$$

where $\langle \cdot \rangle$ denotes an inner product operation.

3.2 Discrete Wavelet Transform

There is a high redundancy in continuous wavelet transform when the parameters (a, b) in Eq.(3.4) are continuous. Hence, the transforms are usually evaluated on a discrete grid in the time-frequency plane. This is performed by selecting integer powers of a fixed dilatation parameter $a_0 > 1$ for a , i.e., $a = a_0^m$, where m is an integer. The value of b is chosen such that high frequency wavelets (small ' a 's) are translated by small steps in order to cover the whole

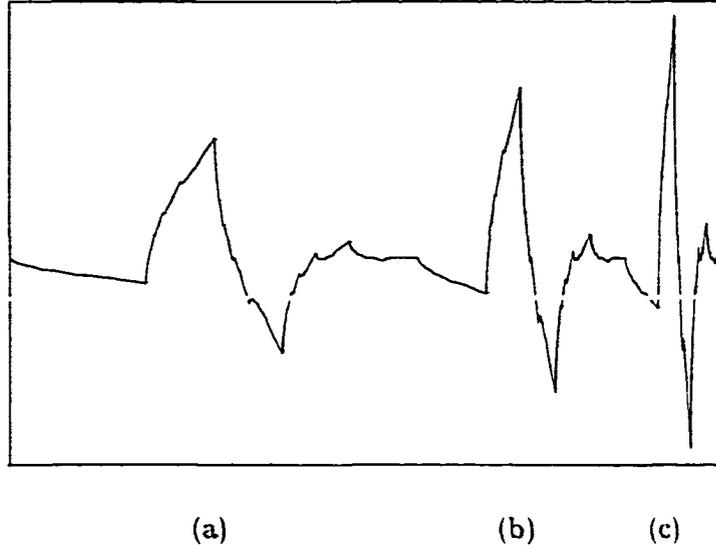


Figure 3.1 Examples of Daubechies wavelet family representing low to high frequency (from left to right). (a) $a = 2^2$. (b) $a = 2^1$. (c) $a = 2^0$.

time range, while low frequency wavelets (large 'a's) are translated by larger steps. Since the duration of $\psi(a_0^{-m}t)$ is proportional to a_0^m , b can be discretized as $b = nb_0a_0^m$, where b_0 is fixed and $n \in \mathcal{Z}$. The corresponding wavelets are then

$$\psi_{m,n}(t) = a_0^{-m/2} \psi(a_0^{-m}(t - nb_0a_0^m)) \quad (3.5)$$

$$= a_0^{-m/2} \psi(a_0^{-m}t - nb_0). \quad (3.6)$$

The parameters a_0, b_0 may be chosen when feasible so that the function $\psi_{m,n}(t)$ form an orthonormal basis in $L^2(\mathcal{R})$ such that

$$\langle \psi_{m,n}, \psi_{m',n'} \rangle = \delta_{mm'} \delta_{nn'} \quad (3.7)$$

where δ_{mn} is a delta function. The most commonly used values of a_0 and b_0 that form an orthonormal basis are 2 and 1, respectively, resulting in

$$\psi_{m,n}(t) = 2^{-\frac{m}{2}} \psi(2^{-m}t - n). \quad (3.8)$$

Here, the scales are on a dyadic grid. Then, a dyadic discrete wavelet transform of a signal $f(t)$ can be obtained, where the signal is split into octave frequency bands. This can be defined as

follows:

$$f(t) = \sum_{m,n} d_{m,n} \psi_{m,n}(t) \quad (3.9)$$

where

$$d_{m,n} = \langle f, \psi_{m,n} \rangle \equiv \int f(t) \psi_{m,n}(t) dt \quad (3.10)$$

are called the wavelet coefficients.

3.3 Multiresolution Decomposition Theory

It is shown that a signal can be evaluated in multiscales by selecting different values of a . If a *resolution* is defined as the number of basis functions used to represent the signal, the wavelet representation can be viewed as a multiresolution representation.

Consider a signal $f \in L^2(\mathcal{R})$. Let P_m a projection operator such that $P_m f$ is the approximation of f at the resolution m . It can be proven that P_m is an orthogonal, linear operator and that $P_m f$ is an element of the vector space V_m defined by [53]

$$V_m = \{P_m f | f \in L^2(\mathcal{R})\}. \quad (3.11)$$

A multiresolution analysis of f is then performed by projecting f on successive lower resolution (higher a) approximations. If all possible values of m is considered, a ladder of nested subspaces can be obtained

$$\dots \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \subset \dots \quad (3.12)$$

with $\bigcup_{m=-\infty}^{m=+\infty} V_m$ is dense in $L^2(\mathcal{R})$ and $\bigcap_{m=-\infty}^{m=+\infty} V_m = \{0\}$. To make this ladder of subspaces a multiresolution representation, additional requirements need to be met:

$$f(t) \in V_m \Rightarrow f(2t) \in V_{m-1}, \forall m \in \mathcal{Z}. \quad (3.13)$$

and

$$f(t) \in V_0 \Rightarrow f(t-n) \in V_0, \forall n \in \mathcal{Z}. \quad (3.14)$$

Finally, there exists a function $\phi(t) \in V_0$ such that the set $\{\phi_{0,n}(t) = \phi(t-n), n \in \mathcal{Z}\}$ form a basis of V_0 . This last equation, together with Eq. (3.13) leads to

$$\phi_{m,n}(t) = 2^{-\frac{m}{2}} \phi(2^{-m}t - n) \quad (3.15)$$

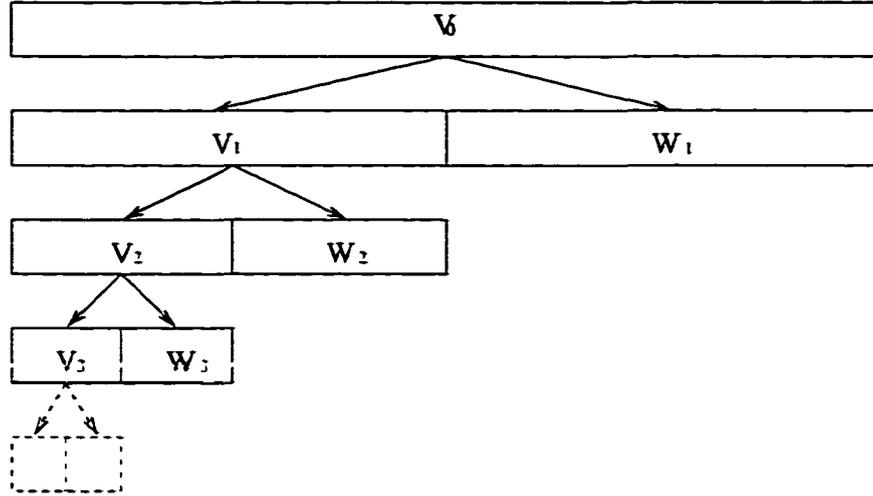


Figure 3.2 Schematic representation of the ladder of subspaces in wavelet transforms.

which is an orthonormal basis for V_m with $n \in \mathcal{Z}$. The function $\phi(t)$ is called *scaling* function.

The process of projecting f from V_{m-1} to V_m results in loss of signal information. The lost information between the transition is contained in the space W_m which is the orthogonal complement of space V_m . That is, $V_{m-1} = V_m \oplus W_m$, where \oplus stands for "orthogonal sum". The space W_m contains the difference information between $P_m f$ and $P_{m-1} f$ and is therefore called the *detail space* at scale m . The process is illustrated in Fig. 3.2.

It is remarkable that the set of functions that spans the complementary space of V_m turn out to be a wavelet basis [53] and can be defined from a mother wavelet $\psi(t)$ such that $\psi_{m,n}(t) = 2^{-\frac{m}{2}} \psi(2^{-m}t - n)$. In other words, the wavelet decomposition of a signal contains the information lost in the projection of a signal at two successive levels of resolution. Suppose, for example, we have a signal $f(t) \in V_m$. Since $V_m = V_{m+1} \oplus W_{m+1}$, it is possible to write

$$f(t) = \sum_n \langle f, \phi_{m,n} \rangle \phi_{m,n} \quad (3.16)$$

$$= \sum_n \langle f, \phi_{m+1,n} \rangle \phi_{m+1,n} + \sum_n \langle f, \psi_{m+1,n} \rangle \psi_{m+1,n}. \quad (3.17)$$

Then,

$$\sum_n \langle f, \psi_{m+1,n} \rangle \psi_{m+1,n} = \sum_n \langle f, \phi_{m,n} \rangle \phi_{m,n} - \sum_n \langle f, \phi_{m+1,n} \rangle \phi_{m+1,n}. \quad (3.18)$$

where $\langle f, \phi_{m,n} \rangle$ results in an approximation of a signal, $s_{m,n}$, and $\langle f, \psi_{m,n} \rangle$ results in a wavelet coefficients, $d_{m,n}$. This clearly shows that the wavelet decomposition of $f(t)$ is the difference between the information contained by projecting two successive levels of resolution.

In order to generate a multiresolution representation of a signal $f(t)$ (let's assume $f(t) \in V_0$), the decomposition is applied recursively to the approximation coefficients $s_{m,n}$, resulting in $s_{m+1,n}$ and $d_{m+1,n}$. If the decomposition is performed up to scale level $m = I$, $f(t)$ can be decomposed as follows:

$$f(t) = \sum_n s_{I,n} \phi_{I,n} + \sum_{i=1}^I \sum_n d_{i,n} \psi_{i,n}, \quad (3.19)$$

and the space V_0 can be represented by

$$V_0 = V_I \oplus \bigoplus_{i=1}^I W_i. \quad (3.20)$$

3.4 Quadrature Mirror Filter (QMF)

Mallat [53] introduces a wavelet transform called the "pyramid wavelet transform" with which computing the orthonormal wavelet basis is not needed in practice. One possibility to construct an orthonormal wavelet basis is to use the scaling function as a starting point. In a multiresolution ladder of spaces V_m , with $\phi_0 \in V_0 \subset V_{-1}$, it follows that

$$\phi(t) = \sqrt{2} \sum_k h_k \phi(2t - k) \quad (3.21)$$

with $h_n = \langle \phi_{0,0}, \phi_{-1,n} \rangle$. Then the orthonormal wavelet basis generated from a mother wavelet ψ_0 is orthogonal to that generated from ϕ_0 in V_{-1} , i.e., $\psi_0 \in W_0 \subset V_{-1}$. Then, there exists a sequence g_n such that

$$\psi(t) = \sqrt{2} \sum_k g_k \psi(2t - k), \quad (3.22)$$

and satisfies

$$g_n = (-1)^n h_{-n+1}. \quad (3.23)$$

Using the discrete Fourier transform, the relation of the two filters can be written as

$$\tilde{g}(\omega) \equiv \sum_k g_k e^{-jk\omega} = \sum_k (-1)^k h_{-k+1} e^{-jk\omega} = \tilde{h}(\pi - \omega). \quad (3.24)$$

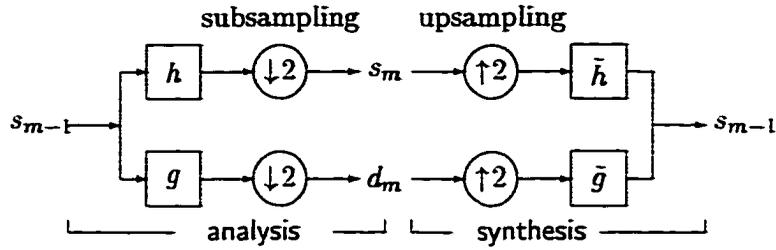


Figure 3.3 A structure of the discrete wavelet transform using Quadrature Mirror Filters.

Therefore $\hat{h}(\omega)$ is a “mirror” of $\hat{g}(\omega)$ with respect to the “half-band” value $\omega = \pi/2$ and consequently g_n and h_n are “quadrature mirror filters” (QMF).

Equations 3.21 and 3.22 can be generalized for successive spaces. Therefore, it is possible to write

$$s_{m,n} = \sum_k h_{k-2n} s_{m-1,k} \quad (3.25)$$

$$d_{m,n} = \sum_k g_{k-2n} d_{m-1,k}. \quad (3.26)$$

This QMF structure is illustrated in Fig. 3.3. The synthesis of the upper scale signal is implemented by upsampling and filtering operations with the filter coefficients: $\tilde{h}_n = h_{-n+1}$, $\tilde{g}_n = g_{-n+1}$. Downsampling is performed by taking every two coefficients, and upsampling is performed by adding zeros between any two consecutive coefficients. This is a very efficient method for computing the discrete wavelet transform (DWT).

Note that the h_n and g_n are associated with a lowpass and a highpass filter, respectively, and they satisfy the following conditions:

$$\tilde{H}(\omega) = H(-\omega), \quad \tilde{G}(\omega) = G(-\omega). \quad (3.27)$$

where $H(\omega)$ and $G(\omega)$ are Fourier transforms of h and g , respectively, and $\tilde{H}(\omega)$ and $\tilde{G}(\omega)$ are Fourier transforms of synthesis filters \tilde{h} and \tilde{g} that are chosen to obtain a perfect reconstruction.

3.5 Extension to Two-Dimensional Wavelet Transforms

Having defined the one dimensional scaling function and wavelet function, a two-dimensional separable scaling function can be obtained as:

$$\phi(x, y) = \phi(x)\phi(y) \quad (3.28)$$

whose dilations and translations form a basis for a subspace V_m . In addition, there exist three mother wavelets generating orthonormal bases for three orthogonal spaces: W_1 , W_2 , and W_3 , respectively. These wavelets are also separable, and are given by

$$\begin{aligned} \psi_1(x, y) &= \phi(x)\psi(y) \\ \psi_2(x, y) &= \psi(x)\phi(y) \\ \psi_3(x, y) &= \psi(x)\psi(y). \end{aligned} \quad (3.29)$$

A commonly used method for 2-D DWT implementation involves a successive application of 1-D QMF filters h_n and g_n along the columns and rows of the image. The subimages resulting from such operations to an image A are three detail images A_{HL} , A_{LH} , and A_{HH} and an approximation image (A_{LL}), where the subscript denotes the frequency information of the rows and the columns, respectively. For example, A_{LH} emphasizes the “low” frequency information along the rows and “high” frequency information along the columns, retaining details of horizontal information. Likewise, the subimage A_{HL} emphasizes vertical information, and A_{HH} emphasizes diagonal information. This process is repeated iteratively on the approximation image A_{LL} , with a translated and dilated version of basis functions for each consecutive scale.

An example of 2-D DWT for a commonly used “Lena” image is shown in Fig. 3.4. Horizontal, vertical, and diagonal details are clearly seen in corresponding subimages. Subimage labeling according to the characteristics of contained information is also shown in Fig. 3.4(b).

The frequency partition for two-level decomposition using 2-D wavelet transform is illustrated in Fig. 3.5. The decomposition of a signal at coarser resolution produces wavelet filters with better localization for low frequency. The number of subimages N becomes $3I + 1$, where I is the scale depth. The combination of a series of filters to produce each subimage is called

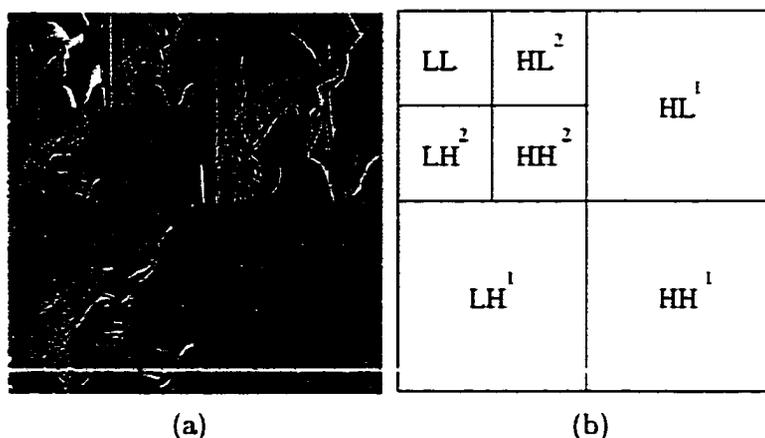


Figure 3.4 An example of a two level DWT of the well-known ‘Lena’ image (a). The letters in (b) indicate the corresponding frequency components of the detail images.

a *channel* of a filterbank. That is, we have $3I + 1$ channels in the case of I -level wavelet decomposition.

3.6 Stationary Wavelet Transform

It is known that a simple integer shift of the input signal results in a nontrivial modification of the discrete wavelet decomposition [53]. This property represents a problem for the texture feature extraction algorithm since “texture” is considered to be shift-invariant (i.e. stationary). A natural way to overcome this limitation is to use a redundant version of DWT which is called the stationary wavelet transform (SWT) in which the down-sampling operation is not performed [80]. The multi-resolutional characteristic of this redundant decomposition is achieved by upsampling 1-D DWT filters before further decomposition. The block diagram of 1-D SWT decomposition of a function $f(t)$ is shown in Fig. 3.6 (a). $H(z)$ and $G(z)$ in the figure represent the Z-domain filter responses of h_n and g_n , respectively.

The extension to 2-D SWT can be also made with the method described in the previous section. The corresponding schematic is shown in Fig. 3.6(b) which uses a notation for denoting channel outputs that will be used in Chapter 7. The diagram illustrates how the 1-D filters are applied to obtain 2-D wavelet representations. The subscript (x or y) at the filter response

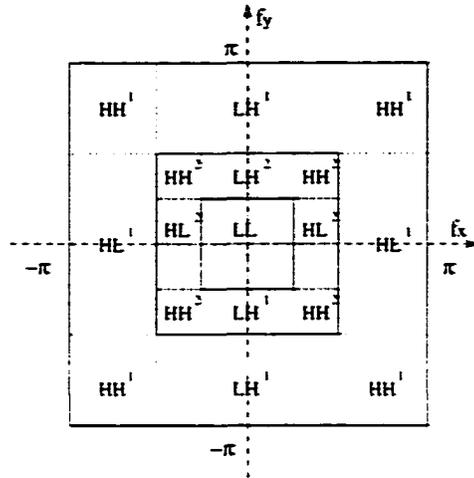


Figure 3.5 Partitions of frequency domain resulting from a 2-level decomposition using DWT. The letters correspond to the ones in Fig. 3.4(b)

denotes that the filter is applied in the direction of the subscript. Each channel is characterized by a successive application of filters in x and y both directions. For example, channel 5 can be characterized by the application of the filter

$$H_y(z^2)G_x(z^2)H_y(z)H_x(z).$$

This redundant wavelet transform has found its usefulness in texture classification and segmentation applications. The classification accuracy obtained by using SWT is found to be superior than the one obtained by using DWT [80]. In addition, SWT has been successfully applied to address texture segmentation problems in [41].

3.7 Wavelet Packets

Wavelet packets are the set of functions generated by splitting each of ϕ and ψ into two new orthonormal sets of basis functions as described in section 3.3. The corresponding schematic representation of the subspace splitting is illustrated in Fig. 3.7. Unlike a pyramid wavelet transform, the packet decomposition decomposes the low- (approximation) and high-frequency (detail) components, constructing a tree-structures multiband extension of the wavelet transform. For example, the number of subimages of three level packet decomposition of an image

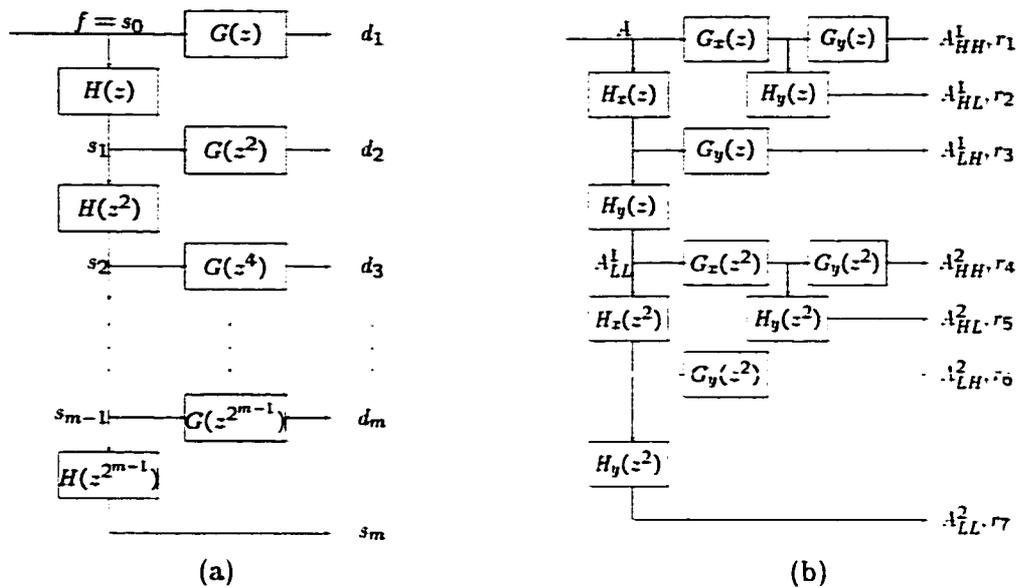


Figure 3.6 A block diagram of stationary wavelet transforms using Quadrature Mirror Filters. (a) 1-D SWT; (b) 2-D SWT. $I = 2$.

is 64. Thus, wavelet packets are well localized in both time and frequency and thus provide an attractive alternative to pure Fourier analysis.

In [40], Laine and Fan show that the use of the wavelet packet transform provides perfect classification results for twenty five textures (size 128×128). The texture features used are energy and entropy measures of the decomposition output. However, most natural texture images do not have energy in the entire frequency domain as shown in Fig. 2.3, and a number of subimages may not be related to the global texture at all. Also, the features tend to be sensitive to texture translation and rotation when the number of transformation levels increases, making the frequency partitions finer. In addition, this approach requires a large texture sample size in order to obtain reasonable classification accuracy. It is worth noting that the energy features described in [40] can be approximated by computing the mean square sum from the uniformly divided cells in the 2-D frequency domain.

Additional examples of the use of wavelet packet transform in texture analysis can be found [21, 46].

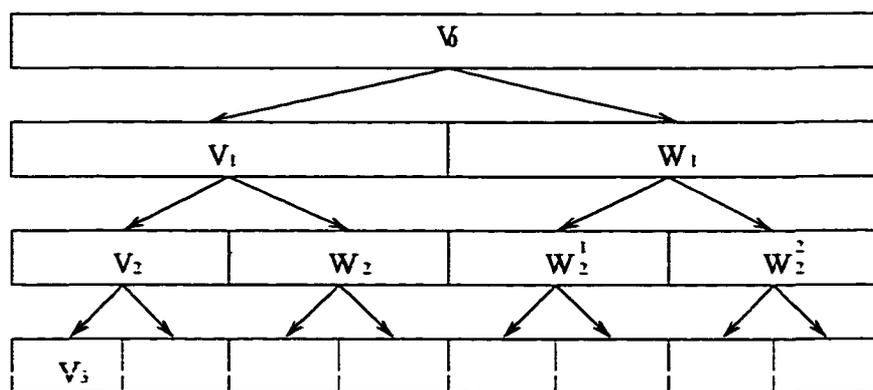


Figure 3.7 Schematic representation of the subspaces in the wavelet packet decomposition.

3.8 Tree-Structured Wavelet Transform (TSWT)

As discussed in Section 3.3, the pyramid wavelet transform recursively decomposes a signal into a set of frequency channels that have narrower bandwidths in the lower frequency region. However, this decomposition may not be useful in texture analysis, since the most significant information of a texture usually appears in the middle frequency channels. Therefore a sophisticated approach to perform the wavelet transform for textures is to detect the dominant frequency channels and then to explore those dominant channels further.

Chang and Kuo [5] introduce a tree-structured wavelet transform (TSWT) to accomplish the frequency selective wavelet transform. Such a decomposition scheme is obtained using an energy measure of subimages as a criterion in decision making for further decomposition. A schematic of a possible four-level TSWT decomposition is shown in Fig. 3.8. This avoids the full decomposition that occurs in wavelet packets, resulting in a computationally efficient algorithm.

More than 95 % accuracy in classifying Brodatz [3] textures is achieved using this adaptive multichannel method [5]. Features employed for texture classification include the locations of the first five dominant channels exhibiting large energy values after a 4-level decomposition. In order to keep the smallest subimage size reasonable, the texture sample needs to be large enough (the authors recommend 128×128). The approximation of the features, however, can

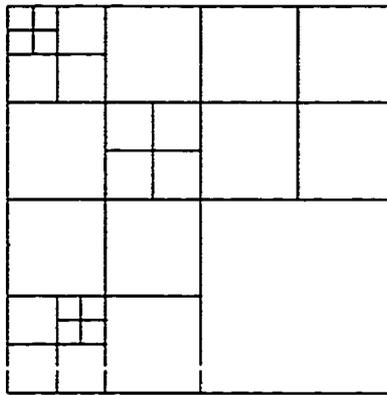


Figure 3.8 Example of a possible four-level TSWT decomposition. Subimages with high energy are further decomposed.

be obtained from 2-D Fourier transforms by dividing the frequency domain with $[\pi/16, \pi/16]$ size cells and identifying the locations of the first five highest energy values. In addition, the features are substantially similar to the locations of dominant frequency channels obtained from the 4-level wavelet packet decomposition.

3.9 Summary

A brief description on wavelet theory is given in this chapter. Definitions of the wavelet transforms both in continuous and discrete time, multi-resolution theory, fast implementation of DWT using QMFs, and its extension to 2-D case are described. In addition, the use of modified DWT that overcomes some of drawbacks associated with DWT is described for texture applications. The modified wavelet transforms include redundant wavelet transform, wavelet packet transform, and tree structured wavelet transform.

CHAPTER 4 ORIENTED TWO-DIMENSIONAL WAVELET FILTERS

In the previous chapter, we described a separable 2-D QMF scheme for texture image decomposition. A drawback of this approach is that the orientation decomposition is limited. The diagonal information captured by a standard wavelet transform contains two diagonal details (45° and 135°) simultaneously. In many texture analysis problems, however, additional rotational information is needed. For example, results independent of orientation are demanded for rotation-invariant classification, where detailed directional information is used to generate rotation-invariant features.

A fine orientation selectivity can be obtained by using Gabor wavelet filters or non-separable oriented wavelet transforms, such as the hexagonal wavelet transform [75] and the steerable wavelet filters [24]. The hexagonal wavelet representation partitions orientations into three bands of 60 degrees equally covering the frequency domain using a hexagonal sampling system and filters at the expense of implementation difficulty. This scheme has been applied to mammographic images for feature enhancement [42]; its application to texture analysis, however has not been actively reported. Steerable filters are overcomplete, and the filter can be steered along any direction using a linear combination of a set of basis functions.

In this chapter, we introduce a new set of oriented and non-separable wavelet filters to obtain better orientation tuning for texture image analysis. In addition, a dual wavelet filterbank is built by combining the standard and oriented wavelet filter sets.

4.1 2-D Wavelet transforms using 2-D Filters

As shown in Section 3.5, the QMFs are typically applied to images in a separable manner. The 1-D QMFs are successively applied in x and y directions, resulting in four subimages

with a mixture of the two orientations. The equivalent 2-D QMFs can be obtained from the products of 1-D wavelet and scaling functions and they are defined as

$$\begin{aligned}
 H_0^s(k, l) &= h(k)h(l) \\
 H_1^s(k, l) &= h(k)g(l) \\
 H_2^s(k, l) &= g(k)h(l) \\
 H_3^s(k, l) &= g(k)g(l)
 \end{aligned} \tag{4.1}$$

where the superscript 's' denotes that it is a standard wavelet filter (SWF). An example of 2-D Haar filters is given by

$$H_0^s = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, H_1^s = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}, H_2^s = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}, H_3^s = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}. \tag{4.2}$$

The application of these filters results in three detail images and an approximation image. Successive decomposition of the approximation image at each scale level results in a pyramidal tree structure with finer resolutions obtained toward the lower frequency bands.

Using a 2-D filtering approach, the channel response can be represented by combining the corresponding 2-D filters. We use $P_k(z_1, z_2)$ to represent the 2-D impulse response of channel k in the Z domain. Hence, the approximate version (channel 0) of an image $I(x, y)$ from a three-level decomposition is obtained by

$$\mathbf{I}(z_1, z_2)P_0(z_1, z_2) = \mathbf{I}(z_1, z_2)\mathbf{H}_H(z_1, z_2)\mathbf{H}_H(z_1^2, z_2^2)\mathbf{H}_H(z_1^4, z_2^4) \tag{4.3}$$

in the Z domain. Although this approach involves more operations than the 1-D approach, it is computationally efficient when the number of channels to be used is limited, since the channel response $P_k(z_1, z_2)$ can be pre-generated and unnecessary successive filtering operations can be avoided. This is particularly useful in packet type transforms where there are many channels to decompose and only a few dominant channels are of interest.

4.2 Rotated Wavelet Filters

We introduce a new method to achieve the spatial-frequency localization of an image in the diagonal directions by rotating 2-D wavelet filters counter-clockwise by 45 degrees so that

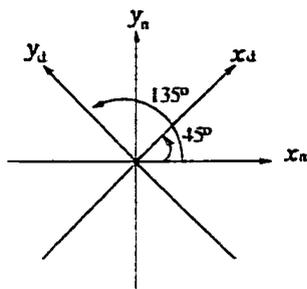


Figure 4.1 Angle reference for the 2-D wavelet transforms used in this study.

the decomposition is performed along the directions x_d and y_d instead of x_n and y_n directions as illustrated in Fig. 4.1. We call the rotated version of standard wavelet filter as a rotated wavelet filter (RWF). The new filter set is oriented and suitable for texture characterization, especially for textures with diagonal information. An example of RWF set for SWF with 4 coefficients in h_n and g_n is presented in Fig. 4.2.

The superscript 'r' represents "rotated", compared to 's' for "standard". The size of the filter is $(2L - 1) \times (2L - 1)$, where L is the length of the 1-D filter. Although the size of the 2-D filter becomes large, the computational complexity associated with the RWF decomposition is the same as that of SWF decomposition, if both are implemented in the 2-D frequency domain. A RWF set that is obtained by rotating 2-D Haar SWF set shown in Eq. (4.2) can be given by

$$H_0^r = \frac{1}{2} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, H_1^r = \frac{1}{2} \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}, H_2^r = \frac{1}{2} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & 0 \end{bmatrix}, H_3^r = \frac{1}{2} \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}. \quad (4.4)$$

The subsets spanned by those four filters are also rotated by 45° . The frequency partitioning of the RWFs are shown in Fig. 4.3(b). A corresponding partitioning for SWF is also shown for comparison. Each covers a quarter of whole frequency region and the set of RWFs retains the orthogonality property, since the following condition holds:

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} S_i(\omega) \overline{S_j(\omega)} d\omega = 0, \quad (i \neq j), \quad (4.5)$$

		a		
	b		c	
d		e		f
g	h		i	j
	k	l		m
		n		o
		p		

(a)

parameter	H_0^r	H_1^r	H_2^r	H_3^r
a	$h_4 h_1$	$g_4 h_1$	$h_4 g_1$	$g_4 g_1$
b	$h_3 h_1$	$g_3 h_1$	$h_3 g_1$	$g_3 g_1$
c	$h_4 h_2$	$g_4 h_2$	$h_4 g_2$	$g_4 g_2$
d	$h_2 h_1$	$g_2 h_1$	$h_2 g_1$	$g_2 g_1$
e	$h_3 h_2$	$g_3 h_2$	$h_3 g_2$	$g_3 g_2$
f	$h_4 h_3$	$g_4 h_3$	$h_4 g_3$	$g_4 g_3$
g	h_1^2	$g_1 h_1$	$h_1 g_1$	g_1^2
h	h_2^2	$g_2 h_2$	$h_2 g_2$	g_2^2
i	h_3^2	$g_3 h_3$	$h_3 g_3$	g_3^2
j	h_4^2	$g_4 h_4$	$h_4 g_4$	g_4^2
k	$h_1 h_2$	$g_1 h_2$	$h_1 g_2$	$g_1 g_2$
l	$h_2 h_3$	$g_2 h_3$	$h_2 g_3$	$g_2 g_3$
m	$h_3 h_4$	$g_3 h_4$	$h_3 g_4$	$g_3 g_4$
n	$h_1 h_3$	$g_1 h_3$	$h_1 g_3$	$g_1 g_3$
o	$h_2 h_4$	$g_2 h_4$	$h_2 g_4$	$g_2 g_4$
p	$h_1 h_4$	$g_1 h_4$	$h_1 g_4$	$g_1 g_4$

(b)

Figure 4.2 Design of rotated wavelet filters (filter length $L = 4$). (a) Rotated 2-D filter. Empty cells are zeros; (b) The filter coefficient values for four 2-D RWFs.

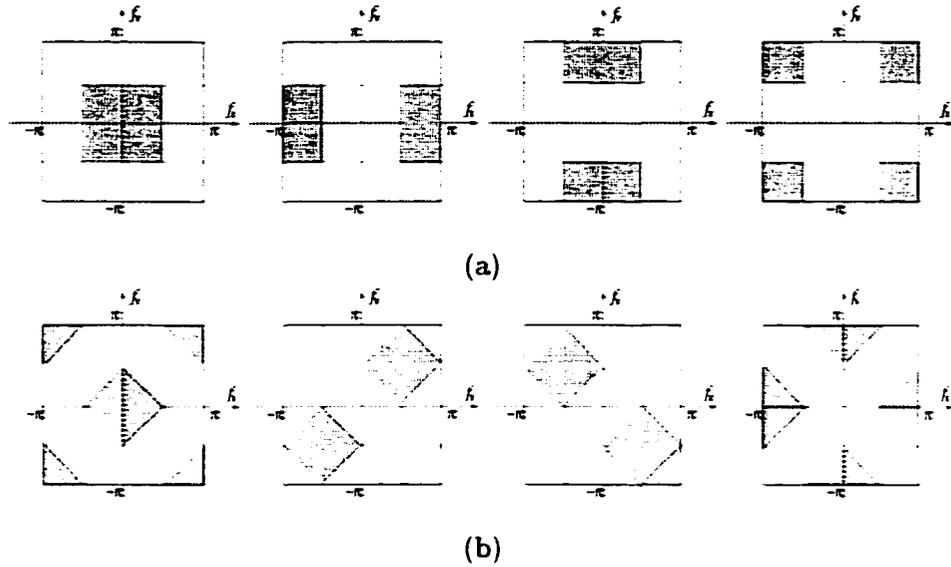


Figure 4.3 Frequency partitioning for the stationary wavelet filters (a) and rotated filters (b).

where $S_i(\omega)$ is the Fourier transform of the 2-D filter $H_i^r(k, l)$.

The one-level RWF decomposition of a natural texture image "D34" from Brodatz texture album is shown in Fig. 4.4(c). Two different diagonal characteristics of the texture are clearly seen at the outputs of filters H_1^r and H_2^r . For comparison, outputs of SWF set are also presented. This characteristics of RWF set provides important complementary information to the SWF set in feature extraction, and provides additional directional information.

The RWF set also keeps the self-inverting property of the standard wavelet filter set, meaning that the filter coefficients that are used for decomposition can be used for reconstruction. The 2-D synthesis filters can be generated by replacing the filter coefficients h_n and g_n in Fig. 4.2 with \tilde{h}_n and \tilde{g}_n :

$$\tilde{h}_n = h_{L-n+1}, \tilde{g}_n = g_{L-n+1}, n = 1, 2, \dots, L. \quad (4.6)$$

Then, the original 2-D signal can be reconstructed by convolving the subband images with the synthesis filters.

Thus, the 2-D RWF set is not separable anymore and provides better freedom of orientation tuning for image analysis, which in turn, provides a useful tool for extracting a directionality

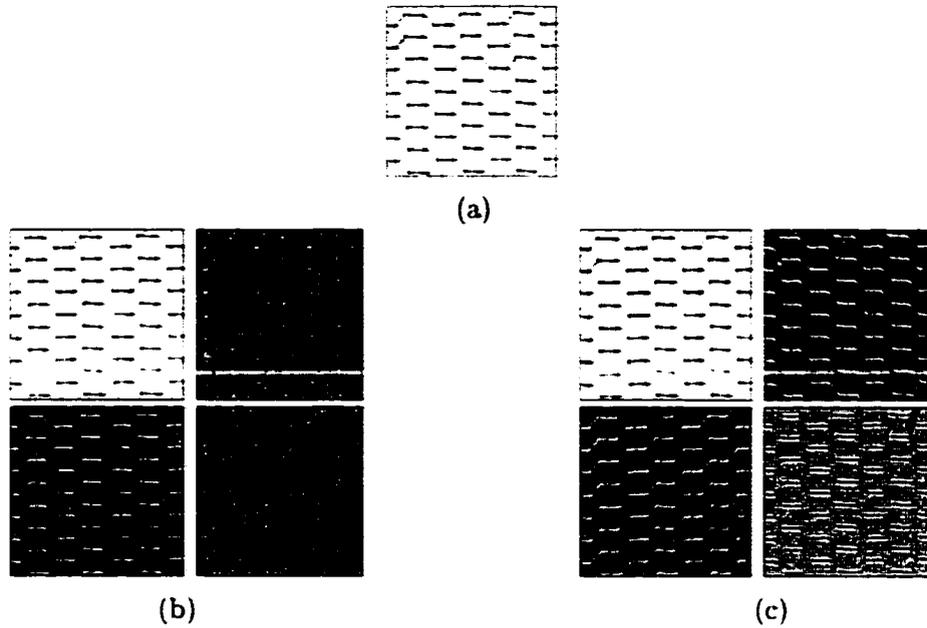


Figure 4.4 Demonstration of the one-level decomposition using SWF and RWF for a natural texture image "D34". (a) Original image: (b) Outputs of SWF: (c) Outputs of RWF. For (b) and (c), from the top left clockwise, filtered image using H_0 , H_1 , H_2 , and H_3

measurement of textures in conjunction with the SWF set. This will be demonstrated in Chapters 5 and 6.

4.3 Dual Wavelet Filter Bank

The cascaded structure of the wavelet decomposition forms a filterbank. We define a dual wavelet filterbank (DWFB) by combining a standard wavelet filterbank (SWFB) and a rotated wavelet filterbank (RWFB). To avoid a shift-variant characteristic of DWT, we employ a redundant version of the DWT. A diagram of the DWFB is presented in Fig. 4.5. The upper bank represents a SWFB and the lower represents a RWFB. Each bank has ten channel outputs (y_k). The output of H_0^s is further decomposed generating eight subband images at the next scale level. The intermediate channel outputs of H_0^r filter are ignored, to have the same number of channels in each filter bank.

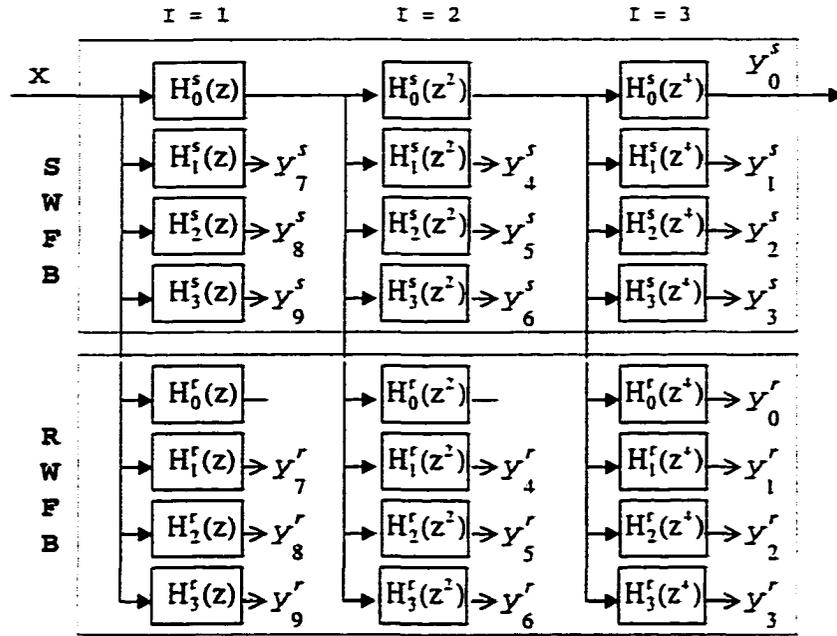


Figure 4.5 A diagram of the dual wavelet filterbank with scale level three. The upper bank represents a SWFB and the lower represents a RWF. The output of H_0^s filter is further decomposed into eight subbands at the next level.

4.4 Summary

A set of rotated wavelet filters (RWF) are proposed to obtain better representation of directional patterns in textures. They are obtained by rotating a set of 2-D standard wavelet filters (SWF) counter-clockwise by 45° . The new 2-D RWF set is not separable. It provides better freedom of orientation tuning for texture analysis, while keeping the orthogonality and self-inverting properties of SWF. However, the subsampling process that appears in DWT causes a partial data loss in this case. Consequently original signal cannot be completely reconstructed using the RWF set. Hence, this may not be a good candidate for a compression applications.

A dual wavelet filterbank (DWFB) is constructed by combining the two filter sets, and the use of this in texture classification applications is described in the following chapters.

CHAPTER 5 WAVELET-BASED TEXTURE CLASSIFICATION

In this chapter, an application of wavelet transforms to texture classification is addressed. We summarize the most successful wavelet-based features that have appeared in literature to date. We also describe the use of the dual wavelet filterbank to address texture classification problems in order to measure the effectiveness of the additional wavelet filterbank. The classification performance has been measured using two large natural texture databases.

5.1 Wavelet-based Texture Features

The most successful texture features that have appeared in texture classification literature to date can be summarized as follows:

- **Energy** The energy distribution across the scale is known to be an important characteristic for texture analysis [51]. The energy of the channel output $y_k(m, n)$ is computed as follows:

$$\varepsilon_k = \frac{1}{MN} \sum_m^{M-1} \sum_n^{N-1} |y_k(m, n)|^2. \quad (5.1)$$

where M and N are the width and length of the image $y_k(m, n)$. It is viewed as an average l_2 -norm. The energy can be also obtained from $\frac{1}{MN} \|\mathbf{Y}_k(\omega_1, \omega_2)\|^2$ using Parseval's theorem, where $\mathbf{Y}_k(\omega_1, \omega_2)$ is the Fourier transform of $y_k(m, n)$. Due to the orthogonality property of wavelet space generated by a set of wavelet filters, the wavelet decomposition preserves the energy through the decomposition scale. That is,

$$E\{\|x(m, n)\|^2\} = \sum_k E\{\|y_k(m, n)\|^2\}, \quad (5.2)$$

where E represents the expectation operator.

The energy is also considered as a variance of the channel output for $k = 1, 2, \dots, K - 1$, since $\sum_k \sum_l H_i(k, l) = 0$ for $i \neq 0$. This results in

$$E\{y_k(m, n)\} = E\{(\mathbf{P}_k * x(m, n))\} = 0. \quad (5.3)$$

where \mathbf{P}_k is the combination of the 2-D filters H_i s that characterize the k th channel.

More examples that use this energy feature for classification problem are found in [40, 8, 5, 68, 83, 46]. Chang and Kuo [5] use an l_1 -norm based energy measure which is defined as

$$E\{\|x(m, n)\|\} = \sum_k E\{\|y_k(m, n)\|\}. \quad (5.4)$$

Wilson and Bayoumi [83] propose a method for the efficient classification of compressed texture images by calculating wavelet energy features directly from a compressed embedded zerotree wavelet [73] symbol stream.

- **Energy ratio** Energy ratio features are used in [37] to characterize continuously varying texture patterns in ultrasound images. The ratio between subimages can be obtained by dividing the energy of a subimage by one obtained from another subimage. For example, $\varepsilon_{hl}^2/\varepsilon_{hl}^1$ is a ratio between two vertical detail subimages in two consecutive scales, 1 and 2. It is found that the features are grayscale transform-invariant and they outperform the GLCM-based features and Fourier features [37].
- **Entropy** This measure, previously proposed in [29] for texture analysis, has been used in [10] to identify a “best basis” for building wavelet packet libraries for signal compression. Entropy measure on the channel output defined as

$$H = -\frac{1}{MN} \sum_m^{M-1} \sum_n^{N-1} |y_k(m, n)|^2 \log |y_k(m, n)|^2. \quad (5.5)$$

has been used as a texture feature in [40, 8].

- **First and second-order statistics** Wouwer *et al.* [84] use the wavelet histogram feature and wavelet co-occurrence features for a classification of thirty VisTex [81] texture images. Their primary goal is to capture the first-order and second-order statistics

of wavelet coefficients as key descriptions for texture discrimination. The first-order statistics are captured by modeling a histogram of wavelet coefficients using

$$h(u) = Ke^{-(|u|/\alpha)^\beta}, \quad (5.6)$$

where β is inversely proportional to the decreasing rate of the peak, while α models the width of the histogram peak. K is a normalization constant to ensure that $\int h(u)du = 1$. The texture signatures used are α and β .

Eight features are computed based on the co-occurrence matrix [29] to capture the second-order statistics of wavelet coefficients. The authors claim that the histogram features outperform the energy features, and the second-order statistics features outperform the histogram features in their texture classification experiment. Some textures seem to be best characterized using first-order statistics, while for others second-order statistics are better. However, the best results are obtained by combining both feature sets.

A disadvantage associated with wavelet approaches is that they require larger size texture samples (128×128) in order to obtain reasonable classification accuracy. The size has to be sufficiently large to ensure that the size of the smallest subimage contains enough texture information. The use of a stationary wavelet transform (SWT) overcomes this drawback. Unser [80] characterizes textures using a set of channel energies estimated at the outputs of the stationary wavelet filterbank. Satisfactory classification accuracy was obtained for twelve Brodatz texture images using a texture sample size of only 32×32 . It is also shown that the SWT outperformed the DWT approach in classifying the twelve texture types [80].

Since one of the main objectives in this study is to demonstrate the effectiveness of the use of RWFB in addition to SWFB in texture analysis, we use the energy measure of the channel output to represent the texture. For the wavelet decomposition, Haar wavelet filters are used because of its computational simplicity. The Haar wavelet is found to provide a fairly good classification performance compared to higher order spline wavelets [80]. Moreover, general procedures for selecting wavelet basis functions have not been identified. The choice of the basis function mainly depends on the application.

5.2 Experimental Implementation

Several issues are involved in selecting the texture database to be used in the experiment. The most fundamental issue is whether to use realworld textures or synthetic textures. Real world textures have one unarguable advantage: ultimately, vision systems must operate in a realworld environment, and should be developed on data with real world properties. On the other hand, synthetic textures offer some advantages: an abundant supply of estimation and validation data can be generated, and the degree of difficulty of the texture problems can be precisely controlled. Synthetic textures also have a significant disadvantage: the algorithm for generating the texture may introduce artifacts which favor particular algorithms in ways which will not be reflected in realworld textures. We favor the use of realworld textures.

Two different classification experiments are performed using different data sets selected from the Brodatz's album [3] that has been used widely by many researchers. The images are scanned from an album of 112 textures photographed by Phil Brodatz. The images are diverse, including grass, pebbles, paper, cloth weave, woods and clouds. These images have been a standard for texture classification and segmentation problems.

The discriminating power of two filter sets (DWFB and RWFB) for various texture images is measured in experiment I, whereas the performance for visually similar textures is measured in experiment II. Another experiment (Experiment III) is conducted on the database used in [84] to evaluate a relative performance of our feature set to the first- and second-order statistics features [84]. It is our intention to provide a comparative study as much as possible to reduce ambiguity in selecting a suitable texture analysis method.

5.2.1 Experiment I and II

The texture data used for experiment I and II are shown in Fig. 5.1 and Fig. 5.2, respectively. The images are selected on a visual basis with each texture maintaining the essential substructure within its 32×32 size. For experiment I, twenty eight textures are selected from Brodatz's album in a way that the data set includes a variety of texture patterns. Among the twenty eight patterns, five textures that have similar texture patterns are selected for experi-

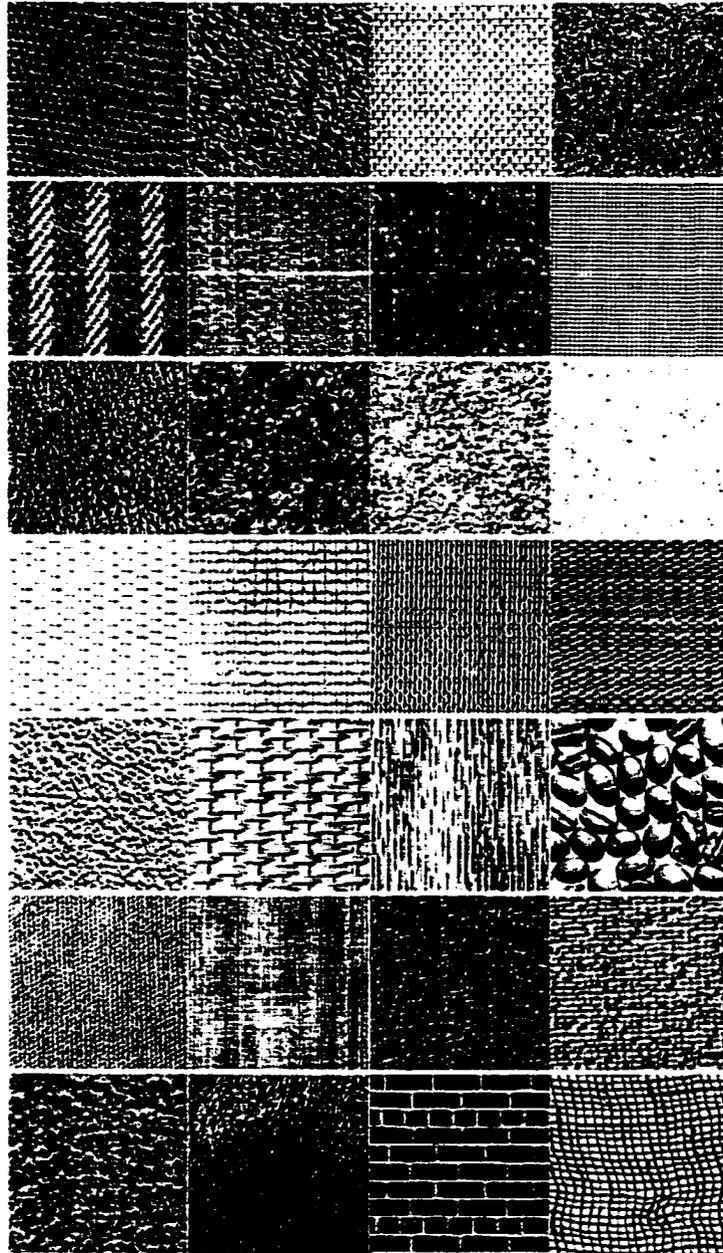


Figure 5.1 Brodatz texture data used in experiment I. First row: D3, D4, D6, D9, second row: D11, D16, D19, D21, third row: D24, D28, D29, D32, fourth row: D34, D52, D53, D55, fifth row: D57, D65, D68, D74, sixth row: D77, D78, D82, D84, seventh row: D92, D93, D95, D103.



Figure 5.2 Visually similar pattern textures used in experiment II: D4, D29, D32, D57, D77.

ment II. They consist of speckle-like patterns and exhibit similar size texture primitives that are not easily discriminated by the human eye.

The original image whose size is 256×256 with 256 gray levels, is subdivided into 64 non-overlapping samples of size 32×32 , resulting in a total of 1792 and 320 samples for experiment I and II, respectively. Fifty percent of samples are used for learning, and the rest are used for testing the classifier. The classifier performance is evaluated using two different randomly selected learning and testing sets. For both experiments, the mean of the sample is extracted from each sample in order to render the sample texture fairly independent of the first order statistics.

5.2.2 Experiment III

A total of 30 textures are obtained from the VisTex database [81] which is used in the study by Wouwer *et al.* [84]. VisTex database is gaining acceptance as a standard image database. The VisTex images offer an advantage over the Brodatz images. Both the Brodatz images and VisTex images are diverse. However, Brodatz tends not to have multiple images of similar scenes, and so very few pairs of Brodatz textures are difficult to distinguish. On the other hand, many of VisTex images are of similar scenes. It is possible, therefore, to find pairs of textures which are difficult to distinguish visually. This allows a more challenging texture problem to be compiled [76]. The results in Section 5.3.3 show that the texture problems drawn from the VisTex images are more difficult than those drawn from the Brodatz images. Again, the sample mean is subtracted from each sample.

As in [84], sixty-four non-overlapping samples size of 64×64 are obtained from an original image size of 512×512 , resulting in a total number of 1920 samples of which half are used for training and the rest are used during the testing phase. Fig. 5.3 shows the 30 texture images whose size is 512×512 .

5.2.3 Bayes Distance Classifier

We use a simple classification method in which the decision function is based on the distance measure. since our main objective in this study is to see how the addition of RWFB features improves the texture classification performance relative to those obtained using only SWFB. Among the distance measures including Euclidian and Mahalanobis distances. Bayes distance is found to provide the best performance in [5]. However. Mahalanobis distance performs as close as the Bayes's [67].

In the learning phase. n samples obtained from the texture class i are decomposed using SWFB and RWFB and a feature vector \mathbf{F} is calculated for each sample. Then. a representative feature map. \mathbf{m}_i . for each texture class i is generated by averaging the features over all n samples. Although the exact form of joint density function of features are not available. we may assume. invoking the law of large numbers. that the density function of features is Gaussian so that the Bayes decision function can be assumed as

$$D(\mathbf{F}, \mathbf{m}_i, \mathbf{C}_i) = (\mathbf{F} - \mathbf{m}_i)^T \mathbf{C}_i^{-1} (\mathbf{F} - \mathbf{m}_i) + \ln |\mathbf{C}_i|. \quad (5.7)$$

where \mathbf{C}_i is the covariance matrix of feature set for texture i measured from n samples.

For the classification phase. we employ the following steps:

1. Decompose an unknown texture and construct a feature vector \mathbf{F} .
2. Calculate the decision function $D_i = D(\mathbf{F}, \mathbf{m}_i, \mathbf{C}_i)$ for the candidate textures.
3. Assign the unknown texture to a class i if $D_i < D_j$ for all $j \neq i$.

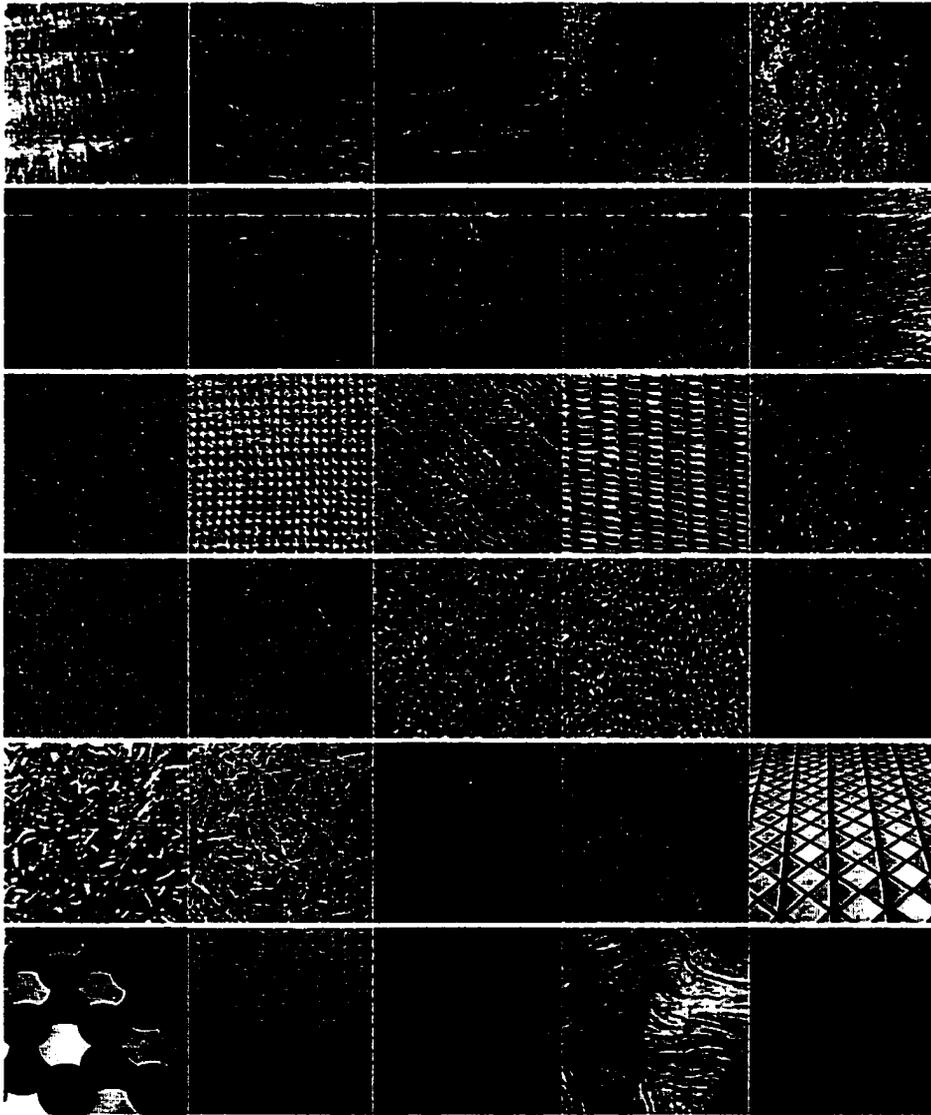


Figure 5.3 VisTex images for experiment III: from left to right: first row: Bark0, Bark4, Bark6, Bark8, Bark9; second row: Brick1, Brick4, Brick5, Fabric0, Fabric4; third row: Fabric7, Fabric9, Fabric11, Fabric13, Fabric16; fourth row: Fabric17, Fabric18, Food0, Food2, Food5; fifth row: Food8, Grass1, Sand0, Stone4, Tile1; sixth row: Tile3, Tile7, Water6, Wood1, Wood2.

5.3 Results and Discussions

The three-level ($I = 3$) wavelet decomposition using DWFB is used for texture feature extraction. The performance of SWFB and RWFB approaches for texture classification are presented and discussed in this section. Each accuracy figure represents the results of an average of two tests conducted with two different randomly selected learning and testing sets. Experiments are performed using feature sets:

- set 1 - SWFB only
- set 2 - RWFB only
- set 3 - DWFB
- set 4 - SWFB and channel outputs of filters H_1^r, H_2^r components of the RWFB

The feature set 4 was selected in a manner that reduces the redundancy in feature set 3. This is possible since the overlap between the filter responses of H_0^r and H_0^s is quite significant in the frequency domain, and this results in a lot of redundant information [for example, see Fig. 4.4(b) and (c)]; the filter response of H_3^r is completely covered by those of the H_1^s and H_2^s filters.

The classification results obtained from the three experiments, showing the percent of correct classification as a function of decomposition level for each feature set, are listed in Tables 5.1, 5.2, and 5.3. The number of features used in each experiment is shown within the parenthesis next to the classification rate.

5.3.1 Experiment I

The performance obtained using the SWFB features (set 1) is better than that obtained using the RWFB features (set 2) for each scale level (I), especially level 1. However, the performances are approximately the same for levels 2 and 3. The classification accuracy improves as the number of scale levels increases for both sets as in [80]. However, the degree of improvement from level 2 to 3 is not significant.

The combination of the two feature sets (set 3) always outperformed each method individually which is consistent with our expectation. In particular, a significant improvement was

Table 5.1 Percent of correct classification results obtained in experiment I as a function of the number of decomposition scale levels I .

number of scale (I)	set 1	set 2	set 3	set 4
1	93.42% (4)	89.34% (4)	97.16% (8)	96.65% (6)
2	95.78% (7)	95.76% (7)	97.10% (14)	97.71% (11)
3	96.13% (10)	95.98% (10)	96.20% (20)	96.82% (16)

observed for scale level 1. It is also found that increasing the scale level affects the classification performance adversely for set 3. This may be due to the increased level of ambiguity in the feature space as the degree of redundancy among the features increases. The reason is also cited by Mao and Jain [57]. Given a fixed number of samples, increasing the number of features for each sample has two conflicting effects: the additional features provide more information that improves a classification accuracy; the sample of instances is less representative of the true distribution which degrades classification accuracy on unseen instances. Typically, as the number of features increases, classification accuracy on unseen instances rises to peak, then degrades. This is known as the peaking phenomenon [31].

Following a systematic study to identify the best combination of the features, feature set 4 was selected based on the reasons that we mentioned earlier. The results show that set 4 provides better performance relative to set 3 except in the first scale in which the number of features did not appear to be adequate. The highest classification rate (97.71%) was obtained using set 4 with level 2. Thus, reducing redundancy in the feature space appears to result in performance improvement.

It is worth noting that the highest classification rate obtained using a single filter bank did not match the lowest rate obtained with the combined methods (set 3 or 4) even at a higher scale level. These findings show that the RWFB method provides additional information relating to the texture and shows the effectiveness of the RWFB in texture classification. The experiments also show that increasing the number of scale levels does not guarantee improved

Table 5.2 Percent of correct classification results obtained in experiment II as a function of the number of decomposition scale levels.

number of scale (I)	set 1	set 2	set 3	set 4
1	94.79% (4)	95.84% (4)	97.92% (8)	97.91% (6)
2	94.26% (7)	96.35% (7)	97.39% (14)	98.69% (11)
3	93.74% (10)	95.56% (10)	96.87% (20)	97.39% (16)

performance for sets 3 and 4.

5.3.2 Experiment II

In this experiment, the RWFB feature set outperformed the SWFB feature set in each scale level. This improvement can be attributed to the fact that some of the discriminating patterns are diagonally distributed in the texture samples and they are better captured using the RWFB set. Combining the two feature sets, we obtain additional improvement as shown in the set 3 column. As in the case of experiment I, set 4 provides better performance relative to set 3 except in the case of scale level 1. However, the performance of set 4 with $I = 1$ is very close to that of set 3. A possible explanation is that the discriminatory power of H_{lh}^r and H_{hl}^r filters is very significant for the specific textures used in this experiment. Again, the highest classification rate (98.69%) was obtained using set 4 with level 2.

The maximum classification rate obtained from a single filter bank did not reach the lowest level obtained by combining the two feature sets despite the fact that it contained a larger number of features. It is also observed that increasing the number of decomposition levels did not improve the classification accuracy in this experiment, especially for sets 1 and 3. This may imply that the discriminatory information exists mostly in the high frequency regions $[\pi/2, \pi]$, i.e., the additional information at the deeper scale does not contribute to the performance while adding uncertainty to the process of discrimination.

Table 5.3 Percent of correct classification obtained in experiment III (thirty VisTex images) as a function of the number of decomposition scale levels.

number of scale (I)	set 1	set 2	set 3	set 4
1	90.73% (4)	86.00% (4)	93.49% (8)	93.54% (6)
2	94.11% (7)	91.61% (7)	94.84% (14)	95.73% (11)
3	93.18% (10)	90.73% (10)	89.48% (20)	92.08% (16)

5.3.3 Experiment III

SWFB features outperform RWFB features in this experiment. The combination of the two filter sets improves the performance for scale levels 1 and 2. However, we observe a significant peaking in performance for Set 3 and Set 4 as the scale level increases. Again, the best performance (95.73%) is obtained from Set 4 with $I = 2$. The highest classification accuracy achieved by using DWFB feature set is higher than combined set of first- and second-order statistic features [84] by 1.75%.

According to [84], the 30 VisTex images are selected such that the classification performance using energy features alone is $< 90\%$, thus creating a sufficiently hard classification problem. In [84], the classification accuracy 83% is achieved by using energy features alone. However, we observe the classification performance using energy features alone is greater than 90% even with one level decomposition. The only difference between the proposed scheme and the one in [84] is the wavelet basis (Haar in our method and second-order spline wavelet in [84]) and the number of decomposition level (3 and 4, respectively).

We also observe that discriminating VisTex images is more difficult than discriminating Brodatz images, even with a larger sample size (64×64 for VisTex and 32×32 for Brodatz). As mentioned earlier, this may be caused by confusion drawn from multiple scenes of the same class in the VisTex database.

5.4 Summary

The effectiveness of the new filter set is illustrated by applying it to image texture classification tasks. The RWFB decomposition provides complementary texture information to the SWFB decomposition. Consequently, a combined feature set results in significantly better accuracy rate than a feature set obtained from a single filter bank despite the fact that a smaller number of features is employed.

For the classification of 28 Brodatz texture classes, we achieved 97.71% classification accuracy using 11 energy measures obtained from 32×32 sample size. We also obtained 98.69% accuracy in the discrimination of five similar pattern texture types using the same feature set used in experiment I.

For the classification of 30 VisTex texture classes, we achieved 95.73% classification accuracy using eleven energy measures obtained from DWFB. As observed from the accuracy figures, we may conclude that the VisTex images are not as well characterized as Brodatz images by using energy parameters. This observation agrees with the observations made in [84]. In addition, the classification accuracy that was achieved by using DWFB is higher than the best feature combination used in [84]. This shows the significance of the proposed texture features and their applicability to general natural texture characterization.

As mentioned earlier, the classification performance can be optimized by employing an adaptive method for selecting the best feature combination to minimize redundancy among features. In addition, research needs to be carried to identify a perceptually meaningful texture feature set using DWFB. An extension of these ideas to other pattern recognition applications such as texture segmentation or denoising applications can be also made.

CHAPTER 6 VIEWPOINT-INVARIANT TEXTURE CLASSIFICATION

In this chapter, the viewpoint-invariant texture classification problem, in particular, the rotation- and grayscale transform- invariance issue, is addressed. We begin with a description of a prior work in the area followed by a discussion of a new approach for obtaining rotation- and grayscale transform-invariance via energy feature transformation. The performance of the proposed texture feature set is evaluated by applying it to two different texture databases that have been used in literature in order to provide a basis for reference.

6.1 Introduction

Several approaches have been introduced to address the texture rotation-invariance problem. Kashyap and Khotanzad [36] developed the “circular autoregressive” model to overcome the inherent rotation dependency of Markov Random Field models. In [9], textures are modeled by different Gaussian Markov Random Fields (GMRF). The rotation and scale parameters were made part of the model through an appropriate transformation of the power spectral density of the GMRF.

A set of Gabor filters has been used for computing rotation-invariant texture features by considering the rotation invariance in the spatial domain as a translation invariance in a frequency domain [78]. Chen and Kundu [8] use the wavelet transform and a Hidden Markov Model (HMM) method. The energy and entropy features are derived from each subband image. They achieve 93.33% classification accuracy using a HMM for the classification of ten Brodatz [3] texture types. The HMM captures the trend caused by rotation and provides effective information concerning dependence among the subbands. Wu and Wei [87] use a spiral resampling technique along with the subband decomposition and a HMM. Prior to the

subband decomposition, the image is spirally resampled into a 1-D signal and the texture features are obtained from high-order statistics of the subband signals. Here, the rotation-invariance is reflected as translation-invariance. The highest accuracy they achieve is 95.14% for 16 Brodatz texture types. However, the texture features in [8] and [87] are not truly rotation-invariant by themselves, since the rotation-invariance is captured by the classification method, HMM.

Porter and Canagarajah [67] provide a comparative study among GMRF, Gabor filter, and wavelet transform approaches to the problem of 16 Brodatz texture classification. To obtain rotation invariance, a circularly symmetric neighborhood in all directions is used for a GMRF model instead of using a conventional neighborhood structure. A circularly symmetric Gabor filter set and discrete wavelet transform are used to compute additional sets of rotation-invariant features. In the wavelet approach, the LH and HL subimages in each scale are grouped together to obtain rotation invariance. It is shown that the wavelet approach is most efficient in terms of computational expense and classification accuracy on 16 Brodatz images. However, it is not safe to conclude that the wavelet approach is superior than Gabor and GMRF approach in obtaining textures' rotation invariance as the authors claim, since the features that the authors compute using the Gabor and GMRF approach may not be the best set that one can obtain using those approaches. For example, the number of Gabor filters can be increased, and the order of GMRF model can be increased. In addition, the performance differs from one texture database to another database.

Recently, Haley and Manjunath [28] proposed a complete space-frequency model to obtain a set of rotation-invariant features. A polar, analytic form of 2-D Gabor wavelets is developed and used to compute the model parameters which are invariant with respect to texture rotation. Unlike other studies mentioned above, the rotation of texture is not obtained by image processing in order to provide more accurate texture rotation. Thirteen classes of Brodatz images with seven rotation angles in each class are used in the experiment, resulting in 96.8% classification accuracy. However, the number of features used is significantly larger than other filtering-based approaches.

The goal of feature extraction is to obtain a texture measure which can be used to discriminate among different texture pattern classes. Most wavelet-based texture analysis algorithms extract features such as energy or entropy from the channel outputs. In this study, we look into an intelligent method to generate a set of features that are independent of texture rotation by transforming the energy measures into another set of feature. The feature set we will be presenting is unique in that rotation- and grayscale transform-invariance are achieved simultaneously and a few of the features are perceptually meaningful.

6.2 Grayscale Transform-Invariance

Texture features that are invariant to grayscale changes due to variations in illumination levels are very desirable in many cases, since it is very unlikely that images are acquired under identical lighting conditions. Using histogram equalization, the undesirable effects can be eliminated to some extent as shown in [8]. The third- and fourth-order moments both normalized with respect to the second-order moment are also used to obtain grayscale transform-invariance [8].

In this study, the grayscale transform-invariance is achieved in two steps. First, the effect of shift is removed by subtracting the mean of the signal to null the effect of β . The effect of a scale constant α is removed in feature formulation as described below.

Instead of using the standard energy measure, we use a normalized energy that is defined as

$$\varepsilon_k = \frac{E\{\|y_k(m, n)\|^2\}}{E\{\|x(m, n)\|^2\}} \quad (6.1)$$

where $y_k(m, n)$ is the channel output of the original image. That is, the channel energy is normalized by the energy of an original image $x(m, n)$. The terminology of 'normalized energy' appears in [8], the normalization, however, is performed by dividing energy by a constant 255^2 .

Then, how do we obtain the invariance through normalization? Let the gray-scale transformed image $x_t = \alpha x + \beta$, where α is a positive scale factor, and β is a shift factor. The effect of shift can be removed by subtracting the mean of x_t . Let \hat{x} be the mean subtracted image.

Then,

$$\hat{x} = x_t - \bar{x}_t = \alpha x + \beta - \bar{x}_t = \alpha x + \beta - \alpha \bar{x} - \beta = \alpha(x - \bar{x}). \quad (6.2)$$

Since the input to the filterbank is zero mean, every channel output $y_k(m, n)$ is also zero-mean except for the lowest approximation ($k = 0$). Then, the averaged l_2 - norm becomes just a variance. The variance of channel output of mean subtracted image, $\hat{y}_k(m, n)$ is

$$\begin{aligned} \text{Var}(\hat{y}_k(m, n)) &= \text{Var}(\hat{x} * \mathbf{H}_k) \\ &= \text{Var}(\alpha(x - \bar{x}) * \mathbf{H}_k) \\ &= \text{Var}(\alpha x * \mathbf{H}_k) - \text{Var}(\alpha \bar{x} * \mathbf{H}_k) \\ &= \alpha^2 \text{Var}(x * \mathbf{H}_k). \end{aligned} \quad (6.3)$$

Then, the normalized energy becomes

$$\hat{\varepsilon}_k = \frac{\text{Var}(\hat{y}_k)}{\text{Var}(\hat{x})} = \frac{\alpha^2 \text{Var}(x * \mathbf{H}_k)}{\alpha^2 \text{Var}(x)} = \varepsilon_k. \quad (6.4)$$

Thus, the effect of α is removed by calculating the normalized energy. Likewise, the normalized energy parameter we defined is independent of linear grayscale transforms. Similarly, the normalized energy obtained using l_1 - norm is also independent of the linear grayscale transform.

6.3 Subband Energy Distribution

In this section, we would like to provide a general idea about the energy distribution of filterbank channel output, especially when the texture is rotated to justify the choice of wavelet-based energy measures as texture features. We begin by showing the energy variations due to texture rotation, particularly for a texture with high directionality. Fig. 6.1 shows the energy of a SWFB channel outputs for the “brick” image [see Fig. 6.6] with a rotation angle range $[0, 2\pi]$. As shown in the figure, the energy is a function of a rotation angle θ , and we define it as $e(\theta)_k$ for channel k . We also adopt the following notation throughout this dissertation: C_{HL}^i , C_{LH}^i , and C_{HH}^i denote channels that have filter H_1 , H_2 and H_3 , respectively at each scale i .

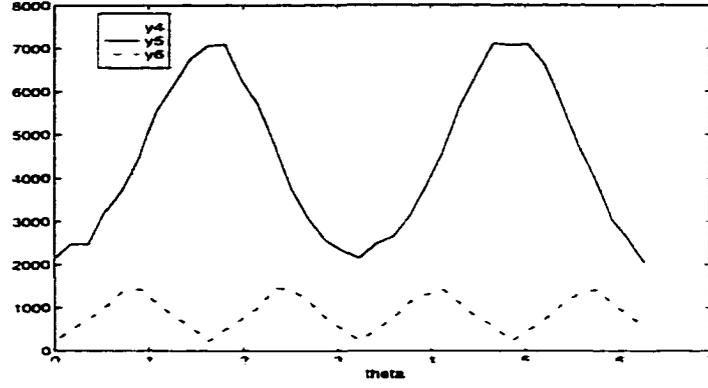


Figure 6.1 Energy distribution at the SWFB outputs $k = 4, 5, 6$ with the “brick” image as an input texture.

Since the filter characteristic of H_1^s is y -axis symmetric, and that of H_2^s is x -axis symmetric, the functions $e(\theta)_{C_{LH}^1}$ and $e(\theta)_{C_{HL}^1}$ have a period $T = \pi$ and satisfy $e(\theta)_{C_{HL}^1} = e(\theta + \pi/2)_{C_{LH}^1}$. Similarly, the $T = \pi/2$ for $e(\theta)_{C_{HH}^1}$. This is also true for the RWFB, but with a delay of $\pi/4$.

We expect to obtain a larger $\max\{e(\theta)_{C_{HL}^1} - e(\theta)_{C_{LH}^1}\}$ for directional textures. This will be clearly demonstrated throughout this section. Fig. 6.2 shows a subband energy distribution from a SWFB for six selected texture images (shown in Fig. 6.6) that have different directionality and coarseness. In each plot, there are seven solid lines that correspond to seven rotation angles from 0 to 210 degrees. For the textures that have round-shape primitives (for example, “bubbles” image), the influence of rotation is not significant, that is, the variation in energy is not significant with respect to rotation. Note that a regular but not directional texture such as “weave” image is also insensitive to the rotation. Fig. 6.3 shows a different view of Fig. 6.2, where the x -axis is the rotation angle and y -axis is the ε_k . The pattern of variations are similar in three scale levels.

As observed from the above figures, the energy parameter itself is not consistent for texture rotation, especially for directional textures. Now, we introduce a new feature set that is independent of texture rotation by making use of the normalized energy measures from a dual wavelet filterbank.

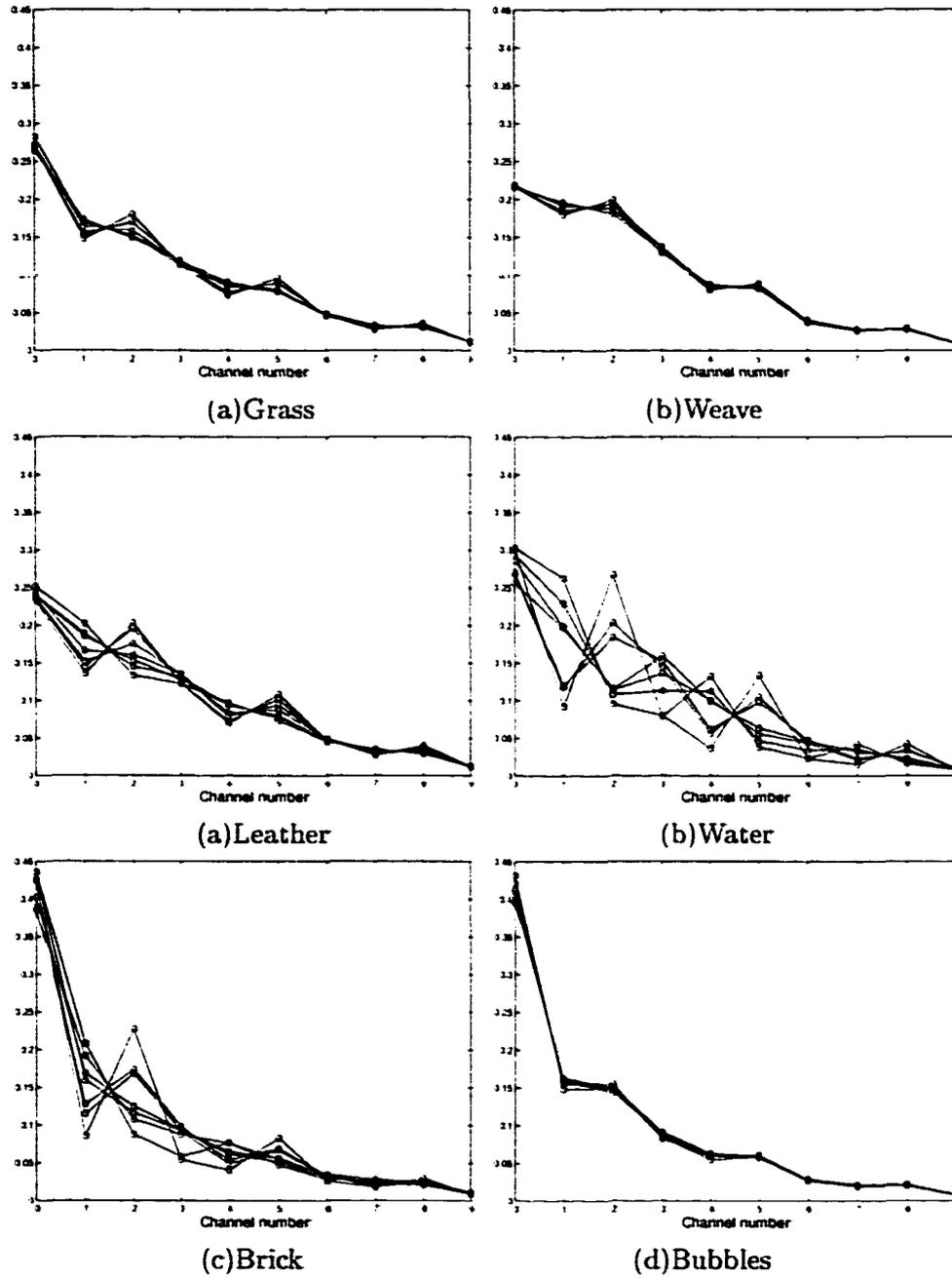


Figure 6.2 Subband energy distribution from a SWFB for selected texture images that have different directionality. Multiple lines in each plot indicate seven rotation angles: 0, 30, 60, 90, 120, 150, and 200 degrees.

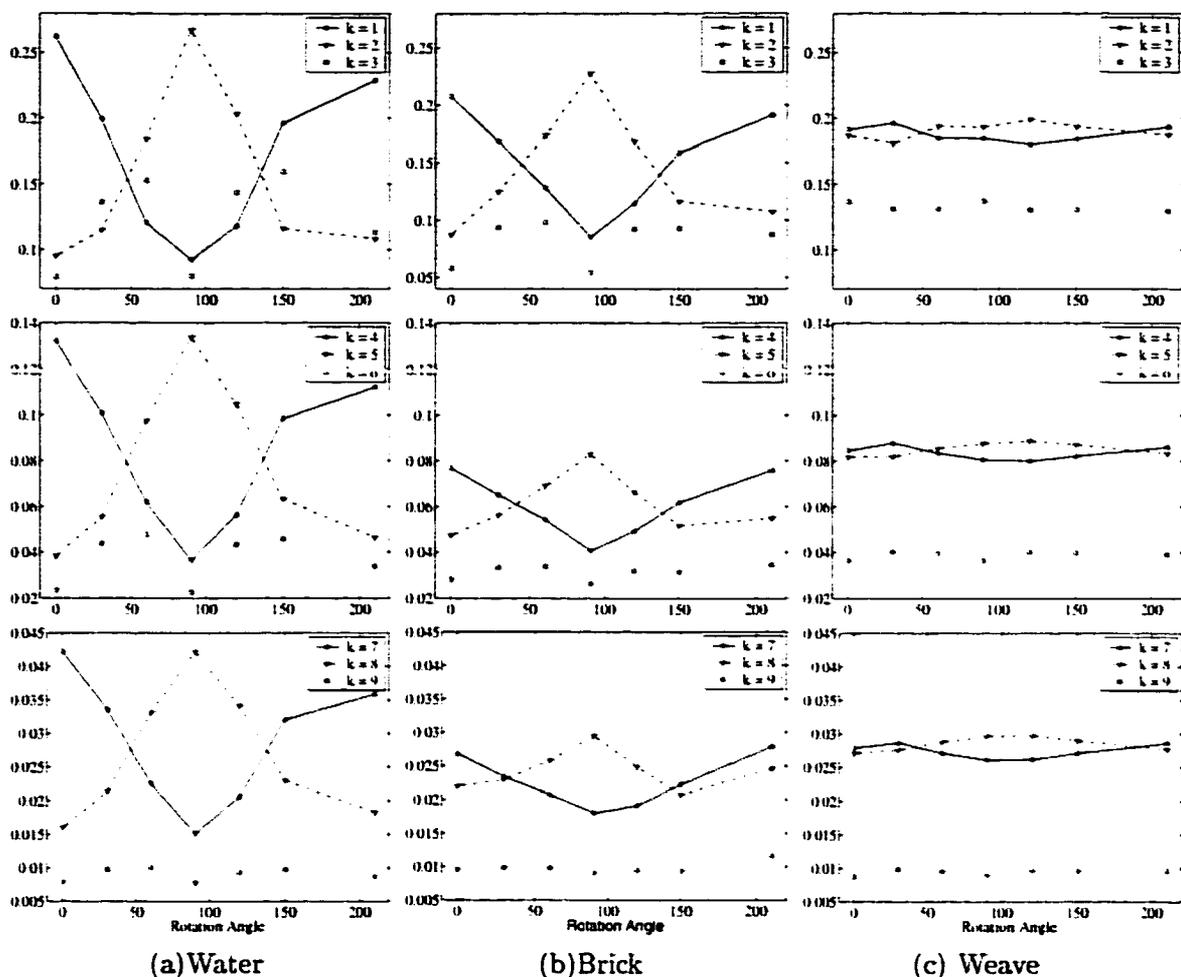


Figure 6.3 Different view of Fig. 6.2 for three directional textures. From top row to bottom, scale $I = 1, 2,$ and $3,$ respectively.

6.4 Rotation-invariant feature extraction

6.4.1 Scale Energy-Based Features

Although the energy vector $\vec{\epsilon}$ itself is not able to provide the rotation-invariant measure of a texture, we can approximate the rotation invariance by adding the three channel energies in each scale. The group of three corresponding frequency regions are illustrated in Fig. 6.4 for the SWFB and RWFB. Although the shape of the region is not exactly circular, the energy within the same scale is roughly constant in spite of texture rotation. For example, the addition of three functions $e(\theta)_{4,5,6}$ in Fig. 6.1 is roughly constant. The area of the corresponding frequency

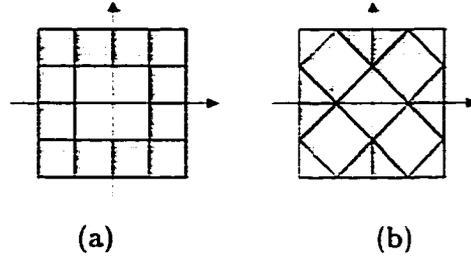


Figure 6.4 Frequency coverage for scale sum energy feature, F_{SSE} for (a) SWFB and (b) RWFB.

region in both filterbanks is same in each scale.

6.4.1.1 Scale Sum Energy

We define the feature *scale sum energy* at each scale i as follows:

$$F_{SSE}^i = \sum_{k=3i-2}^{3i} \varepsilon_k = \sum_{k=3i-2}^{3i} \frac{\|y_k(m, n)\|^2}{MN}, \quad i = 1, 2, \text{ and } 3. \quad (6.5)$$

The F_{SSE} features can be also obtained by $R\varepsilon^T$, where R is a matrix

$$R = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}$$

6.4.1.2 Scale Energy Ratio

We define *scale energy ratio* features that measure the relative energy amount in two consecutive scales. The scale energy ratio is defined as

$$F_{SER}^{i-1,i} = \frac{F_{SSE}^{i-1}}{F_{SSE}^i}, \quad i = 1, 2, \text{ and } 3. \quad (6.6)$$

where F_{SSE}^0 is the energy of smoothest version of the image, y_0 , and it can be also obtained by $1 - \sum_{i=1}^3 F_{SSE}^i$. Fig. 6.5 helps visualization of this feature. For example, a low $F_{SER}^{2,3}$ value implies that the texture has higher energy in the 3rd scale compared to the 2nd scale, which implies that the texture has more frequency components within $[\pi/2, \pi]$ than the $[\pi/4, \pi/2]$ range.

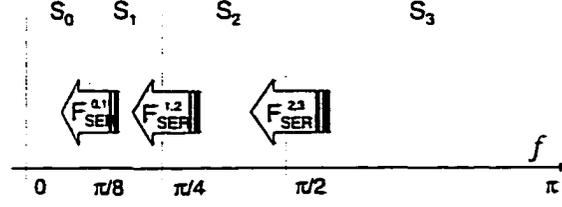


Figure 6.5 Visualization of SER feature.

Energy spread across scales are different for different textures. Hence, we introduce another feature set that measure the relative energy change across scales by computing the ratio of F_{SER} features, incorporating energy information from three consecutive scales. The parameters are defined as follows:

$$F_{RSE}^1 = \frac{F_{SER}^{0,1}}{F_{SER}^{1,2}} = \frac{F_{SSE}^0 F_{SSE}^2}{(F_{SSE}^1)^2} \quad (6.7)$$

$$F_{RSE}^2 = \frac{F_{SER}^{1,2}}{F_{SER}^{2,3}} = \frac{F_{SSE}^1 F_{SSE}^3}{(F_{SSE}^2)^2}$$

Since these features measure the relative energy change from high to low frequency, these parameters are expected to be good indices of texture coarseness/fineness.

6.4.2 Directionality Feature

Channels C_{HL}^i and C_{LH}^i are very sensitive to rotation as shown in Fig. 6.1 through Fig. 6.3 for both filter banks. The more directionality we have, the higher energy variation we obtain from the channel outputs. There is a small energy change for a texture image with round objects in it or a very structured texture with various texture orientations.

Let $D(\epsilon_{C_{HL}^i}, \epsilon_{C_{LH}^i})$ represent the difference in energy between the two channels (C_{HL}^i and C_{LH}^i), where $D(a, b) = |a - b|$. Then the difference (we call it D_ϵ^i) provides an indication of how much a texture is oriented in the corresponding scale i . However, this value also varies as a texture rotates [imagine $D(\epsilon_4, \epsilon_5)$ in Fig. 6.1]. Hence, the D_ϵ^i from a single filterbank cannot serve as a rotation-invariant feature. The deficiency can, however, be met by introducing D_ϵ^i from RWFB where the energy difference is obtained from the 45° shifted frequency regions. For example, if a texture that is strongly directional in vertical direction is rotated by 45° ,

the D_{ε}^i from SWFB reaches at its minimum, while that from RWFB reaches at its maximum.

Hence, the combination of the energy difference between C_{HL}^i and C_{LH}^i from both filter banks is expected to be constant. This is computed as follows:

$$F_{DM}^i = \left[\frac{D(\varepsilon_{C_{HL}^i} \cdot \varepsilon_{C_{LH}^i})}{F_{SSE}^i} \right]_{SWFB} + \left[\frac{D(\varepsilon_{C_{HL}^i} \cdot \varepsilon_{C_{LH}^i})}{F_{SSE}^i} \right]_{RWFB}. \quad (6.8)$$

In summary, the new feature set F includes F_{SSE} , F_{SER} , F_{RSE} , and F_{DM} . The features are rotation-invariant as well as grayscale transform-invariant, since they are all computed based on the normalized energy measure which is grayscale transform-invariant.

6.5 Experimental Setup

The proposed feature set is evaluated by conducting two different classification experiments using two different data sets. The texture databases are chosen mainly because we would like to evaluate the performance of the proposed feature set relative to results obtained with alternative approaches proposed in recent years [28, 84]. For the first experiment (experiment I), we use the data set used by Haley and Manjunath [28]. The rotated texture images are obtained from the USC-SPIP Image database [74]. The database consists of thirteen of the Brodatz texture images, each digitized at seven different rotation angles: 0, 30, 60, 90, 120, 150, and 200 degrees. The images are not obtained by artificial rotation which may contribute to distortion in textures, particularly when the texture is very finely structured. In addition, the range of textures is broad in terms of coarseness, directionality, and regularity. The effect of texture rotation on energy measures is significant when the texture is strongly directional, causing difficulties in obtaining rotation invariance. For perfectly homogeneous and isotropic texture, texture features ought to be ideally rotation-invariant. This point is often overlooked in the literature when texture selection is made.

The total number of images is 91 and the images are all of size 512×512 pixels with 8 bits/pixel. The rotated texture images were scanned using a 512×512 pixel video digitizing camera. Fig. 6.6 presents these textures showing 256×256 size samples with a rotation angle of 60° . The original image is then subdivided into sixteen 128×128 samples without overlapping.

resulting in a total of $91 \times 16 = 1456$ samples. Examples of the rotated “straw” texture are shown in Fig. 6.7, indicating sixteen 128×128 samples. As seen from the figure, the straw image is not considered as a homogeneous texture. Several samples possess different texture properties relative to the rest of them. Such problems render the classification task even more difficult.

For the second experiment, thirty VisTex images used in [84] are used to evaluate the robustness of the proposed rotation- and grayscale transform-invariant feature set with respect to the problem of larger and different databases. The experimental setup is described in Section 5.2.2, and is identical to the one used in [84].

A three-level ($I = 3$) wavelet decomposition using DWFB is used for rotation-invariant texture feature extraction. The texture sample with its mean subtracted is decomposed through 20 channels of DWFB and two normalized energy vectors are generated for each sample. Again, the Haar wavelet is used for its computational simplicity. The energy vectors are calculated using both average l_1 -norm and l_2 -norm methods. Then, viewpoint-invariant texture features are generated by transforming the normalized energy vectors using the methods described in the previous section. The feature vector is then classified using a Bayes classifier described in Section 5.2.3. The overall texture classification scheme is presented in Fig. 6.8. Again, half the samples are used for training while the rest are used during the testing phase.

6.6 Feature Analysis

In this section, the correlation of the viewpoint-invariant features to the texture classes is presented. Thirteen Brodatz textures, used in the experiment I, were used to demonstrate the effectiveness of the approaches. One sample is extracted from each rotation angle, and hence, we have seven samples for each texture class.

6.6.1 Scale Energy-based Features

The correlation plot of the F_{SSE}^i parameters with thirteen texture types is shown in Fig. 6.9. The parameters are approximately constant with respect to the rotation. The parameters from

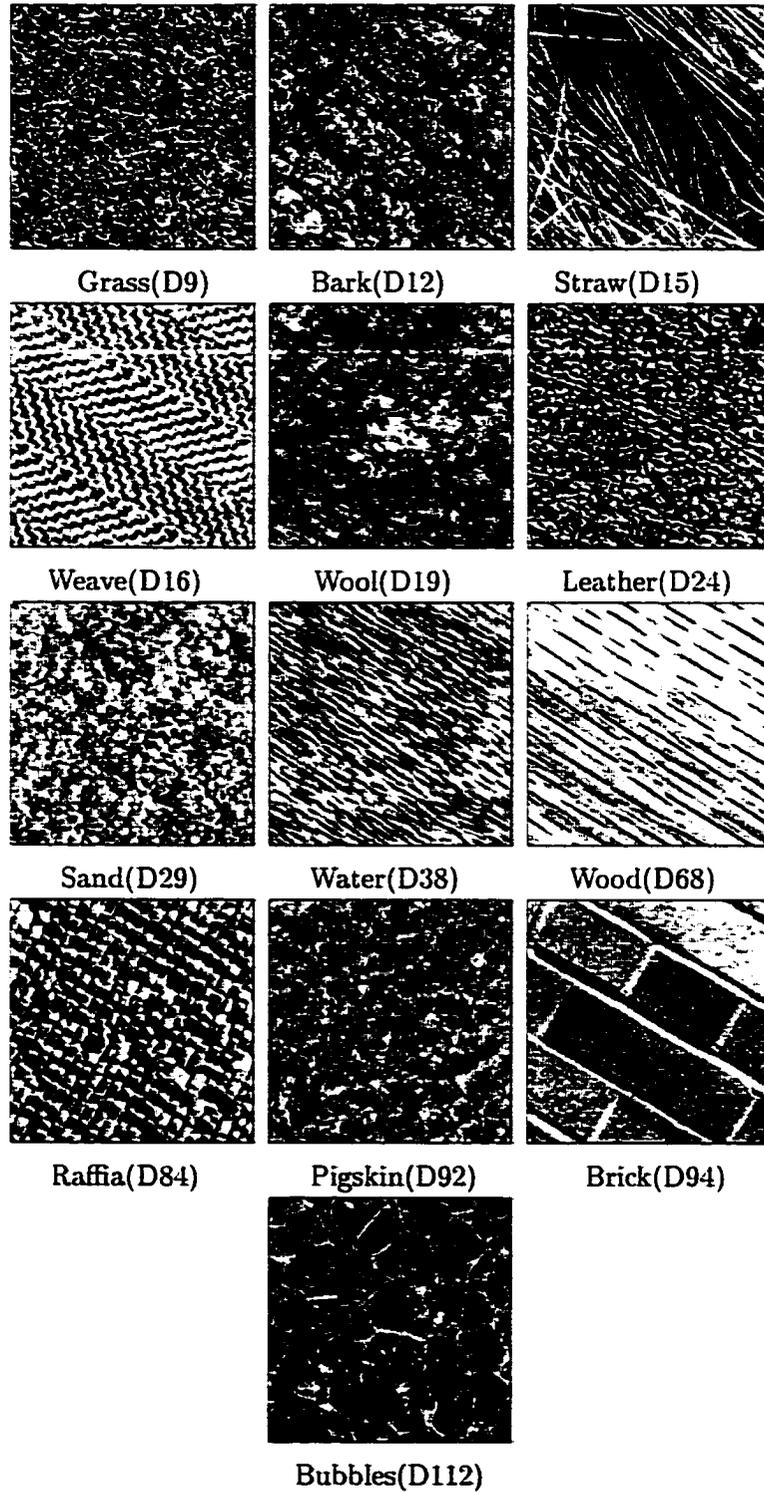


Figure 6.6 Textures used in the experiment. Each texture was digitized at seven different rotation angles: 0, 30, 60, 90, 120, 150, and 200 degrees. The size displayed here is 256×256 .

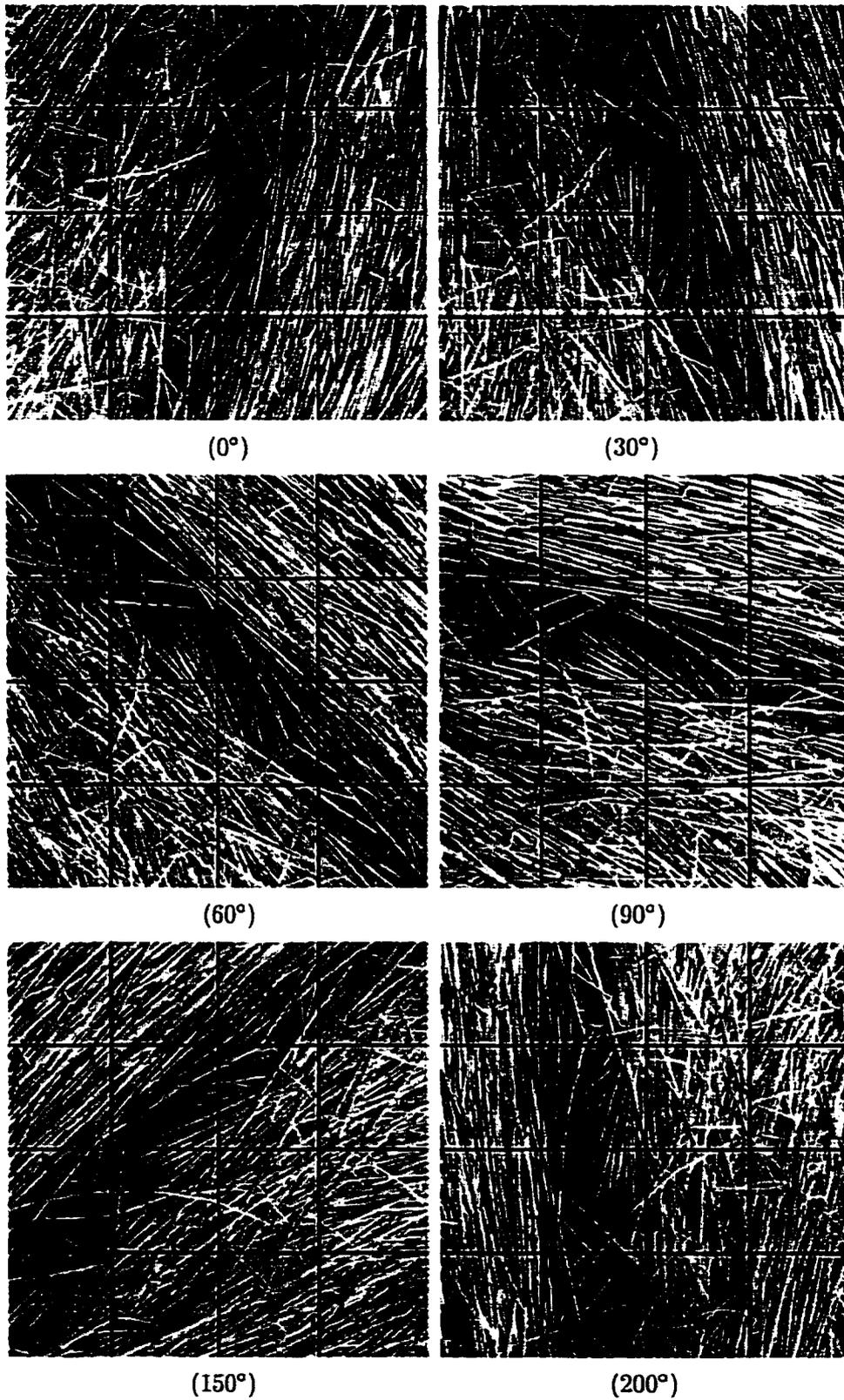


Figure 6.7 Examples of rotated texture image "straw". Sixteen non-overlapping samples of size 128×128 are shown.

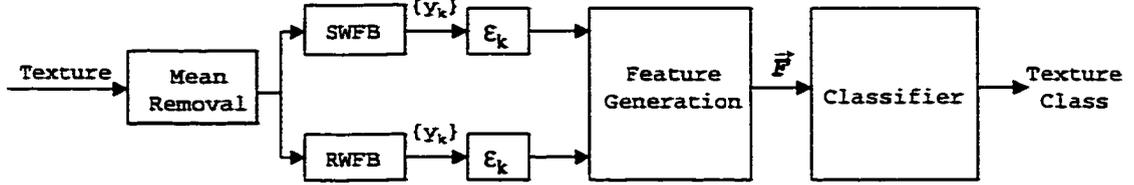


Figure 6.8 Overall diagram of rotation-invariant texture classification procedure.

the two filter banks are somewhat similar in magnitude. Fig. 6.10 presents the correlation between the F_{SER} parameters and the thirteen texture classes. They are also roughly constant within the same texture range. The difference between the two FBs is slightly amplified compared to the F_{SSE} parameters. As mentioned earlier, these features measure the relative energy in the two consecutive scales and the combination of three F_{SER} features serves as a measure of coarseness of a texture. For example, the texture “wool” has low, medium, and high value of $F_{SER}^{2,3}$, $F_{SER}^{1,2}$, and $F_{SER}^{0,1}$, respectively, which implies that the texture is very coarse. In contrast the “weave” texture has high, medium, and low values for the three features. This variation is reflected in the parameter F_{RSE} . A correlation plot for the parameters is shown in Fig. 6.11. The parameter F_{RSE}^2 appears to measure the coarseness of texture very well. For example, “bark”, “wool”, and “brick” images which are coarse have high F_{RSE}^2 for both SWFB and RWFB. Since the effect of F_{SSE}^1 is dominant in F_{RSE}^1 , it is highly correlated with $1/F_{SSE}^1$ and is not used in the experiment.

6.6.2 Directionality Measure

In order to gain a better understanding of the F_{DM} feature, we show how a constant value of directionality is obtained through an example presented in Fig. 6.12. D_{SWFB} and D_{RWFB} in the figure represents the first term and second term of Eq. (6.8), respectively. Although either D_{SWFB} or D_{RWFB} varies due to texture rotation, the sum of the two terms stays fairly constant, which implies that $D_{SWFB}(D_{RWFB})$ provides complementary information to $D_{RWFB}(D_{SWFB})$. The correlation plot of F_{DM}^i features with 13 texture types is shown in Fig. 6.13. The feature matches well with human perception. As expected, “straw”, “water”,

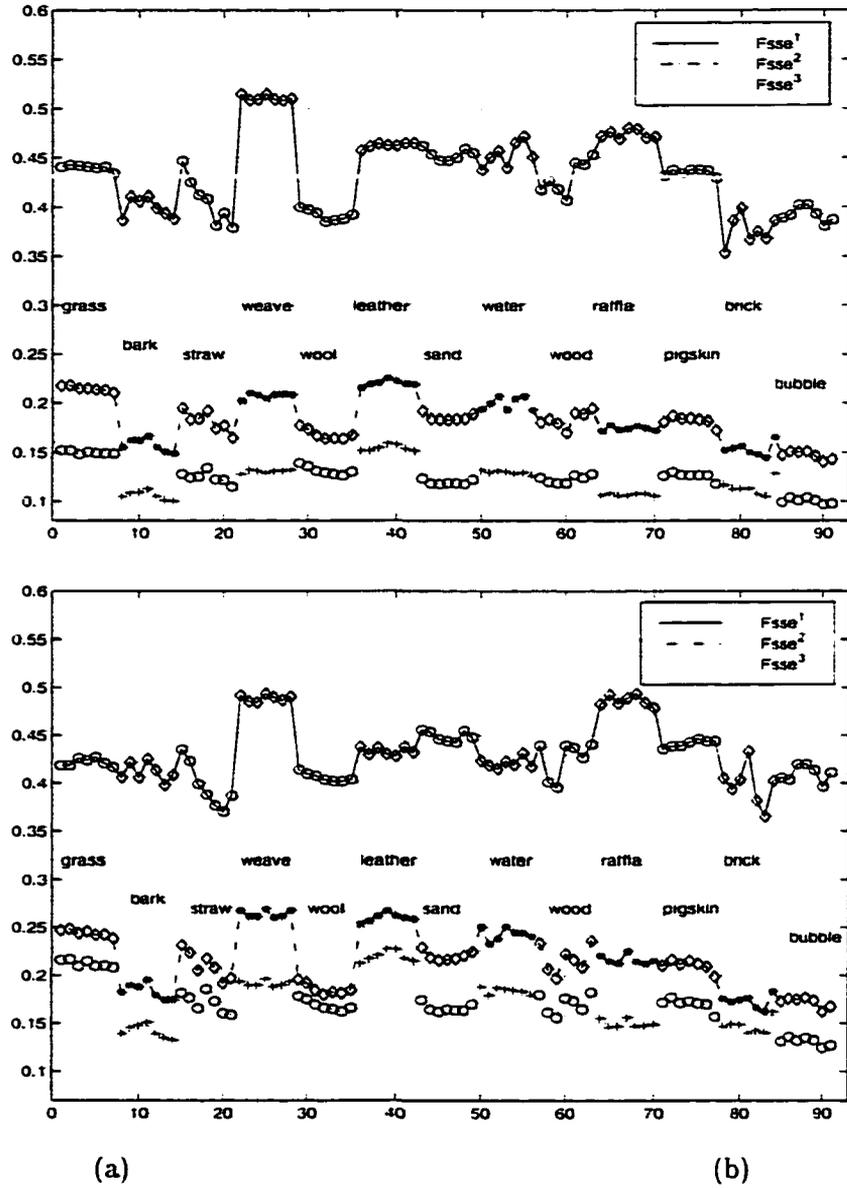


Figure 6.9 Scale sum energy features from the output of (a) SWFB and (b) RWFB. The F_{SSE}^3 was amplified two times for better visualization.

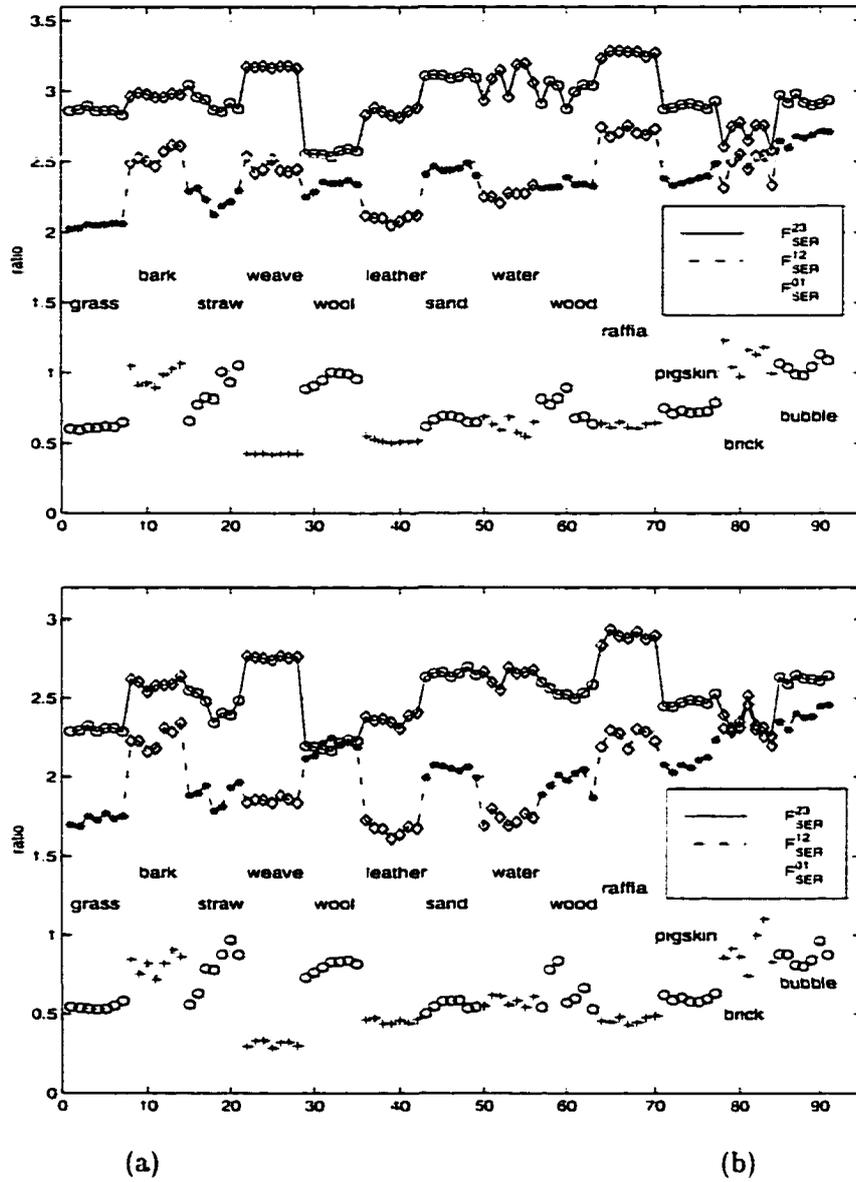


Figure 6.10 Scale energy ratio features from the output of (a) SWFB and (b) RWFB.

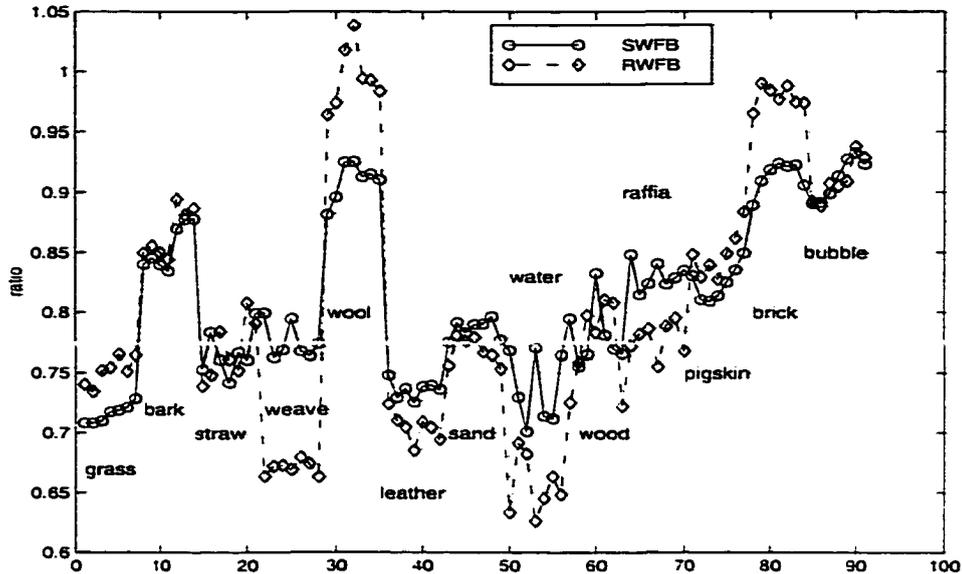


Figure 6.11 Correlation plot of F^2_{RSEr} feature with thirteen texture classes.

“wood”. and “brick” show relatively high directionality in all the scales. It is interesting to note the changes across the scales. The parameter for “brick” and “bark” increases as the scale goes down, which implies that frequency components associated with the directionality lie towards the low frequency. The F_{DM} parameter for “water” decreases as the scale goes from 2 to 1 showing that the texture loses its directionality in the lower frequency range. It follows human perception, since the “water” image contains more high frequency components than the “wood” image.

6.6.3 Feature Selection

Some of the features presented previously may be highly correlated with each other. Since redundant and ambiguous information may adversely affect the classification performance, the choice of features is important. Most feature selection methods involve high computational complexity, and do not always lead to an optimal selection. In this study, we use a very simple feature selection algorithm described as follows:

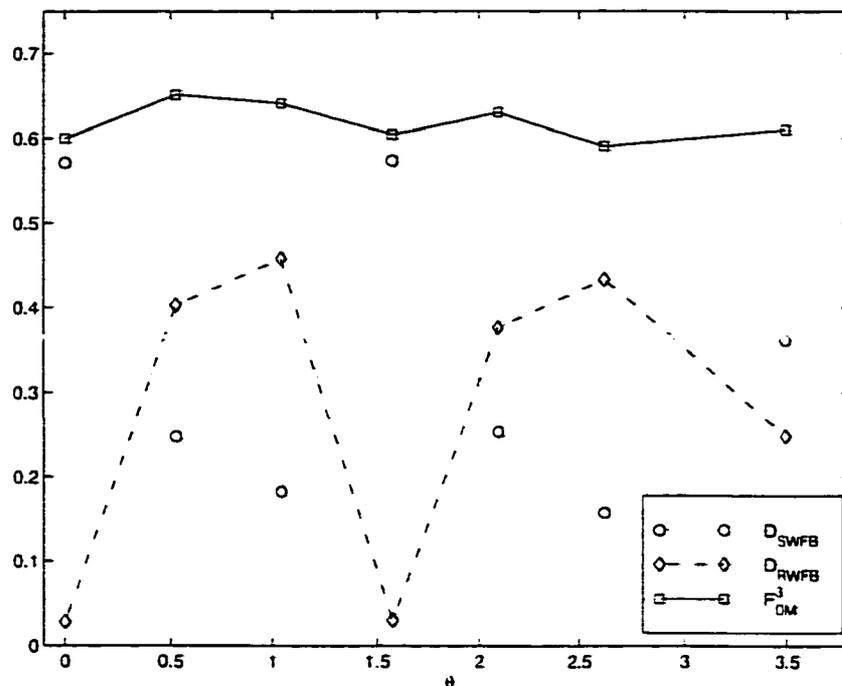


Figure 6.12 Normalized distance D for both filter banks and the summation of the two for directionality measurement (F_{DM}^3) of "wood" texture.

for $i = 1$:feature dimension n

for $j = 1$: n

compute $r_{ij} = \frac{Cov(f_i, f_j)}{\sigma_i \sigma_j}$

if $r_{ij} > p$ ($i \neq j$)

compute v_{ik} and v_{jk} ,

where $v_{ik} = \frac{\sigma_{ik}^2}{m_{ik}}$ for each texture class k

discard f with higher $\sum_k v_{ik}$

end

end

end

where m_i and σ_i are a mean and a standard deviation of f_i , respectively, and p is a constant determined by the user. In this experiment, p was set at 0.95.

To evaluate each feature category, we performed classification experiments using a different

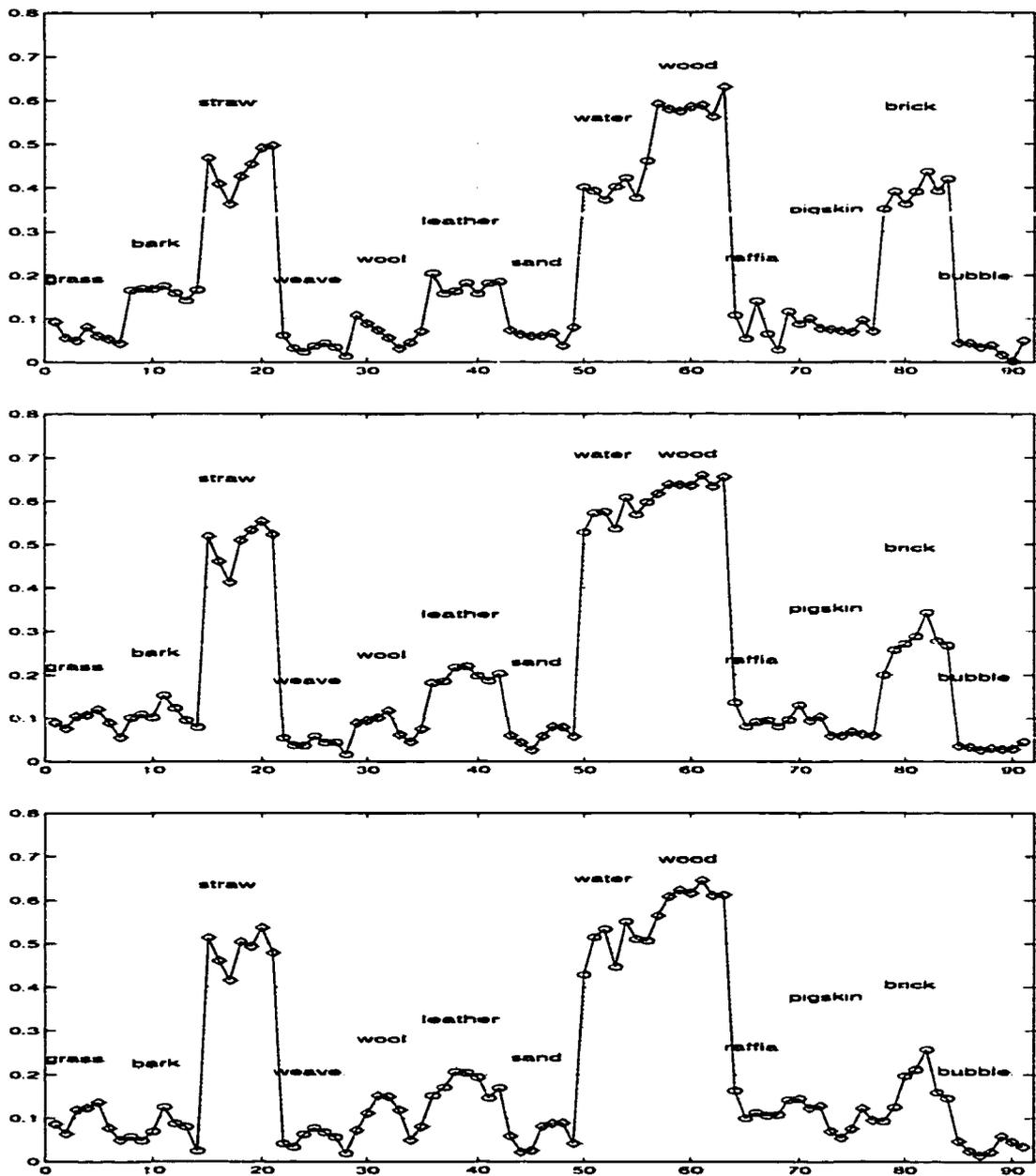


Figure 6.13 Correlation of F_{DM}^i with thirteen texture classes, where i denotes the scale level. (From top to bottom) F_{DM}^1 ; F_{DM}^2 ; F_{DM}^3 . Note the variations across the scale level.

Table 6.1 Nine feature combinations studied in the experiment.

	Features	No. of Features
Set 1	$[F_{SSE}, F_{SER}, F_{RSE}^2]_{SWFB}$	7
Set 2	$[F_{SSE}, F_{SER}, F_{RSE}^2]_{RWFB}$	7
Set 3	$[F_{DM}]$	3
Set 4	Set 1 + Set 2	14
Set 5	Set 1 + Set 3	10
Set 6	Set 2 + Set 3	10
Set 7	Set 1 + Set 2 + Set 3	17
Set 8	$[F_{RSE}^2]_{DWFB}$ + Set 3	5
Set 9	Selected Feature Set	10

combination of features listed in Table 6.1.

The ten features selected for Set 9 using the selection method include $[F_{SSE}^1, F_{SER}^{1,2}, F_{SER}^{2,3}, F_{RSE}^2]$ from SWFB, $[F_{SER}^{1,2}, F_{SER}^{2,3}, F_{RSE}^2]$ from RWFB, F_{DM}^1, F_{DM}^2 , and F_{DM}^3 .

6.7 Classification Results

6.7.1 Experiment I

The classification performance of each feature set is presented in Table 6.2 showing the percent correctly classified using two energy measure methods. As shown in the table, there is an advantage of using l_1 -norm over l_2 -norm for energy calculation. Except for Set 1 and Set 2, l_1 -norm-based energy measures provide better classification performance than l_2 -norm measures. This can be attributed to the fact that the inter-class variance becomes larger when the l_2 -norm based energy is used. Hence, discussions are limited to results obtained using the l_1 -norm.

The features from a single FB (Set 1 and Set 2) provided about 80% classification accuracy, and F_{DM} features provided around 60% accuracy. However, a combination of Set 1 and Set 2 provided an average 14% improvement over the single FB. Another significant improvement is obtained by adding 3 more F_{DM} features to the Sets 1 and 2 (see Sets 5 and 6). Combining all three feature categories, we obtained 96.70% accuracy. The Set 8 that consists of only five perceptual features, which measure the texture's coarseness and directionality, provided an

Table 6.2 Classification performance of viewpoint-invariant texture features for thirteen Brodatz textures using two different energy measures.

Feature Set	Classification Rate (%)	
	l_1 -norm	l_2 -norm
Set 1	78.43%	80.63%
Set 2	76.99%	79.12%
Set 3	60.03%	56.94%
Set 4	91.14%	88.46%
Set 5	95.95%	94.23%
Set 6	95.40%	92.93%
Set 7	96.70%	94.78%
Set 8	94.92%	94.71%
Set 9	97.53%	95.47%

accuracy of 95%. The performance of Set 3 was improved by 35% by adding two more F_{RSE} features. Thus, these perceptual features play a very important role in classification.

The best performance achieved is 97.53 % using Set 9. This is better than the best performance (96.8%) presented in [28] in which the same experimental data and significantly more texture features (order of ten) are used. From two different experiments, only 36 samples out of a total of 1456 samples are misclassified. Major misclassification includes “grass \rightarrow straw”, “bark \leftrightarrow bubble”, “sand \leftrightarrow pigskin”, and “wood \leftrightarrow straw”, where \rightarrow and \leftrightarrow denote that an error occurs in one and both directions, respectively. However, “weave”, “leather”, “water”, “raffia”, and “brick” images are perfectly classified. Note that “water” and “wood” images that are very close each other are also perfectly discriminated. In addition, none of the single texture types showed less than 93% accuracy with Set 7 and Set 9. The result obtained using Set 9 is summarized in Table 6.3.

6.7.1.1 Evaluation of Perceptual Features

Among the features presented in the previous sections, F_{RSE}^2 and F_{DM} features show good connection to the perceptual meaning of textures. Another experiment was performed to evaluate those perceptual features. We define a coarseness measure feature (F_{CM}) parameter by taking the mean of F_{RSE}^2 features from SWFB and RWFB to obtain a single measurement

Table 6.3 Classification performance for each texture class using ten selected features.

	D9	D12	D15	D16	D19	D24	D29	D38	D68	D84	D92	D94	D112
D9	105	—	—	—	—	—	—	—	—	—	—	—	—
D12	—	109	—	—	1	—	—	—	—	—	2	—	5
D15	5	—	109	—	—	—	—	—	5	—	—	—	—
D16	—	—	—	112	—	—	—	—	—	—	—	—	—
D19	—	—	—	—	108	—	—	—	—	—	2	—	—
D24	1	—	—	—	—	112	—	—	—	—	—	—	—
D29	1	—	—	—	—	—	109	—	—	—	2	—	—
D38	—	—	1	—	—	—	—	112	—	—	—	—	—
D68	—	—	2	—	—	—	—	—	107	—	—	—	—
D84	—	—	—	—	—	—	—	—	—	112	—	—	—
D92	—	—	—	—	2	—	3	—	—	—	106	—	—
D94	—	—	—	—	—	—	—	—	—	—	—	112	—
D112	—	3	—	—	1	—	—	—	—	—	—	—	107
% Rate	93.8	97.3	97.3	100	96.4	100	97.3	100	95.5	100	94.6	100	95.5

of coarseness. Then, we divide every sample into three categories: very coarse, coarse, and fine based on the F_{CM} parameter. The results are tested with F_{DM} parameters at the next level and, then divided into three categories: very directional, directional, and non-directional. The fine texture is tested with F_{DM}^3 ; coarse with F_{DM}^2 ; and very coarse with F_{DM}^1 . The procedure and the classified textures are presented in Fig. 6.14. The results show that 81% of the samples fall into the categories shown in the figure. Almost half the “straw” images were classified as category ‘coarse, very directional’ instead of ‘fine, very directional’, and the half of “bark” samples are classified as ‘coarse, non-directional’ instead of ‘very coarse, non-directional’. Increasing the number of thresholds for F_{CM} improves classification accuracy.

6.7.1.2 Feature Evaluation for Rotation Invariance Only

The rotation-invariant features developed by Haley and Manjunath [28] are not grayscale transform-invariant. To provide a fair comparison between their method and the approach discussed in this dissertation, an additional test has been conducted. Grayscale transform-invariance is not considered in this test. This is accomplished by ignoring the energy normalization procedure. The results obtained by using rotation-invariant features are summarized in Table 6.4.

The improvement obtained in Set 1 and Set 2 over the viewpoint-invariant (rotation- and

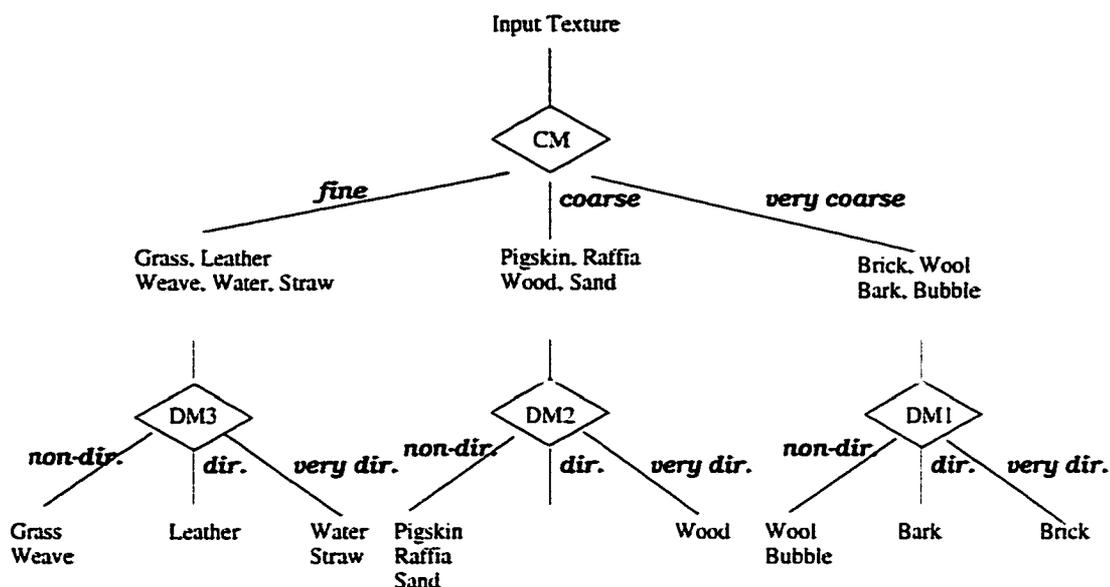


Figure 6.14 Example of tree-pruning results obtained using four perceptual texture features.

grayscale transform- invariant) feature is significant. A 12% increase in the classification accuracy is obtained. The highest classification accuracy, 99.31% is achieved using the ten selected feature set. This offers a significant improvement over the method proposed by Haley and Manjunath [28].

6.7.2 Experiment II

For this experiment, thirty VisTex images are used to evaluate the performance when a larger and different databases are used. Again, two tests have been conducted: Test I uses features with viewpoint-invariance (rotation and grayscale invariance both), while Test II uses features with rotation-invariance property only. The classification results are summarized in Table 6.5.

The first observation is that the performance of Test I is not comparable to the result obtained by using energy features developed in Chapter 5 [see Table 5.3]. This implies that the viewpoint-invariant features are not favored unless texture rotation is involved. However, by ignoring grayscale-invariance, we achieve a significant improvement in all feature sets except for Set 3 and Set 8 in which the number of features are not sufficient to discriminate thirty

Table 6.4 Classification performance of features with rotation-invariance only.

Feature Set	Classification Rate (%)
Set 1	90.93%
Set 2	88.80%
Set 3	60.03%
Set 4	95.40%
Set 5	97.53%
Set 6	97.66%
Set 7	98.01%
Set 8	94.92%
Set 9	99.31%

Table 6.5 Performance of the proposed feature set for 30 VisTex texture classification.

Feature Set	Classification Rate (%)	
	Test I	Test II
Set 1	66.77%	86.77%
Set 2	62.55%	85.57%
Set 3	37.55%	38.70%
Set 4	77.00%	90.63%
Set 5	80.57%	92.29%
Set 6	76.09%	91.93%
Set 7	81.98%	93.22%
Set 8	64.27%	66.93%
Set 9	82.60%	91.51%

texture classes. The highest accuracy obtained from Test II is 93.22%. The performance is close to the highest classification accuracy 94% obtained by Wouwer *et al.* [84].

6.8 Summary

A new set of viewpoint-invariant texture features are proposed in this chapter. Texture features are computed based on the normalized energy vectors that are obtained from the channel outputs of DWFB. The features are applied to two texture databases each used by Haley and Majunath [28] and by Wouwer *et al.* [84], respectively. They include thirteen Brodatz texture classes with 7 rotation angles in each class and thirty VisTex images with no

rotation. In order to conduct a comparative study, the experimental setup for each database is identical to the setup used by other researchers.

Our results also show that l_1 -norm based energy estimates provides better performance relative to the l_2 -norm based estimates. The highest classification accuracy we achieved for Brodatz and VisTex databases are 97.53% and 82.60%, respectively. The 97.53% accuracy is higher than the highest classification accuracy reported in [28], while 82.60% for VisTex falls short of the figure (94%) reported in [84] and the figure (95.73%) reported in Chapter 5. The number of features used are 10, 300, and 120 for the proposed approach, Haleys, and Wouwer's approach, respectively. However, significant improvement can be made by ignoring grayscale transform-invariance. Ignoring the invariance makes it fair to compare with the other methods. By doing so, we could achieve 99.31% and 93.22% for Brodatz and VisTex database, respectively. The improvement in performance is obvious, since the brightness information is lost as a result of the energy normalization procedure, which removes one of the important image properties.

In addition to improvement in viewpoint-invariant texture discrimination, two important perceptual properties, coarseness and directionality are captured by the proposed texture features. This is a very desirable characteristic in a texture classification system, since it is good to have perceptual image indices for an automated image query system.

CHAPTER 7 UNSUPERVISED TEXTURE SEGMENTATION USING FEATURE DIMENSION REDUCTION

Texture segmentation is another important texture analysis issue. One of the difficulties associated with segmenting textures is the high computation demand. In this chapter, a new and efficient method is developed to reduce the dimensions of the feature space, and thereby the computational effort is described. By reducing the amount of ambiguity among features, we obtain better performance.

7.1 Introduction

Texture segmentation involves identifying regions with uniform textures in a given image. Appropriate texture measures are needed in order to decide whether a given region has a uniform texture or otherwise. Several different texture measures for segmentation purposes are experimentally compared in [4] and [70]. Laine [41] applied the wavelet packet frame method and an envelope detection algorithm for texture segmentation. The number of channels increases at the rate of power of two along with the decomposition level, thereby increasing the computational complexity substantially. Lin *et al.* [48] have proposed a texture segmentation scheme using a tree-structured wavelet transform in which the decomposition mechanism is based upon the global energy of the transformed images, avoiding unnecessary decomposition that occurs during the packet transformation.

However, the multiscale-based segmentation approaches previously mentioned do not consider the discriminating power of each feature image. Several methods have been introduced for feature reduction and separability measure of texture features [80, 32, 63, 22]. Etemad and Chellappa [22] present unsupervised texture analysis methods for which the further wavelet

decomposition is performed based on the overall class separability instead of energy measure. Unser [80] uses the Karhunen-Loève transform to reduce the dimension of feature space that is obtained from variance estimates of wavelet frame filter outputs. Pichler *et al.* [63] apply the Gabor filtering method and use a measure of feature contrast as a tool for feature weighting and for reducing the feature space dimensions. Jain and Farrokhnia [32] also use Gabor filters and reduced the dimension of feature space by calculating energy of each filter output. The channels with largest energy are selected for a clustering procedure. However, wavelet decomposition does not guarantee that the dominant filtered images contain discriminatory information since it employs a different scheme of frequency partitioning. In many cases, lower scale channels (larger i in Fig. 3.6) are likely to have more energy, although high scale channels (smaller i) may have the most amount of discriminating information (Examples in Section 7.2.5 will demonstrate this). Hence, a reliable and efficient channel selection method is strongly recommended in the wavelet domain for more accurate and computationally efficient segmentation.

The goal in this study is to find a minimum number of measurements that can discriminate textures that are perceived to be different. The measurements should also remain approximately constant in a region when the texture is considered to be homogeneous. In an attempt to do so, a novel approach for channel selection is proposed by evaluating the histogram of each channel output. The histogram of an image with gray levels in the range $k \in [0, L - 1]$ is a discrete function $p(k) = n_k/n$, where n_k is the number of pixels in the image with that gray level, and n is the total number of pixels in the image. We use a wavelet-based multiscale filter bank, and the gray values in the filtered images transformed nonlinearly serve as texture features. A fusion of the extracted features at different scales permits a compact texture representation.

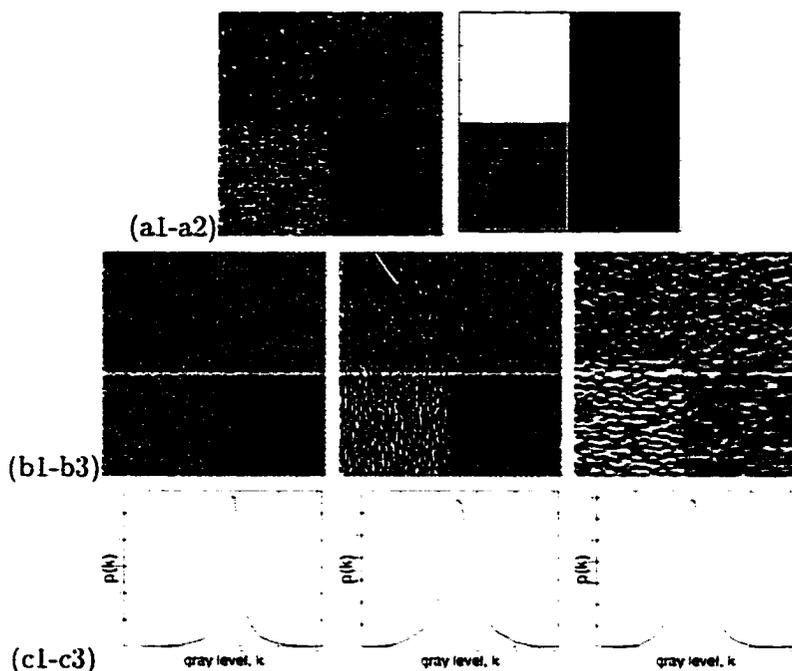


Figure 7.1 (From left to right) (a1-a2) A 128×128 artificial image and region labels; (b1-b3) channel outputs, $r_n(x, y)$, $n = 1, 3,$ and 5 , respectively; (c1-c3) image histograms of (b).

7.2 Unsupervised Texture Segmentation Algorithm in Wavelet Domain

7.2.1 Feature Extraction

Texture features that are used in this study are basically obtained from the wavelet filterbank outputs, $r_n(x, y)$, $n = 1, \dots, N$, where N is the number of channels in the filterbank [see notations in Fig. 3.6(b)]. The channel outputs are obtained by decomposing the zero-mean original image into several filtered images with limited spectral information. Four different texture images (two Gaussian Markov Random Field based, a fractal based, and a paint wall image) are used for demonstration purposes. These images were also used by Ohanian and Dubes [60] for evaluating commonly used texture feature extraction methods. These textures do not necessarily represent general texture types. They, however, provide a set of test images for evaluating algorithms. Fig. 7.1 presents an example of 2-D stationary wavelet transform (SWT) of an original image that consists of four different texture classes [Fig. 7.1(a1)] and examples of channel outputs [Fig. 7.1(b1-b3)]. Fig. 7.1(a2) shows their region labels. The

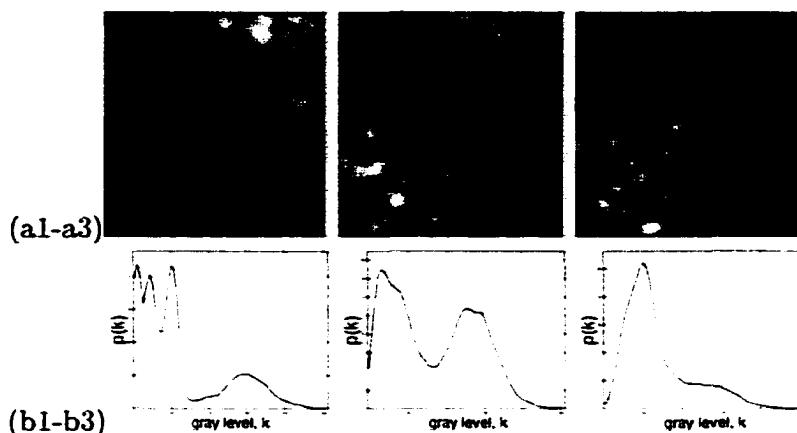


Figure 7.2 (From left to right) (a1-a3) Feature images $e_n(x, y)$ that correspond to $r_n(x, y)$ in Fig. 7.1. where $n = 1, 3,$ and $5,$ respectively: (b1-b3) image histograms of (a) with $\sigma = 2.5.$

histogram of the channel outputs shown in Fig. 7.1(c1-c3) has a Gaussian shape and, therefore the channel output itself is not a good candidate for a feature image. The goal, therefore, is to obtain a meaningful feature image from $r_n(x, y)$ such that the resulting image has better separability in terms of gray levels. A histogram of pixel values for a successful segmentation algorithm should clearly distinguish peaks corresponding to segmented regions. The necessary degree of separability for unsupervised clustering purposes can be obtained by applying a series of functions on $r_n(x, y)$, resulting in a significant change in the number of modes in the histogram. We describe our approach for transforming the histogram of a channel output such that it contains distinct peaks.

First, $r_n(x, y)$ is transformed using one of the commonly used nonlinear functions [32], $\zeta(x)$ which is defined as

$$\zeta(x) = \tanh(\alpha x) = \frac{1 - e^{-2\alpha x}}{1 + e^{-2\alpha x}}$$

followed by another nonlinear function $abs(\cdot)$. Application of $\zeta(x)$ expands the histogram and, therefore, expands the dynamic range of gray levels. Then, the feature image $e_n(x, y)$ corresponding to the filtered image $r_n(x, y)$ can be represented by

$$e_n(x, y) = \frac{1}{M^2} \sum_{(a,b) \in W_{xy}} abs\{\zeta(r_n(a, b))\}.$$

where W_{xy} is an $M \times M$ window centered at the pixel (x, y) . The window size needs to be carefully selected. A smaller window localizes texture boundaries more accurately with poorer region localization, while a larger window works better for region segmentation. The choice of window weights also affects the segmentation result. Generally, a Gaussian weighted window is found to provide better boundary localization [32]. The Gaussian function has the form

$$G(x, y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x^2 + y^2)}, \quad (7.1)$$

where σ is a constant that controls a sharpness of its lowpass filtering characteristics. This smoothing function plays a role as an envelope detector in conjunction with the previous application of nonlinear functions. It forces the pixels within each texture class to have a similar gray levels. Through these steps, a new histogram function is obtained. As shown in Fig. 7.2, we can observe a significant change in the histogram shape from $r_n(x, y)$ to $e_n(x, y)$, especially in the case of $n = 1$ and 3. This significantly improves the separability of the different texture classes in the image. The feature $e_1(x, y)$ has four distinct peaks in the histogram. The peaks correspond to the four different texture classes. A simple thresholding scheme can be applied for segmentation at this stage. The segmentation system employed in the paper is schematically summarized in the block diagram shown in Fig. 7.3.

7.2.2 Selection of parameter σ

The selection of the σ parameter is very important in order to generate a meaningful feature image. Fig. 7.4 demonstrates the effect of a parameter σ on a feature image, $e_2(x, y)$, and its histogram. As shown in the figure, when σ increases, the boundary regions become blurred, while the separability in the histogram plot becomes significant due to the increase in homogeneity within each texture category. For example, $e_2(x, y)$ becomes a meaningful feature image with a larger σ , having three distinct gray levels. Hence, σ needs to be selected such that the resulting feature image has a well-separated set of modes in the histogram. Since unsupervised clustering algorithm classifies the features largely based on distance measures, a histogram with distinctly separable peaks is required to obtain a good classification result. However, a larger value of σ does not always lead to a better segmentation result, since the

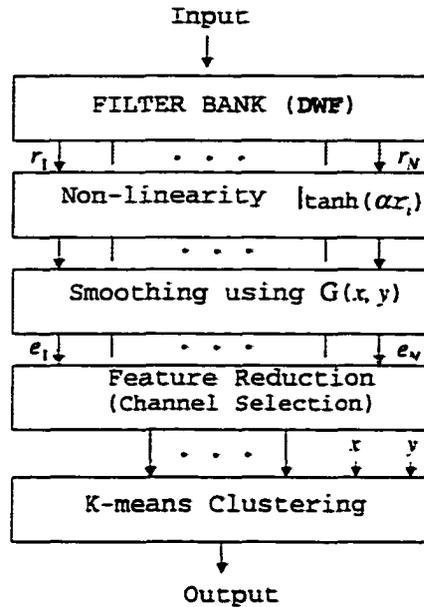


Figure 7.3 Block diagram of an image processing system for wavelet-based texture segmentation.

boundary localization has to be considered at the same time. This leads to a trade-off between choosing a good region segmentation and a good boundary segmentation. This parameter is determined during the channel selection algorithm which is described in the next section. For a Gaussian weighted window, the effect of window size is not critical.

7.2.3 Channel Selection using Histogram Analysis

It is important to reduce the dimensions of the feature space, not only because it reduces the computational effort associated with the clustering procedure, but also because it improves the classification accuracy by reducing ambiguity. It has been shown that the separation performance is not monotonically dependent on the dimensions of the feature space. Indeed, the classification accuracy decreases beyond a certain dimension [57]. We present a new method for reducing feature space dimensions by analyzing a histogram of each feature image, $e_n(x, y)$.

As described earlier, the histogram function of $r_n(x, y)$ is transformed in a manner that increases the number of modes by successively applying two nonlinear functions followed by a smoothing function on $r_n(x, y)$. We measure the separability of the feature image by ex-

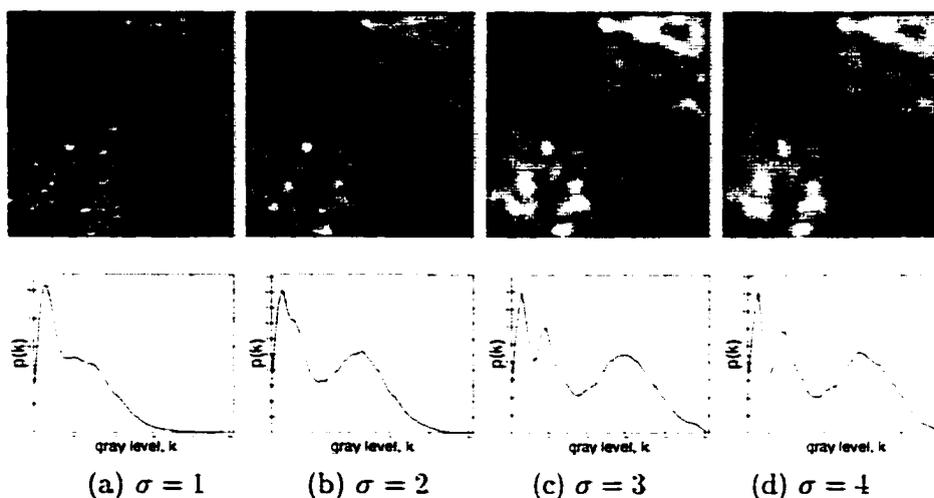


Figure 7.4 Illustration of the effect of σ for Gaussian function on the feature image and its histogram. The first row represents $e_2(x, y)$ with increasing σ (from left to right), and the second row represents the corresponding histograms.

aming distinct peaks in its histogram function. Peaks in the histogram could potentially be associated with a specific texture category. The “best” feature image would have the same number of peaks as the number of texture classes. The procedure for channel selection with a predetermined number of classes, N_c , is as follows:

1. Obtain a histogram function $p(k)$ from feature images $e_n(x, y)$, $n = 1, \dots, N$.
2. Compute the derivative of $p(k)$ for each n .
3. Count the number of zero crossings (N_{ZC}) from step 2.

If the area of derivative between two consecutive zero crossings is smaller than a constant K , then $N_{ZC} = N_{ZC} - 1$. The constant K can be considered as a threshold value for counting zero crossings from larger peaks.

4. If $N_{ZC} \geq \beta$, where β be a constant that is determined by a user, then select the channel n . If not, GOTO step 1 with $\sigma = \sigma + \gamma$, where γ represents an increment of the Gaussian function parameter σ .

Steps 2 and 3 together count the number of modes in the histogram function $p(x)$. An ideal

feature image offers a good boundary and a region localization simultaneously with $N_{ZC} = 2N_c$. As long as N_{ZC} does not increase with σ , the selection of a smaller value of σ is recommended to obtain good boundary localization. Normally, values of σ beyond 4 did not lead to identification of additional peaks in the histogram. In addition, we did not observe a significant smoothing effect for the purpose of segmentation with $\sigma < 2$. The parameter β can be determined depending upon the choice of N_{ZC} and the number of selected channels. The increment for γ can be chosen to be an arbitrarily small positive value.

7.2.4 Feature Clustering

The final step in the segmentation procedure involves the use of a clustering technique. We use the K-means algorithm [11] for clustering. Each feature is normalized with respect to its maximum value so that every input feature has the same weight. The main shortcoming of filtering-based segmentation algorithms is that they do not take into consideration the fact that neighboring pixels are very likely to belong to the same texture class. A simple method for addressing this problem is to postprocess the segmented image by employing a median filter or the relaxation method [30]. Postprocessing the images improves the degree of segmentation homogeneity at the expense of loss in precision with respect to the boundary estimates. We use a simple method that has been proposed in [32] to incorporate spatial neighborhood information directly in the clustering process. This is achieved by including the spatial coordinates of the pixels as two additional features (see Fig. 7.3).

7.2.5 Texture Segmentation

For texture segmentation, we use the Haar wavelet filter, since it is computationally efficient and it has shown its usefulness in texture analysis applications [80]. A two level 2-D SWT decomposition ($N = 7$) is used. The proposed segmentation algorithm was evaluated using two texture images [Fig. 7.1 and Fig. 7.6(a)]. The threshold K is taken as 10% of maximum area under the derivative of histogram function. The increment in σ was set at $\gamma = 0.5$, and α for $\zeta(x)$ was 0.01.

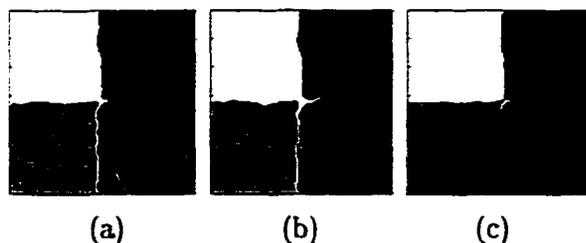


Figure 7.5 Segmentation results of an artificial image using $e_n(x, y)$: (a) $n = 1, 2, \text{ and } 3$; (b) $n = \text{all}$; (c) $n = 1$

Fig. 7.5 shows segmentation results obtained with the artificial image with and without employing the channel selection method. The channels that meet the selection criteria ($\beta = 6$) were channels 1, 2, and 3. The result obtained using the selected channels are shown in Fig. 7.5(a). We also present the result obtained using channel 1 alone for comparison purposes in Fig. 7.5(c). As observed from the results, texture segmentation obtained using the selected feature images provides better boundary localization than the one obtained using every channel [Fig. 7.5(b)].

The efficiency of the proposed algorithm was again evaluated using a composite Brodatz texture image [3] which has been commonly used as one of the standard texture images to evaluate texture segmentation techniques. The image has five different texture classes with more complex boundaries [Fig. 7.6(a)]. Channels 1, 2, 4, and 7 meet the selection criteria ($\beta = 6$) and the histograms of the channel outputs are shown in Fig. 7.6 (d1-d4). None of the single $e_n(x, y)$ has five peaks in its histogram. However, by combining the four selected feature images, good localization of texture categories [See Fig. 7.6 (b)] is obtained, although there are some misclassifications along the boundary regions as seen by different gray shade between textures. The segmentation result obtained using all channel outputs was found to be satisfactory for certain texture categories [Fig. 7.6(c)]. However, it showed more misclassifications at the boundary between textures.

Through these two test experiments, we observe that many higher scale channels (smaller n) were selected and they provide more discriminating power than lower scale channels, although they have smaller levels of energy. It is also shown that a better localization of different texture

categories can be achieved by combining the segmentation algorithm with the channel selection method.

7.3 Summary

A method for an efficient wavelet-based texture segmentation is proposed. A novel approach for wavelet channel selection is proposed in order to achieve a reduction in feature dimensions. The selection has been made by measuring the separability of the modes in the histogram of a feature image that is transformed from the channel output. The reduced feature space results in better segmentation results.

The limitation associated with the proposed feature reduction method is that the algorithm requires a set of user-defined parameters (K and β) that need to be decided *a priori*. An improvement in segmentation accuracy could be also achieved by using wavelet packets in conjunction with the channel selection method.

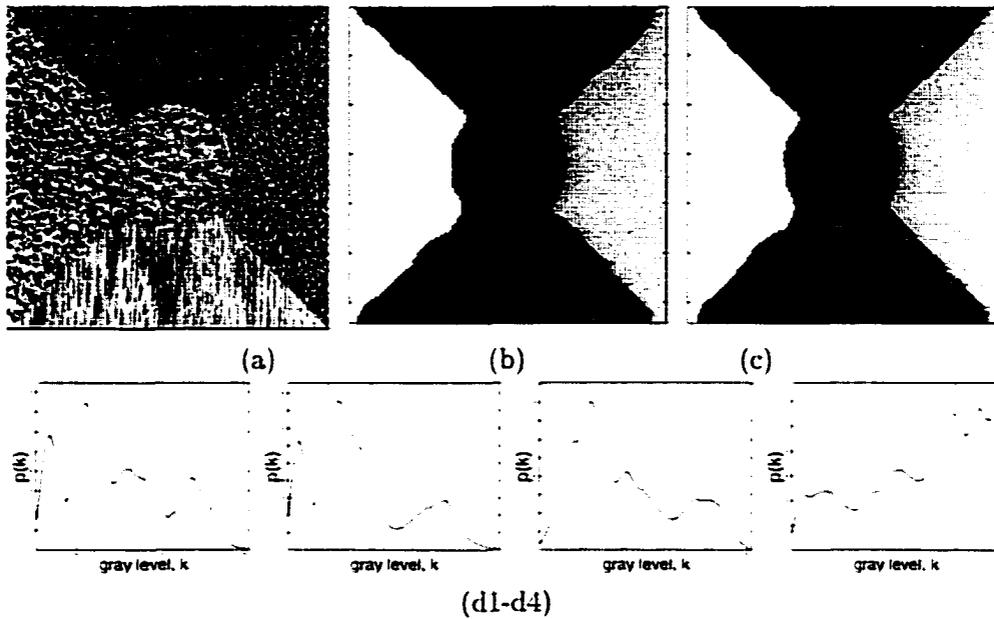


Figure 7.6 Segmentation experiment for a natural texture image. (a) original image (256×256); (b) segmentation result using selected channels ($n = 1, 2, 4, 7$); (c) segmentation result using the all channels; (d) histograms of $e_n(x, y)$ for the selected channels used for (b). (d1) $n = 1, \sigma = 2.5$; (d2) $n = 2, \sigma = 2$; (d3) $n = 4, \sigma = 2.5$; (d4) $n = 7, \sigma = 3$

CHAPTER 8 TEXTURE ANALYSIS FOR TENDON INJURY EVALUATION

In this chapter, an application of the texture segmentation algorithm to ultrasound tendon image characterization is described. The feature reduction algorithm developed in the previous chapter is employed for segmenting injured tendon regions from normal tendons. The results show that a more objective assessment of tendon injury can be obtained by using wavelet-based texture analysis techniques.

8.1 Introduction

Sports-related connective tissue injuries can be classified as macrotraumatic acute tissue destruction or microtraumatic chronic abusive cyclic loading. During strenuous exercise as in training and racing, the equine superficial digital flexor tendon (SDFT) is subjected to heavy cyclic loads. The microtrauma can lead to mechanical strain and cyclic loading can lead to degenerative changes which can result in irreversible tissue damage and a chronic tendinosis cycle with microtears and ultimately tissue failure. In humans and other animal models, the term tendinosis has been used to characterize a focal area of intratendinous degeneration that is initially asymptomatic. Histological findings include noninflammatory intratendinous collagen degeneration with fiber disorientation, hypocellularity, scattered vascular ingrowth, and occasional local necrosis or calcification [45].

Injury to the SDFT in racehorses accounts for about 8% – 12% of all racing injuries. With traditional modes of therapy, there is approximately a 40% – 50% chance that the horse will return to athletic activity after significant SDFT injury. Consequently the problem represents a major economic loss to racehorse owners. Although clinicians can see the evidence of tendon

disruption, they still do not have a method for objectively quantifying the full extent of the damage. Hence, it is necessary to have an objective assessment of both the degree of initial tendon injury and the extent of healing in order to make a correct decision as to whether, or when, a horse can resume training and racing. More importantly, if minor injuries (degenerative changes) can be detected in their early stages, the incidence of clinical tendinitis in racehorses can presumably be reduced by altering the management of these horses.

The value of ultrasonic imaging in detecting tendon injury has been well established in equine veterinary medicine [69]. Real-time medical ultrasound scanners provide a two-dimensional B-mode image of the tissue along the plane of sound wave propagation. Histologic examination of normal tendons demonstrates the axial alignment of fiber bundles [69]. This parallelism of fibers is a consistent finding in midline sagittal (longitudinal or parallel to the length of the tendon) ultrasound scans of the tendons of the metacarpus and metatarsus of normal horses. The cross-sectional images of the tendons show fine dot patterns since sound beam interacts with the fiber bundles approximately perpendicularly. In the dorsal plane, the echogenicity of the tendon structure is mottled where the circular-shaped reflectors are related to bundles of collagen fibers in transverse section [69]. In the sagittal plane, echogenic lines corresponding to longitudinal sections of bundles of collagen fibers give each tendon and ligament a striated echogenic pattern [69],[77]. With tendinosis, there is collagen degeneration with fiber disorientation and disruption in alignment. At present, sonographic demonstration of injuries and the assessment of healing are based largely upon a qualitative interpretation of the image. Typically, tendon lesions are graded on a scale of 1 to 4. A grade 1 correlates with only minor damage; a grade 4 lesion correlates with major damage.

In an attempt to provide a semi-quantitative evaluation, Reef *et al.* [71] measure the length of the lesion and the damaged cross-sectional area. Genovese *et al.* [25] link the cross-sectional area and a qualitative grading of lesion echogenicity and express them as a severity grading. All these techniques, however, still require a subjective visual assessment of the area of damaged tendon. Nicoll *et al.* [59] evaluate the merits of measuring the mean gray-level of the image area for obtaining a more quantitative assessment. However, there is a strong potential for extract-

ing additional information from ultrasound images using texture analysis algorithms. Texture analysis has shown its usefulness in many medical applications, for example, in discriminating normal from abnormal liver [85],[58] and in characterizing intramuscular fat percentage in ribeye muscle of cattle [37].

In this chapter, we apply the texture segmentation algorithm developed in the previous chapter to ultrasound B-scan tendon images for objective assessment of changes in texture that would help clinician assess the equine tendon injury. In addition, energy ratio features will be obtained from DWT subimages and used in characterizing tendon fiber alignment as an additional injury index.

8.2 Tendon Image Acquisition and Clinical Evaluation

Ultrasound scans were obtained from horses that injured their tendons while racing at a local racetrack in the 1996 season. Images were collected from 24 horses with superficial digital flexor tendon (SDFT) injury. Of these 24 horses, 7 had bilateral tendon injuries, 5 injured the right front SDFT, and 12 injured the left front SDFT. The injury was diagnosed on clinical evaluation including subjective evaluation of ultrasound images by an experienced clinician. Gray-scale digital images of B-mode ultrasound were acquired using a veterinary ultrasound machine, the PIE Medical Scanner 200 (Classic Medical Supply Inc., Tequesta, Florida) with a 7.5 MHz linear array transducer.

Tendon scans were performed at six different sites on the leg (zones 1A - 3B in Fig. 8.1) in order to provide a representative sample of the entire SDFT. The collected image size is 4 cm \times 7 cm (309 \times 426 pixels). The sites chosen included sites of predilection for injury based upon previous experience. Zones 2A, 2B, and 3A are common sites of SDFT injury. Since the deep digital flexor tendon (DDFT) is rarely injured, it is used as a normal reference image. A cross-sectional view (with a transducer perpendicular to long axis of the tendon or leg) and a longitudinal view (with a transducer parallel to long axis of the tendon or leg) were obtained at each site. Whenever an injured tendon was scanned, the opposite clinically normal tendon was also scanned to facilitate a comparison, and to minimize the effects of biological variation

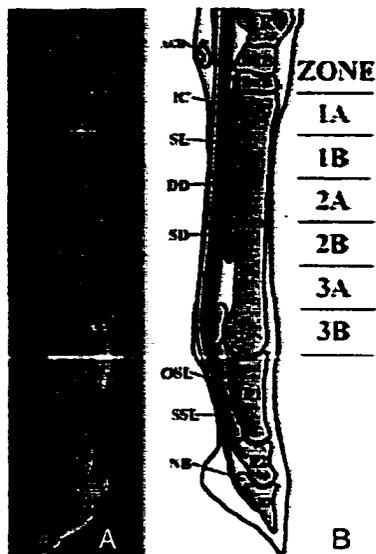


Figure 8.1 Sagittal view of the metacarpus. A. Frozen section. B. Schematic of frozen section showing zones of scanning. DD = deep digital flexor tendon; SD = superficial digital flexor tendon.

between horses. Some of these horses were scanned more than once to evaluate subsequent healing that took place over several months. For each ultrasound image, the SDFT and DDFT were traced by an experienced ultrasonographer for further analysis. The whole area of the SDFT and DDFT was traced and selected from the cross-sectional image. A rectangular area in the longitudinal view containing adequate texture information was selected for further analysis.

The SDFT and DDFT are identified in images shown in Fig. 8.2. The upper region of interest (ROI) represents the SDFT and the lower ROI represents the DDFT. The darker area within the cross-sectional image of the SDFT is the lesion and has different textural characteristics relative to the normal area. For the longitudinal view image, the amount of fiber alignment is measured from the selected ROI as an injury index. The normal tendon, in general, has parallel horizontal lines in the SDFT as mentioned in Section 8.1. The injured tendon is characterized with disruptions of the alignment in the longitudinal view images. Subjective grades were assigned at the time of ultrasound examination for each image. The subjective fiber alignment was graded as : 1 if > 75% of the fibers in the lesion are parallel, 2 if

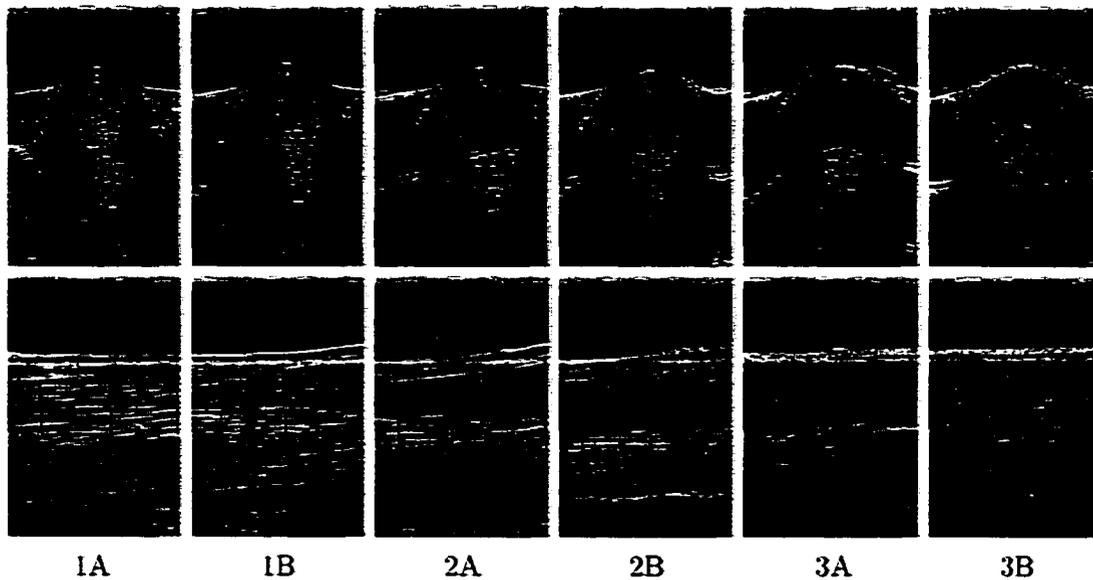


Figure 8.2 Typical ultrasound images of equine leg showing SDFT and DDFT. Transducer is at top edge of images. Actual image size is 4 cm wide and 7 cm deep (309 × 426 pixels). First row: Cross-sectional images from zone 1A through 3B. The subjective grades are 1, 1, 3, 3, 2, and 2, respectively. Second row: Corresponding longitudinal view images. The fiber alignment grades are 1, 1, 3, 3, 3, and 2. The upper regions of interest (ROI) represent the SDFT and the lower ROIs represent the DDFT. The transducer is at the top edge of the images.

> 50 but \leq 75% of the fibers in the lesion are parallel. 3 if > 25% but \leq 50% of the fibers are parallel, and 4 if \leq 25% of the fibers were parallel. However, since the grades are somewhat subjective, they may differ between examiners. Hence, an objective assessment of tendon injury for both cross-sectional and longitudinal images would provide better information about the injury status. However, a true objective reference of tendon injury (e.g., biopsy or histology of a tendon slice or possibly a high resolution MRI image) was not available for the horses scanned. Therefore, the objective of the present study is to develop an “objective” algorithm that will at best match the accuracy of visual classification by an experienced clinician. Fig. 8.3 shows the procedures of our experiments for two types of images.

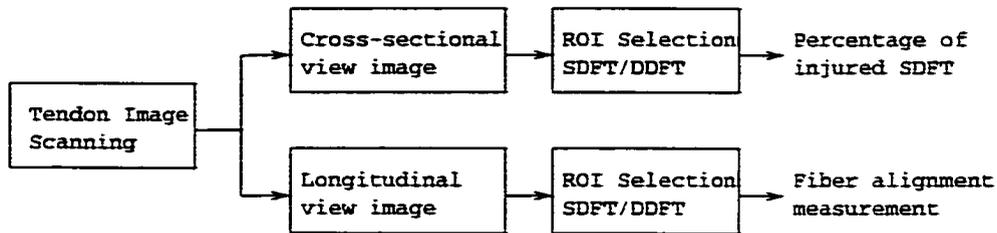


Figure 8.3 Injury evaluation procedure for two types of B-scan tendon images.

8.2.1 Lesion Segmentation for Cross-sectional Images

One of the important tasks for evaluating cross-sectional images is the detection and measurement of the amount of injured SDFT that lies at the outer tendon region. The injured tendon does not reflect ultrasound in the way the normal tendon does, which makes the texture property of the two regions different. The subjective injury grades are largely based on the visual appearance of the B-scan images, and they are not consistent. To provide more objective data on the amount of SDFT damage, a wavelet packet frame method was applied without employing the feature selection method [38]. The dimensions of the feature space needs to be very large to obtain a reasonably good texture differentiation using the packet-based method. We apply the texture segmentation procedure proposed in Section III to differentiate the damaged lesion and to quantify the percentage of the damaged SDFT. The rectangular shaped ROI that contains the SDFT is used as an input to the segmentation system.

8.2.2 Ultrasonic Fiber Measurement for Longitudinal Images

For the longitudinal image, fiber alignment in the SDFT was characterized using energy-based texture parameters obtained from a SWFB. The relations between the energy from different channels provide a measure of differences in dominant sizes and orientations of texture pattern [37]. Since the main focus in analyzing the longitudinal image is the quantification of horizontal line patterns, some of the energy ratio (ER) parameters used in [37] were adopted in this experiment. The texture parameters used for measuring fiber alignment were calculated by first computing the ER parameters from the SDFT and DDFT of the same B-scan image

using

$$ER_i = E(A_{LH}^i)/E(A_{HL}^i), i = 1, 2, \dots, I, \quad (8.1)$$

where E denotes the energy defined in equation (5.1), A_{LH}^i and A_{HL}^i denote a horizontally emphasized and a vertically emphasized subimage, respectively at scale level i . The energy ratio of the two channels characterizes the relative amount of directional (horizontal/vertical) components in the original image. The ER of SDFT is then divided by that of DDFT and the resulting parameter is denoted as normalized energy ratio ($NER = ER_{SDFT}/ER_{DDFT}$). Since the DDFT is rarely injured, it is a good reference for the purpose of normalization.

8.3 Results

We present the injured lesion segmentation results obtained by using the proposed approach in Chapter 7 to analyze cross-sectional images, as well as fiber alignment measurement results obtained by analyzing longitudinal view images.

A redundant wavelet filterbank with ten channels ($I = 3$) was used for decomposing cross-sectional SDFT images. The SDFT images in Fig. 8.2 are presented again in Fig. 8.4 (a1-a6). We applied the channel selection method to ten feature images of each SDFT image for obtaining feature reduction. The number of texture classes associated with the tendon images is two (normal/abnormal), and we set the parameter β at 4. Although there are variations in the channel numbers selected for each SDFT image, channel 10 was selected in all cases, which allows low frequency components associated with the two textures to be differentiated. In addition, channel 6 was selected to ensure that horizontal lines that are visible in the normal tendon area are well emphasized.

The lesion segmentation results are presented in Fig. 8.4 (b1-b6) and the percentage of the injured SDFT is summarized in Table 8.1. The numbers in parentheses are the subjective grades. The SDFT damaged percentage has good correlation with the subjective grades except for the zone 3A. This may be attributed to either subjectiveness of the visual evaluation or misclassification of the proposed segmentation algorithm.

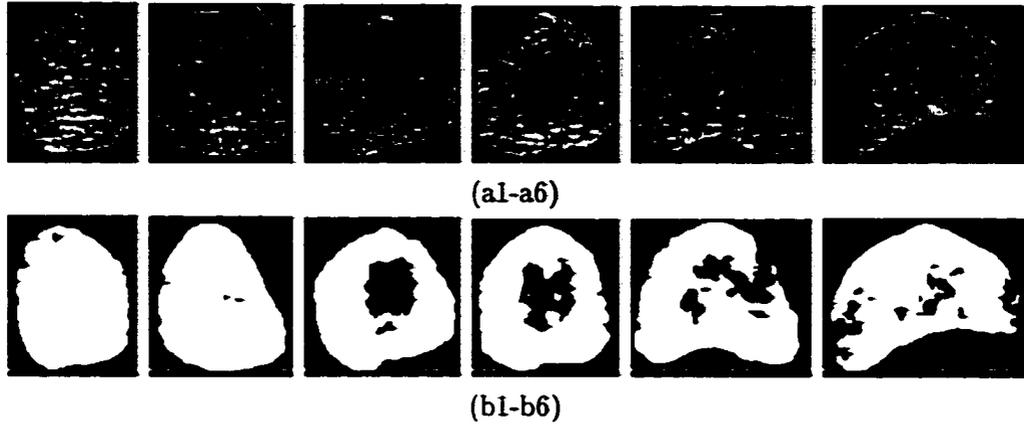


Figure 8.4 Lesion segmentation results. (a1-a6) The SDFTs from cross-sectional images in Fig. 8.2: (b1-b6) the corresponding results indicating the percentage of the injured SDFT.

Table 8.1 Percentage of damaged SDFT and the fiber alignment measure for the ROIs in fig. 8.2. The numbers in parentheses are the subjective grades.

Zone	SDFT damage percentage	NER_1
1A	2.23 (1)	0.9116 (1)
1B	1.08 (1)	0.9889 (1)
2A	14.09 (3)	0.3213 (3)
2B	13.87 (3)	0.3789 (3)
3A	14.03 (2)	0.5399 (3)
3B	8.87 (2)	0.8329 (2)

For the analysis of the longitudinal image, we use a seven channel SWFB, and we measure the parameter $NER_i, i = 1, 2$. Eighty two images that were selected from the database depending on the number of available images for each subjective grade were analyzed. The NER_i parameter was calculated and tested for correlation with four different grades of injury assigned by a clinician. Although the subjective grading may not be a good basis for evaluating results, it provides an insight into the usefulness of the texture parameters. NER_2 was found to be poorly correlated with the grades. However, NER_1 was found to be highly correlated with the grades, which implies that the information relating to the horizontal fiber lines are captured in the first scale. In other words, the frequency range of the vertical variations of the horizontal fiber lines lies between $[\pi/2, \pi]$.

Table 8.2 Results of fiber alignment grading for longitudinal image analysis. *standard deviation

Subjective grade	No. of images	Average NER_1	$\pm\sigma^*$
1	24	0.97	0.15
2	29	0.63	0.10
3	20	0.43	0.07
4	8	0.31	0.05

NER_1 values for the images in Fig. 8.2(b) are shown in Table. 8.1. The SDFT damage percentage and NER_1 parameter in the table are negatively correlated with fairly high correlation. since the injured tendon normally tends to have a larger damaged tendon area and a smaller value of parallel fiber alignment index. As we observe, the subject grading for cross-sectional and longitudinal images are not consistent for zone 3A. Either subjectiveness of the grading or the different range of image scanning area could have caused this problem.

The result of the test obtained with 82 scans using NER_1 is summarized in Table 8.2. The mean value in each category does not exactly fall in the range of the current grading criteria. However, the NER_1 parameter provides a tool for objective assessment of fiber alignment in the longitudinal images. In summary, the parameters obtained from the two methods provided great potential for quantitative measurements of tendon injury.

8.4 Summary

The wavelet-based texture analysis algorithm was applied to ultrasound images of racing horses for evaluating tendon injury. A measure of the percentage of injured SDFT is obtained by applying the texture segmentation algorithm to cross-sectional images. Similarly, a measure of the fiber alignment is obtained from the longitudinal images using a directionally sensitive energy ratio parameter. The results agree well with subjective grades estimated by an experienced clinician for both cross-sectional and longitudinal images. Of course, this potential technique will have to be thoroughly evaluated against an objective reference, such as biopsy or histology of a tendon slice or possibly a high resolution MRI image. Eventually, this approach may help the veterinarian in objectively quantifying image textures which could then

potentially aid in evaluating progress of injury or healing process.

It is also desirable to develop methods for combining the quantitative results obtained from both types of image analysis.

CHAPTER 9 CONCLUSIONS

The main focus of this dissertation is on filtering-based texture analysis, using the wavelet filterbank approach. Efficient algorithms are developed for the classification and segmentation of texture images. The viewpoint-invariance problem, which is one of the most challenging tasks, is tackled by introducing a dual wavelet filterbank that consists of a standard wavelet filterbank (SWFB) and a rotated wavelet filterbank (RWFB). The RWFB is developed in order to achieve better characterization of oriented texture patterns and is obtained by rotating the 2-D standard wavelet filters by 45° . The classification performance is evaluated using a large number of natural texture images. The performance is compared with results obtained by using other successful filtering-based approaches [84, 28].

An efficient method for reducing a dimensions of the feature space is proposed for texture segmentation problems where the computational complexity is one of the main issues. The feature reduction is obtained by analyzing histograms of feature images. The proposed segmentation algorithm has been applied to ultrasound tendon images for an objective assessment of racehorse tendon injury.

The major contributions of this dissertation, the major findings from the experiments, and the list of suggestions for future research are described in the following sections.

9.1 Major Contributions

The major contributions of this dissertation can be summarized as follows:

1. A new oriented wavelet filterbank, RWFB, that is more efficient than SWFB in characterizing oriented texture patterns is introduced. The oriented filters are non-separable and retain the orthogonality and inversibility of the standard wavelet filterbank. The texture

classification results obtained in Chapter 5 show that RWFB provides complementary information to the SWFB, resulting in an improvement of classification accuracy. The performance of the combined feature set is better than most of the successful wavelet-based texture approaches [84].

2. A dual wavelet filterbank (DWFB) is introduced in order to obtain a set of rotation-invariant texture features. The features seem to have a physical connection with human perception. The “coarseness” and “directionality” properties of a texture seem to be captured by a few of the proposed features. These features are first introduced in this study (except the F_{SSE} feature). The features offer results that are superior to a recently proposed filter-based approach [28] for rotation-invariant texture classification, even though we employ significantly smaller number of features. This implies that the proposed feature set is better representative of textures than others.
3. Directionality feature is invariant with respect to texture rotation. In addition, computing this feature in different scales provides directionality of a texture in multi-resolution. This multi-resolution features are particularly useful for a texture that contains different amount of directional patterns in multiple directions. To the best of our knowledge, this is the first multi-channel/multi-resolucional filtering approach that has been proposed to obtain the directionality of a texture image.
4. The rotation-invariant texture features are also grayscale transform-invariant. This is a significant accomplishment, since it reduces an effort for a pre-processing procedure that needs to be performed for normalization.
5. Although we have not conducted a complete comparative study, two well accepted filtering approaches are used to facilitate a partial comparison.
6. A channel selection algorithm has been presented for achieving feature reduction in texture segmentation problems. The selection process improves the segmentation performance and reduces the computational complexity. The channel selection is made by analyzing histogram functions of the channel output.

9.2 Major Findings

Major findings observed from the classification experiments that were conducted can be summarized as follows:

1. It is still very difficult to represent textures completely using a limited number of parameters. The incompleteness arises partially due to the fact that a number of texture samples used in the training and testing phases are not considered homogeneous.
2. It is observed that the l_1 -norm based energy measure represents textures more efficiently than the l_2 -norm based energy measure. This has not been reported in literature yet. and warrants a thorough investigation using different texture databases.
3. Another consistent finding we have observed is the superior performance of Haar wavelet basis over the higher-order Daubechies wavelet basis. although the Haar wavelet basis is not recommended for applications such as denoising and compression. A justification needs to be provided

9.3 Suggestions for Further Research

We have introduced DWFB for obtaining an effective set of texture features that are useful for both general and viewpoint-invariant texture classification problems. Experimental results obtained show that the proposed approach is very promising in terms of effectiveness, simplicity, and robustness. Suggestions for future work include are:

1. The rotation-invariant texture features obtained from DWFB can be also obtained from other oriented wavelet transforms, such as Gabor filter set, Hexagonal wavelet transform [75], and a steerable wavelet transform [24] at the expense of frequency and orientation fine-tuning. A study comparing these oriented wavelet transform needs to be conducted.
2. Current work on DWFB has so far focused on algorithm development and experimental justification. A more thorough theoretical analysis needs to be carried out. In addition, applications such as denoising and image enhancement can be explored using DWFB.

3. The limitation of the texture features obtained from DWFb is that they cannot capture the structure of texture primitives, as in most filtering-based feature extraction methods. Hence, “regularity” measure was not obtained using the proposed filterbank scheme. This is due to the fact that the energy parameter captures the global information within the filtered images. However, a measure of second-order statistics, such as a covariance matrix or a set of GLCM features, captures structural information on the wavelet coefficients. This information in multiple scales may serve as a measure of regularity in multiresolution.
4. As we have observed from texture classification experiments, the peaking phenomenon is significant. A systematic approach needs to be carried out to obtain an optimal set of features for a given texture data set.
5. The texture segmentation algorithm proposed in Chapter 7 can benefit from a systematic approach for obtaining parameters for channel selection. This will render the user-independent approach.
6. Finally, theoretical research needs to be carried out for defining textures. Surveys in characteristics of all possible filters and the amount of textural information that can be obtained using each of the filters would significantly benefit the texture analysis community and assist it in selecting a transform domain for a given task to be solved.

BIBLIOGRAPHY

- [1] J. Beck, A. Sutter, and R. Ivry. Spatial frequency channels and perceptual grouping in texture segregation. *CVGIP*, 37:299–325, 1987.
- [2] A. C. Bovik, M. Clark, and W. S. Geisler. Multichannel texture analysis using localized spatial filters. *IEEE Trans. Patt. Anal. and Machine Intell.*, 12:55–73, 1990.
- [3] P. Brodatz. *Textures: A Photographic Album for Artists and Designers*. Dover Publications, New York, 1966.
- [4] J. M. H. D. Buf, M. Kardan, and M. Spann. Texture feature performance for image segmentation. *Pattern Recognition*, 23(3/4):291–309, 1990.
- [5] T. Chang and C. J. Kuo. Texture analysis and classification with tree-structures wavelet transform. *IEEE Trans. Image Process.*, 2(4):429–441, 1993.
- [6] B. Chaudhuri, N. Sarkar, and P. Kundu. Improved fractal geometry based texture segmentation technique. *IEE Proceedings part E*, 140:233–241, 1993.
- [7] R. Chellappa and S. Chatterjee. Classification of textures using gaussian markov random fields. *IEEE Trans. ASSP*, ASSP-33(4):959–963, 1985.
- [8] J. L. Chen and A. Kundu. Rotation and gray scale transform invariant texture classification using wavelet decomposition and hidden markov model. *IEEE Trans. Image Proc.*, 16(2):208–214, 1994.
- [9] F. S. Cohen, Z. Fan, and M. A. Patel. Classification of rotated and scaled textured image using gaussian markov random field models. *IEEE Trans. Pattern Anal. Machine Intell.*, 13(2):192–202, 1991.

- [10] R. R. Coifman and M. V. Wickerhauser. Entropy-based algorithms for best basis selection. *IEEE Trans. Info. Theory*, 38(2):713–718. 1992.
- [11] G. B. Coleman and H. C. Andrews. Image segmentation by clustering. *Proc. of the IEEE*, 67(5):773–785. 1979.
- [12] R. W. Connors and C. A. Harlow. A theoretical comparison of texture algorithms. *IEEE Trans. PAMI*. PAMI-2(3):204–222. 1980.
- [13] I. J. Cox. F. T. Leighton. and T. Shamoan. Secure spread spectrum watermarking for multimedia. *IEEE Trans. Image Process.*, 6:1673–1687. 1997.
- [14] G. Cross and A. Jain. Markov random field texture models. *IEEE Trans. Pattern Anal. Machine Intell.*, 5:25–39. 1983.
- [15] I. Daubechies. The wavelet transform, time-frequency localization and signal analysis. *IEEE Trans. Info. Theory*. (9):961–1005. 1990.
- [16] J. G. Daugman. Two-dimensional spectral analysis of cortical receptive field profile. *Vision Res.*, 20:847–856. 1980.
- [17] J. G. Daugman. Uncertainty relation for resolution in space, spatial-frequency, and orientation optimized by two-dimensional visual cortical filters. *J. Opt. Soc. Am.*, 2:1160–1169. 1985.
- [18] H. Derin and H. Elliott. Modeling and segmentation of noisy and textured images using gibbs random fields. *IEEE Trans. Patt. Anal. and Machine Intell.*, 9:39–55. 1987.
- [19] R. C. Dubes and A. K. Jain. Random field models in image analysis. *J. Appl. Statist.*, 6:131–164. 1989.
- [20] D. Dunn and W. E. Higgins. Optimal gabor filters for texture segmentation. *IEEE Trans. Image Process.*, 4(7), 1995.
- [21] K. Etemad and R. Chellappa. Separability based tree structured local basis selection for texture classification. *Proc. IEEE Int'l Conf. Image Process.*, 3:441–445, 1994.

- [22] K. Etemad and R. Chellappa. Separability-based multiscale basis selection and feature extraction for signal and image classification. *IEEE Trans. Image Process.*, 7(10):1453–1465, 1998.
- [23] O. Faugeras and W. Pratt. Decorrelation methods of texture feature extraction. *IEEE Trans. Pattern Anal. Machine Intell.*, 2:323–332, 1980.
- [24] W. T. Freeman and E. H. Adelson. The design and use of steerable filters. *IEEE Trans. Patt. Anal. and Machine Intell.*, 13(9):891–906, 1991.
- [25] R. Genovese and N. Rantanen. Clinical experience with quantitative analysis of superficial digital flexor tendon injuries in thoroughbred and standardbred racehorses. *Vet. Clinics N. Am. Equine Pract.*, 6:129–145, 1990.
- [26] R. C. Gonzalez and R. E. Woods. *Digital Image Processing*. Addison Wesley Publishing Co., 1992.
- [27] T. O. Gulrud and S. Kjode. Optimal filter for detection of stellate lesions and circumscribed masses in mammograms. *Proc. SPIE's Visual Comm. and Image Process.*, pages 430–440. Orlando, Florida, Mar. 1996.
- [28] G. M. Haley and B. S. Manjunath. Rotation-invariant texture classification using a complete space-frequency model. *IEEE Trans. Image Process.*, 8(2):255–269, 1999.
- [29] R. M. Haralick, K. Shanmugam, and I. Dinstein. Textural features for image classification. *IEEE Trans. System Man Cybernet.*, SMC-3(6):610–621, 1973.
- [30] J. Y. Hsiao and A. A. Sawchuk. Supervised textured image segmentation using feature smoothing and probabilistic relaxation techniques. *IEEE Trans. Patt. Anal. and Machine Intell.*, 11(12):1279–1292, 1989.
- [31] A. K. Jain and B. Chandrasekaran. *Dimensionality and sample size considerations in pattern recognition practice*. Chapter 39, North-Holland, Amsterdam, 1982.

- [32] A. K. Jain and F. Farrokhnia. Unsupervised texture segmentation using gabor filters. *Pattern Recognition*, 24(12):1167–1186, 1991.
- [33] A. K. Jain and B. Yu. Automatic text location in images and video frames. *Pattern Recognition*, 31(12):2055–2076, 1998.
- [34] M. S. Kankanhalli and R. K. Ramakrishnan. Content based watermarking of images. *Proc. of ACM Multimedia*, pages 61–70, Bristol, UK, 1998.
- [35] R. L. Kashyap, R. Chellappa, and A. Khotanzad. Texture classification using features derived from random field models. *Pattern Recognition Letters*, 1(1):43–50, 1982.
- [36] R. L. Kashyap and A. Khotanzad. A model-based method for rotation invariant texture classification. *IEEE Trans. Pattern Anal. Machine Intell.*, 8(7):472–481, 1986.
- [37] N.-D. Kim, V. Amin, D. Wilson, G. Rouse, and S. Udpa. Ultrasound image texture analysis for characterizing intramuscular fat content of live beef cattle. *Ultrasonic Imaging*, 20:191–205, 1998.
- [38] N.-D. Kim, L. Booth, V. Amin, J. Lim, and S. Udpa. Ultrasound image processing for tendon injury evaluation. *Proc. IEEE-SP Int'l symposium on TFTS*, (Pittsburgh, PA):241–244, 1998.
- [39] J. E. Koss, F. D. Newman, T. K. Johnson, and D. L. Kirch. Abdominal organ segmentation using texture transforms and a hopfield neural network. *IEEE Trans Medical Imaging*, 18(7):640 – 648, 1999.
- [40] A. Laine and J. Fan. Texture classification by wavelet packet signatures. *IEEE Trans. Patt. Anal. and Machine Intell.*, 15(11):1186–1191, 1993.
- [41] A. F. Laine and J. Fan. Frame representations for texture segmentation. *IEEE Trans. Image Process.*, 5(5):771–780, 1996.
- [42] A. F. Laine, S. Schuler, J. Fan, and W. Huda. Mammographic feature enhancement by multiscale analysis. *IEEE Trans. Medical Imaging*, 13(4):725–740, 1994.

- [43] N. Lamei, K. D. Hutchinson, M. M. Crawford, and N. Khazenie. Cloud-type discrimination via multispectral textural analysis. *Optical Engineering*, 33(4):1303–1313, 1994.
- [44] K. I. Law. Rapid texture identification. *SPIE Image Processing for Missile Guidance*, 238:376–380, 1980.
- [45] W. B. Leadbetter. Cell-matrix response in tendon injury. *Clinical Sports Medicine*, 11:533, 1992.
- [46] M.-C. Lee and C.-M. Pun. Texture classification using dominant wavelet packet energy features. *Proc. 4th IEEE Southwest Symp. on Image Anal. Interpretation*, pages 301–304, 2000.
- [47] H.-C. Lin, L.-L. Wang, and S.-N. Yang. Regular-texture image retrieval based on texture-primitive extraction. *Image and Vision Computing*, 17:51–63, 1999.
- [48] Y. Lin, T. Chang, and C. C. Kuo. Texture segmentation using wavelet packets. *SPIE Mathematical Imaging*, 2034:277–286, 1993.
- [49] L. Linnett, S. Clarke, C. Graham, and D. Langhorne. Remote sensing of the seabed using fractal techniques. *Elect. Comm. Eng. J.*, 3(5):195–203, 1991.
- [50] F. Liu and W. Picard. Periodicity, directionality, and randomness: Wold features for image modeling and retrieval. *IEEE Trans. Patt. Anal. and Machine Intell.*, 18(7):722–733, 1996.
- [51] S. Livens, P. Scheunders, D. W. G. Van, and D. D. Van. Wavelets for texture analysis, an overview. *IEE Conf. Image Proc. and Anal.* Dublin, July:581–585, 1997.
- [52] J. Malik and P. Perona. Preattentive texture discrimination with early vision mechanisms. *J. Opt. Soc. Amer.*, 7:923–932, 1990.
- [53] S. Mallat. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans. Patt. Anal. and Machine Intell.*, 11(7):674–693, 1989.

- [54] S. Mallat. Wavelets for vision. *Proc. of the IEEE*, 84(4):604–614. 1996.
- [55] B. S. Manjunath and R. Chellappa. Unsupervised texture segmentation using markov random field models. *IEEE Trans. PAMI*, 13(5):478–482. 1991.
- [56] B. S. Manjunath and W. Y. Ma. Texture features for browsing and retrieval of image data. *IEEE Trans. PAMI, Special Issue Digital Libraries*, 18:837–842. 1996.
- [57] J. Mao and A. K. Jain. Texture classification and segmentation using multiresolution simultaneous autoregressive models. *Pattern Recognition*, 25:173–188. 1992.
- [58] A. Mojsilović, M. Popović, S. Marković, and M. Krstić. Characterization of visually similar diffuse diseases from b-scan liver images using nonseparable wavelet transform. *IEEE Trans. Medical Imaging*, 17(4):541–549. 1998.
- [59] R. Nicoll, A. Wood, and R. Rothwell. Ultrasonographical and pathological studies of equine superficial digital flexor tendons: initial observations, including tissue characterization by analysis of image gray scale, in a thoroughbred gelding. *Equine Vet. J.*, 24(4):318–320. 1992.
- [60] P. Ohanian and R. Dubes. Performance evaluation for four classes of textural features. *Pattern Recognition*, 25:819–833. 1992.
- [61] T. Ojala, M. Pietikinen, and D. Harwood. A comparative study of texture measures with classification based on feature distributions. *Pattern Recognition*, 29(1):51–59. 1996.
- [62] A. Pentland. Fractal-based description of natural scenes. *IEEE Trans. Patt. Anal. and Machine Intell.*, 6(6):661–674. 1984.
- [63] O. Pichler, A. Teuner, and B. J. Hosticka. An unsupervised texture segmentation algorithm with feature space reduction and knowledge feedback. *IEEE Trans. Image Process.*, 7(1):53–61. 1998.
- [64] I. Pitas and C. Kotropoulos. A texture-based approach to the segmentation of seismic images. *Pattern Recognition*, 25(9):929–945. 1992.

- [65] D. A. Pollen and S. F. Ronner. Visual cortical neurons as localized spatial frequency filters. *IEEE Trans. SMC*, 13:907–916, 1983.
- [66] M. Porat and Y. Y. Zeevi. The generalized scheme of image representation in biological and machine vision. *IEEE Trans. Patt. Anal. and Machine Intell.*, 10(7):452–468, 1988.
- [67] R. Porter and N. Canagarajah. Robust rotation-invariant texture classification: wavelet, gabor filter and gmrf based schemes. *IEE Proc.-Vis. Image Process.*, 144(3):180–188, 1997.
- [68] T. Randen and J. H. Husøy. Filtering for texture classification: A comparative study. *IEEE Trans. PAMI*, 21(4):291–310, 1999.
- [69] N. Rantanen, M. Hause, and R. Genovese. Superficial digital flexor tendinitis: diagnosis using real-time ultrasound imaging. *J. Equine Vet. Sci.*, 5:115–119, 1985.
- [70] T. R. Reed and J. M. H. du Buf. A review of recent texture segmentation and parameter extraction techniques. *CVGIP:Image Understanding*, 57:359–372, 1993.
- [71] V. Reef, B. Martin, and A. Elser. Types of tendon and ligament injuries detected with diagnostic ultrasound: detection and follow-up. *Proc. Am. Ass. Equine Pract.*, 34:245–248, 1993.
- [72] S. Santini and R. Jain. Similarity measures. *IEEE Trans. PAMI*, 21(9):871–883, 1999.
- [73] J. M. Shapiro. Embedded image coding using zerotrees of wavelet coefficients. *IEEE Trans. Signal Process.*, 41(12):3445–3462, 1993.
- [74] Signal and L. A. Image Processing Institute. University of Southern California. (url: <http://sipi.usc.edu/services/database/Database.html>).
- [75] E. P. Simoncelli and E. H. Adelson. Non-separable extensions of quadrature mirror filters to multiple dimensions. *Proceedings of the IEEE: Special issue on Multi-dimensional Signal Processing*, 78(4):652–664, 1990.

- [76] G. Smith and I. Burns. Measuring texture classification algorithms. *Pattern Recognition Letters*, 18:1495–1501, 1997.
- [77] K. Spaulding. Ultrasonic anatomy of the tendons and ligments in the distal metacarpal-metatarsal region of the equine limb. *Vet. Radiol.*, 25:265–273, 1984.
- [78] T. N. Tan. Rotation invariant texture features and their use in automatic script identification. *IEEE Trans. PAMI*, 20(7):751–756, 1998.
- [79] M. Tuceryan and A. K. Jain. Texture analysis. *Handbook of Pattern Recognition and Computer Vision*. C.H. Chen, L.F. Pau, and P.S.P. Wang, Eds.(Singapore:World Scientific):235–276, 1993.
- [80] M. Unser. Texture classification and segmentation using wavelet frames. *IEEE Trans. Image Process.*, 4(11):1549–1560, 1995.
- [81] VisTex. Vistex texture database. <http://www-white.media.mit.edu/vismod/imagery/visiontexture/vistex.html>. *MIT Media Lab*.
- [82] J. S. Weszka, C. R. Dyer, and A. Rosenfeld. A comparative study of texture measures for terrain classification. *IEEE Trans. SMC*, SMC-6(4):269–285, 1976.
- [83] B. Wilson and M. A. Bayoumi. Compressed domain texture classification from a modified ezw symbol stream. *Proc. Data Compression Conference*, page 586, 2000.
- [84] G. V. Wouwer, P. Scheunders, and D. V. Dyck. Statistical texture classification from discrete wavelet representations. *IEEE Trans. Image Process.*, 8(4):592–598, 1999.
- [85] C. Wu, Y. Chen, and K. Hsieh. Texture features for classifications of ultrasonic liver images. *IEEE Trans. Medical Imaging*, 11:141–152, 1992.
- [86] V. Wu, R. Manmatha, and E. M. Riseman. Textfinder:an automatic system to detect and recognize text in images. *IEEE Trans. PAMI*, 21(11):1224–1229, 1999.

- [87] W. R. Wu and S. C. Wei. Rotation and gray-scale transform invariant texture classification using spiral resampling, subband decomposition, and hidden markov model. *IEEE Trans. Image Process.*, 5(10):1423–1434. 1996.
- [88] B. Zhu, M. Ramsey, and H. Chen. Creating a large-scale content-based airport image digital library. *IEEE Trans. Image Process.*, 9(1):163–167. 2000.