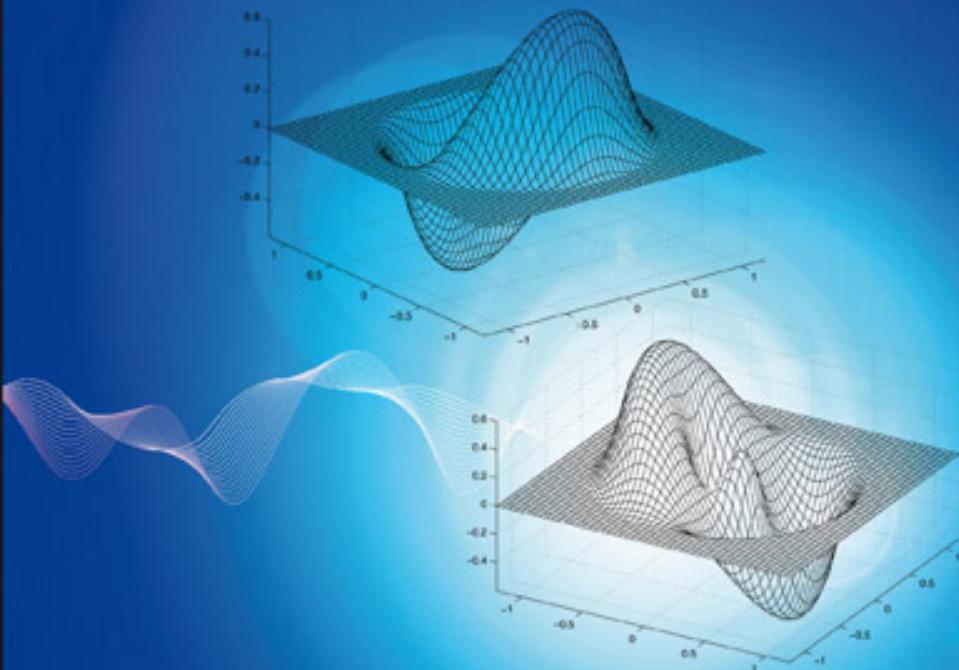


WAVELET THEORY APPROACH TO PATTERN RECOGNITION

2nd Edition

Yuan Yan Tang



SERIES IN
MACHINE PERCEPTION
ARTIFICIAL INTELLIGENCE
Volume 74

**WAVELET THEORY
APPROACH TO
PATTERN RECOGNITION**

2nd Edition

SERIES IN MACHINE PERCEPTION AND ARTIFICIAL INTELLIGENCE*

Editors: **H. Bunke** (Univ. Bern, Switzerland)
P. S. P. Wang (Northeastern Univ., USA)

- Vol. 58: Computational Web Intelligence: Intelligent Technology for Web Applications
(Eds. *Y. Zhang, A. Kandel, T. Y. Lin and Y. Yao*)
- Vol. 59: Fuzzy Neural Network Theory and Application
(*P. Liu and H. Li*)
- Vol. 60: Robust Range Image Registration Using Genetic Algorithms and the Surface Interpenetration Measure
(*L. Silva, O. R. P. Bellon and K. L. Boyer*)
- Vol. 61: Decomposition Methodology for Knowledge Discovery and Data Mining: Theory and Applications
(*O. Maimon and L. Rokach*)
- Vol. 62: Graph-Theoretic Techniques for Web Content Mining
(*A. Schenker, H. Bunke, M. Last and A. Kandel*)
- Vol. 63: Computational Intelligence in Software Quality Assurance
(*S. Dick and A. Kandel*)
- Vol. 64: The Dissimilarity Representation for Pattern Recognition: Foundations and Applications
(*Elżbieta Pękalska and Robert P. W. Duin*)
- Vol. 65: Fighting Terror in Cyberspace
(Eds. *M. Last and A. Kandel*)
- Vol. 66: Formal Models, Languages and Applications
(Eds. *K. G. Subramanian, K. Rangarajan and M. Mukund*)
- Vol. 67: Image Pattern Recognition: Synthesis and Analysis in Biometrics
(Eds. *S. N. Yanushkevich, P. S. P. Wang, M. L. Gavrilova and S. N. Srihari*)
- Vol. 68: Bridging the Gap Between Graph Edit Distance and Kernel Machines
(*M. Neuhaus and H. Bunke*)
- Vol. 69: Data Mining with Decision Trees: Theory and Applications
(*L. Rokach and O. Maimon*)
- Vol. 70: Personalization Techniques and Recommender Systems
(Eds. *G. Uchygigit and M. Ma*)
- Vol. 71: Recognition of Whiteboard Notes: Online, Offline and Combination
(Eds. *H. Bunke and M. Liwicki*)
- Vol. 72: Kernels for Structured Data
(*T Gärtner*)
- Vol. 73: Progress in Computer Vision and Image Analysis
(Eds. *H. Bunke, J. J. Villanueva, G. Sánchez and X. Otazu*)

*For the complete list of titles in this series, please write to the Publisher.

WAVELET THEORY APPROACH TO PATTERN RECOGNITION

2nd Edition

Yuan Yan Tang

Chongqing University, China

and

Hong Kong Baptist University, Hong Kong



World Scientific

Published by

World Scientific Publishing Co. Pte. Ltd.

5 Toh Tuck Link, Singapore 596224

USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601

UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library.

WAVELET THEORY APPROACH TO PATTERN RECOGNITION

(2nd Edition)

Series in Machine Perception and Artificial Intelligence — Vol. 74

Copyright © 2009 by World Scientific Publishing Co. Pte. Ltd.

All rights reserved. This book, or parts thereof, may not be reproduced in any form or by any means, electronic or mechanical, including photocopying, recording or any information storage and retrieval system now known or to be invented, without written permission from the Publisher.

For photocopying of material in this volume, please pay a copying fee through the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, USA. In this case permission to photocopy is not required from the publisher.

ISBN-13 978-981-4273-95-4

ISBN-10 981-4273-95-3

Printed in Singapore.

I dedicate this book to my parents, wife,
daughter, brother and sisters

This page intentionally left blank

Preface

This book is an update of the book “Wavelet Theory and its Application to Pattern Recognition” which was published in 2000. Three new chapters are added to this new book. The objective is to attack a challenging research topic that is related to both areas of wavelet theory and pattern recognition.

Wavelet analysis and its applications have become one of the fastest growing research areas in recent years. This is in part attributed to the pioneering work by the researchers as well as practitioners in the fields of mathematics and signal processing. Wavelet theory has been employed in many fields and applications, such as signal and image processing, communication systems, biomedical imaging, radar, air acoustics, theoretical mathematics, control system, and endless other areas. However, the research on applying the wavelets to pattern recognition is still too weak; only a few publications deal with this topic at the present. This book focuses on this challenging research topic.

The most fascinating area of signal/image processing with practical applications is pattern recognition. Making computers see and recognize objects like humans has captured the attention of many scientists in different disciplines. Indeed, machine recognition of different patterns such as printed and handwritten characters, fingerprints, biomedical images, etc. has been intensively and extensively researched by scientists in different countries around the world. The area of pattern recognition, after over five decades of continued development, is now definitely playing a very major role in advanced automation in the 21st century. Although a lot of achievements have been made in the area of pattern recognition, many problems

still have to be solved. The goal of this book is to, through mathematically sound derivations and experiments, develop some new application-oriented techniques in wavelet theory, and thereafter, apply these new techniques to solve some particular problems in the area of pattern recognition.

This book is organized into two groups, namely, Chapters 1 - 4 extend wavelet theory, while Chapter 5 through Chapter 12 deal with the application of the wavelet theory to pattern recognition, which is the core of this book. Considering the major readers of this book are scientists and engineers, thus, in some chapters/sections, we give up the exactness in mathematics temporarily.

Initially, in Chapter 1, a brief description of wavelet theory is introduced, and a comparison between the wavelet and Fourier transforms is discussed along with several pictorial examples. This chapter reviews established applications of the wavelet theory to pattern recognition. The review is not detailed, since this book concentrates on the novel research results developed by the author. However, the references are cited if additional details are desired.

Throughout Chapters 2 and 3, both the wavelet transform and wavelet bases are of critical concerns, which formulate the basic wavelet theory. In Chapter 2, the general theory of the continuous wavelet transform is addressed, and its major properties are investigated including the characterization of Lipschitz regularity of signals by the wavelet transform. Chapter 2 also primarily examines an important property by relating the processing to matched filtering concepts. Chapter 3 considers multiresolution analysis (MRA) and wavelet bases, where the basic concepts of the both are presented as well as the construction of them. As an important algorithm for implementing the discrete wavelet transform, Mallat algorithm is introduced.

After studying these basic concepts of the wavelet theory, some typical wavelet bases including the orthonormal and nonorthonormal bases are provided in Chapter 4. They benefit the application of the wavelet theory to the engineering, such as pattern recognition, image processing, etc.

By formulating the above wavelet theory and the general applications with wavelet theory, the second group of this book (Chapters 5 - 12) demonstrates more detailed applications, which become the core chapters. All of these applications were made by our research group.

Chapter 5 develops a method to identify different structures of the edges and design an algorithm to detect the step-structure edges. This technique

can be employed to contour extraction in document processing as well as 2-D object recognition.

Chapter 6 aims at studying the characterization of Dirac-structure edges with wavelet transform, and selecting the suitable wavelet functions to detect the Dirac edges. A mapping technique is applied in this chapter to construct such a wavelet function. In this way, a low-pass function is mapped onto a wavelet function by a derivation operation. In this chapter, the quadratic spline wavelet is utilized to characterize the Dirac-structure edges and an algorithm to extract the Dirac-structure edges by wavelet transform is also developed.

Chapter 7 introduces a new wavelet function called Tang-Yang wavelet, which is constructed by our research group. The characteristics of the Tang-Yang wavelet with curves are discussed. They are grey-level invariant, slope invariant and width invariant. The application of new wavelet function to curve analysis is presented.

In Chapter 8, skeletonization of Ribbon-like shapes based on the Tang-Yang wavelet function is presented. Characterization of the Ribbon-like shape with wavelet transform using the Tang-Yang wavelet is investigated. Some useful algorithms are also provided.

Chapter 9 presents an approach to feature extraction. In this way, the wavelet decomposition is used to produce wavelet sub-patterns, and thereafter, the fractal divider dimensions are utilized to find the numerical features from these sub-patterns.

Chapter 10 applies 2-D multiresolution analysis (MRA) and Mallat algorithm to form document analysis. The HL and LH sub-images are utilized to find the reference lines in a form document, furthermore, the useful information can be extracted in accordance with these reference lines. This application is verified by several concrete examples of bank checks.

Chapter 11 uses B-spline wavelet for Chinese computing, which consists of three operations, namely, (1) compression of Chinese characters, (2) enlargement of type size with arbitrary scales, and (3) generation of type styles of Chinese fonts.

Finally, Chapter 12 deals with the classification of patterns with wavelet theory, where the orthogonal wavelet series are used for the probability density estimation in the classifier design.

The main components of this book are the achievements in our research group with visiting research scholars. The professors from various universities and the graduate students at Hong Kong Baptist University have

made contributions to this book. Actually, they are co-authors of this book: Professors Jiming Liu and P. C. Yuen at Hong Kong Baptist University, Professor Seong-Whan Lee at Korea University of Korea, Professor Lihua Yan at Zhongshan (Sun Yat-Sen) University of China, Professor Ching Y. Suen at Concordia University of Canada, Professor Xinge You at Huazhong University of Science and Technology of China, Professor Z. K. Chen at Chongqing University of China, Professors Hong Ma and Bing-Fa Li at Sichuan University of China, Professor Feng Yang at the South Medical University of China.

To review the established applications of the wavelet theory to pattern recognition, in Chapter 1, some materials including figures from the published papers are quoted. I would like to thank the following authors to release the copyrights to this book: S. H. Yoon, J. H. Kim, W. E. Alexander, S. M. Park and K. H. Sohn [Yoon et al., 1998], F. Murtagh and J.-L. Starck [Murtagh and Starck, 1998], K. H. Liang, F. Chang, T. M. Tan and W. L. Hwang [Liang et al., 1999].

A specific international journal called “International Journal on Wavelets, Multiresolution, and Information Processing (IJWMIP)” was founded by myself in 2003. In Chapter 1, the following papers from IJWMIP are quoted. I would like to record my appreciation to the authors for their contributions to this book: J. Daugman [Daugman, 2003], L. H. Yang, T. D. Bui and C. Y. Suen [Yang et al., 2003b], A. Z. Kouzani and S. H. Ong [Kouzani and Ong, 2003], S. Kumar, D. K. Kumar, A. Sharma and N. McLachlan [Kumar et al., 2003], S. Kumar and D. K. Kumar [Kumar and Kumar, 2005], A. Sharnia, D. K. Kumar and S. Kumar [Sharnia et al., 2004], R. S. Kunte and R. D. S. Samuel [Kunte and Samuel, 2007], K. Muneeswaran, L. Ganesan, S. Arumugam and P. Harinarayan [Muneeswaran et al., 2005], R. Ksantini, D. Ziou, F. Dubeau and P. Hari-narayan [Ksantini et al., 2006], S. E. El-Khamy, M. M. Hadhoud, M. I. Dessouky, B. M. Salam and F. E. A. El-Samie [El-Khamy et al., 2006].

I would like to express my deep gratitude to the students, Y. W. Man, Limin Cui, Yan Tang, D. H. Xi and L. Feng.

I would like to also thank the many people at Hong Kong Baptist University, especially the Department of Computer Science for their strong support.

Many research projects involved in this book were supported by the research grants received from Research Grant Council (RGC) of Hong Kong and Faculty Research Grant (FRG) of Hong Kong Baptist University. This

work was also supported by the National Natural Science Foundation of China.

This page intentionally left blank

Contents

Preface	vii
Chapter 1 Introduction	1
1.1 Wavelet: A Novel Mathematical Tool for Pattern Recognition	1
1.2 Brief Review of Pattern Recognition with Wavelet Theory	18
1.2.1 Iris Pattern Recognition	21
1.2.2 Face Recognition Using Wavelet Transform	22
1.2.3 Hand Gestures Classification	28
1.2.4 Classification and Clustering	31
1.2.5 Document Analysis with Wavelets	37
1.2.6 Analysis and Detection of Singularities with Wavelets .	43
1.2.7 Wavelet Descriptors for Shapes of the Objects	46
1.2.8 Invariant Representation of Patterns	49
1.2.9 Handwritten and Printed Character Recognition	56
1.2.10 Texture Analysis and Classification	61
1.2.11 Image Indexing and Retrieval	67
1.2.12 Wavelet-Based Image Fusion	70
1.2.13 Others	72
Chapter 2 Continuous Wavelet Transforms	75
2.1 General Theory of Continuous Wavelet Transforms	75
2.2 The Continuous Wavelet Transform as a Filter	91
2.3 Characterization of Lipschitz Regularity of Signal by Wavelet .	94
2.4 Some Examples of Basic Wavelets	98

Chapter 3 Multiresolution Analysis and Wavelet Bases	105
3.1 Multiresolution Analysis	105
3.1.1 Basic Concept of Multiresolution Analysis (MRA)	105
3.1.2 The Solution of Two-Scale Equation	109
3.2 The Construction of MRAs	120
3.2.1 The Biorthonormal MRA	127
3.2.2 Examples of Constructing MRA	133
3.3 The Construction of Biorthonormal Wavelet Bases	143
3.4 S. Mallat Algorithms	148
Chapter 4 Some Typical Wavelet Bases	153
4.1 Orthonormal Wavelet Bases	156
4.1.1 Haar Wavelet	156
4.1.2 Littlewood-Paley (LP) Wavelet	158
4.1.3 Meyer Wavelet	160
4.1.4 Battle-Lemaré-spline Wavelet	162
4.1.5 Daubechies' Compactly Supported Orthonormal Wavelets	166
4.1.6 Coiflet	178
4.2 Nonorthonormal Wavelet Bases	181
4.2.1 Cardinal Spline Wavelet	181
4.2.2 Compactly Supported Spline Wavelet	196
Chapter 5 Step-Edge Detection by Wavelet Transform	201
5.1 Edge Detection with Local Maximal Modulus of Wavelet Transform	202
5.2 Calculation of $W_s f(x)$ and $W_s f(x, y)$	211
5.2.1 Calculation of $W_s f(x)$	213
5.2.2 Calculation of $W_s f(x, y)$	215
5.3 Wavelet Transform for Contour Extraction and Background Removal	220
5.3.1 Basic Edge Structures	225
5.3.2 Analysis of the Basic Edge Structures with Wavelet Transform	228
5.3.3 Scale-Independent Algorithm	239
5.3.4 Experiments	242

Chapter 6 Characterization of Dirac-Edges with Quadratic Spline Wavelet Transform	251
6.1 Selection of Wavelet Functions by Derivation	252
6.1.1 Scale Wavelet Transform	253
6.1.2 Construction of Wavelet Function by Derivation of the Low-Pass Function	255
6.2 Characterization of Dirac-Structure Edges by Wavelet Transform	258
6.2.1 Slope Invariant	263
6.2.2 Grey-Level Invariant	265
6.2.3 Width Light-Dependent	267
6.3 Experiments	272
Chapter 7 Construction of New Wavelet Function and Application to Curve Analysis	279
7.1 Construction of New Wavelet Function — Tang-Yang Wavelet	282
7.2 Characterization of Curves through New Wavelet Transform . .	285
7.3 Comparison with Other Wavelets	295
7.3.1 Comparison with Gaussian Wavelets	296
7.3.2 Comparison with Quadratic Spline Wavelets	300
7.4 Algorithm and Experiments	301
7.4.1 Algorithm	301
7.4.2 Experiments	305
Chapter 8 Skeletonization of Ribbon-like Shapes with New Wavelet Function	311
8.1 Tang-Yang Wavelet Function	315
8.2 Characterization of the Boundary of a Shape by Wavelet Transform	317
8.3 Wavelet Skeletons and Its Implementation	322
8.3.1 Wavelet Transform in the Discrete Domain	322
8.3.2 Generation of Wavelet Skeleton in the Discrete Domain	325
8.3.3 Modification of Primary Wavelet Skeleton	327
8.4 Algorithm and Experiment	331
8.4.1 Algorithm	331
8.4.2 Experiments	332

Chapter 9 Feature Extraction by Wavelet Sub-Patterns and Divider Dimensions	343
9.1 Dimensionality Reduction of Two-Dimensional Patterns with Ring-Projection	346
9.2 Wavelet Orthonormal Decomposition to Produce Sub-Patterns	351
9.3 Wavelet-Fractal Scheme	358
9.3.1 Basic Concepts of Fractal Dimension	359
9.3.2 The Divider Dimension of One-Dimensional Patterns	365
9.4 Experiments	367
9.4.1 Experimental Procedure	367
9.4.2 Experimental Results	368
Chapter 10 Document Analysis by Reference Line Detection with 2-D Wavelet Transform	381
10.1 Two-Dimensional MRA and Mallat Algorithm	383
10.2 Detection of Reference Line from Sub-Images by the MRA	387
10.3 Experiments	390
Chapter 11 Chinese Character Processing with B-Spline Wavelet Transform	399
11.1 Compression of Chinese Character	404
11.1.1 Algorithm 1 (Global Approach)	404
11.1.2 Algorithm 2 (Local Approach)	406
11.1.3 Experiments	407
11.2 Enlargement of Type Size with Arbitrary Scale Based on Wavelet Transform	408
11.2.1 Algorithms	409
11.2.2 Experiments	414
11.3 Generation of Chinese Type Style Based on Wavelet Transform	415
11.3.1 Modification	416
11.3.2 Composition	424
Chapter 12 Classifier Design Based on Orthogonal Wavelet Series	429
12.1 Fundamentals	429
12.2 Minimum Average Loss Classifier Design	433
12.3 Minimum Error-Probability Classifier Design	435

12.4 Probability Density Estimation Based on Orthogonal Wavelet Series	437
12.4.1 Kernel Estimation of a Density Function	437
12.4.2 Orthogonal Series Probability Density Estimators	439
12.4.3 Orthogonal Wavelet Series Density Estimators	441
List of Symbols	449
Bibliography	451
Index	461

This page intentionally left blank

Chapter 1

Introduction

1.1 Wavelet: A Novel Mathematical Tool for Pattern Recognition

Wavelet analysis is a relatively recent development of applied mathematics in 1980s. Independent from its developments in harmonic analysis, A. Grossmann, J. Morlet and their coworkers studied the wavelet transform in its continuous form and initially applied it to analyze geological data [Grossmann and Morlet, 1984; Grossmann and Morlet, 1985; Grossmann et al., 1985; Morlet et al., 1982a; Morlet et al., 1982b]. However, at that time, the roughness of wavelets made the mathematicians suspect the existence of a “good” wavelet basis until two great events took place in 1988, namely:

- Daubechies, a female mathematician, constructed a class of wavelet bases, which are smooth, compactly supported and orthonormal. They are referred to as Daubechies bases, and successfully applied to many fields today.
- French signal analysis expert, S. Mallat, with mathematician, Y. Meyer, proposed a general method to construct wavelet bases. It is termed multiresolution analysis (MRA) and is intrinsically consistent with sub-band coding in signal analysis.

The above achievements play important roles in mathematics as well as engineering. They established the theoretical frame for wavelet analysis in mathematics. On the other hand, they improved the traditional theory of frequency analysis in engineering. As a result, a novel method

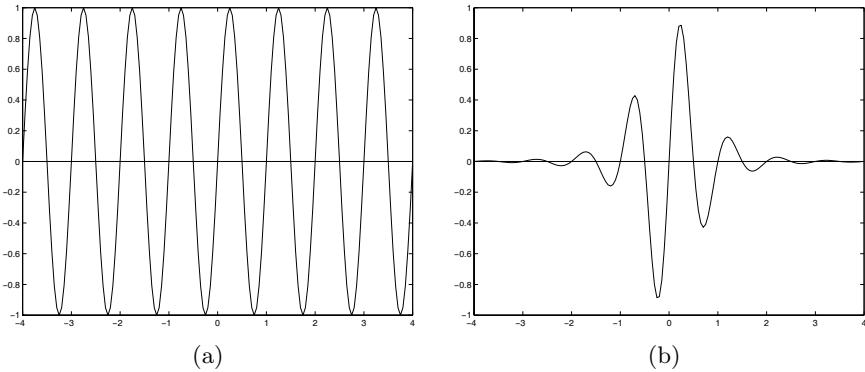


Fig. 1.1 Simple examples of wave and wavelet. (a) A wave is produced by a non-damped simple harmonic vibration; (b) A wavelet is formed by a damped oscillation.

for time-frequency analysis has been formed. In this way, signals can be locally characterized in both the time domain and frequency domain simultaneously and self-adaptively. According to this property, wavelet analysis can be efficiently applied to analyze and process the non-stationary signals. The engineers and scientists have paid ample attention to wavelet theory, and carried out the research of wavelets on very wide areas, in recent years. As a matter of fact, significant success of wavelet analysis has achieved on a variety of disciplines, such as, signal processing, image compressing and enhancement, pattern recognition, communication systems, control systems, biomedical imaging, air acoustic, radar, theoretical mathematics, and endless other signal processing fields. Moreover, the potentiality of wavelets in both the research and development would be immeasurable.

What is a wavelet? The simplest answer is a “short” wave (wave + let \Rightarrow wavelet). The suffix “let” initially means “a small kind of” in a general English dictionary [Procter, 1993]. However, in the mathematical term “wavelet”, the meaning of the suffix “let” comes “short”, which indicates that the duration of the function is very limited. In other words, it is said that wavelets are localized. Look at the following simple examples in Fig. 1.1, to obtain a brief idea: what is a wave, and what is a wavelet.

We consider two typical mechanical movements:

- non-damped simple harmonic vibration, and
- damped oscillation.

A wave can be produced by a non-damped simple harmonic vibration, which is illustrated in Fig. 1.1(a). It forms a sinusoidal wave with non-attenuated amplitude. By contrast, a wavelet can be formed by a damped oscillation, as shown in Fig. 1.1(b), its amplitudes quickly decay to zero in both the positive and negative directions. In other words, the wavelet is a special signal, in which, two conditions have to be satisfied.

First, the wavelet must be oscillatory (wave).

Second, its amplitudes are nonzero only during a short interval (short).

The required oscillatory condition leads to sinusoids as the building blocks. The quick decay condition is a tapering or windowing operation. These two conditions must be simultaneously satisfied for the function to be a wavelet [Young, 1993]. According to these conditions, we can justify whether a function is a wavelet.

Consider the functions in Fig. 1.2. As applying these two conditions, it is clear that the functions in Fig. 1.2(a) and (b) are wavelets, while the ones in Fig. 1.2(c) and (d) are not. The functions in Fig. 1.2(a) and (b) are oscillatory and have amplitudes which quickly decay to zero with time. Thus, they are wavelets. The functions in Fig. 1.2(c) and (d) do not satisfy both wave and short simultaneously. Fig. 1.2(c) decays quickly to zero but does not oscillate, while Fig. 1.2(d) is wave but not short. Therefore, they are not wavelets.

The next question is what we can do with this simple wave? It is well-known that any complex movement, which is described by $f(x)$, can be represented by the summation of several simple harmonic vibrations with different frequencies:

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx), \quad (1.1)$$

where, $\cos kx$ and $\sin kx$ denote the simple harmonic vibrations.

J. Fourier first discovered this important fact in 1807, when he was studying the equation of heat conduction. This is why the above series expansion is called Fourier expansion. It infers that any complex wave $f(x)$ can be formed by the linear combination of the basic elements, which are produced by the dilations and translations of $\sin x$. $\sin x$ is the simplest basic wave and is viewed as a basic “atom”. Therefore, through the investigation of rules, which indicate how a complex function $f(x)$ can be

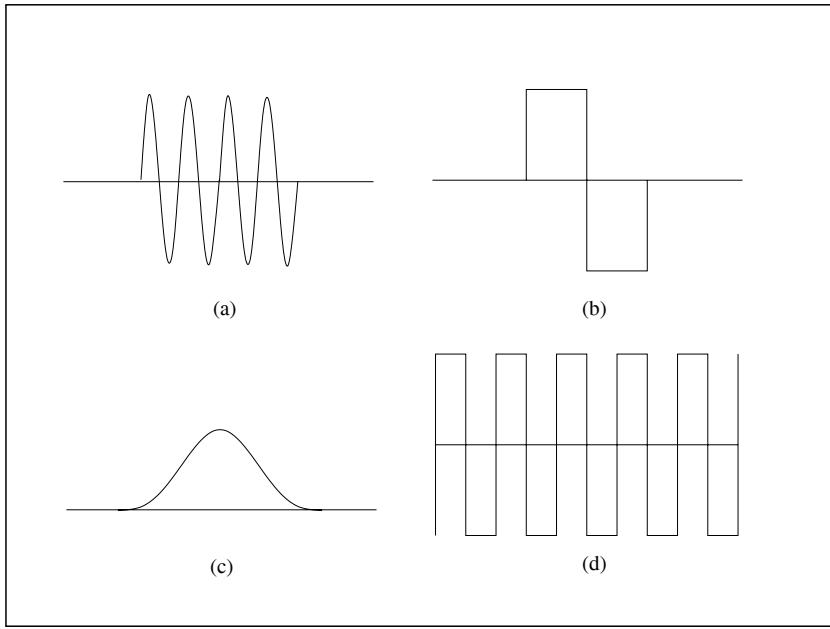


Fig. 1.2 Simple examples of wave and wavelet. (a) and (b) are wavelets, since they are oscillatory and in limited duration; (c) and (d) are not wavelets, because the function in (c) decays quickly to zero but does not oscillate, the function in (d) is wave but not short.

composed of such small atoms, we can understand function $f(x)$ itself well. It has been mathematically proved that the sum of the finite number of terms, which are the fore ones in Fourier series, is the best approximation of $f(x)$ under the consideration of energy. That is:

$$\begin{aligned} & \int_0^{2\pi} |f(x) - [\frac{a_0}{2} + \sum_{k=1}^N (a_k \cos kx + b_k \sin kx)]|^2 dx \\ &= \min_{c_0, c_1, d_1, \dots, c_N, d_N} \int_0^{2\pi} |f(x) - [\frac{c_0}{2} + \sum_{k=1}^N (c_k \cos kx + d_k \sin kx)]|^2 dx \rightarrow 0 \end{aligned} \quad (N \rightarrow \infty). \quad (1.2)$$

For these reasons, Fourier analysis has been a traditional and efficient tool in many fields of science and engineering, in past two hundred years. It will also play an indispensable role in the future. However, the development of science and engineering has no limits. Fourier analysis is not a solution that will be good for all time and in all situations. Meanwhile, Fourier analysis has its own deficiency. In fact, it has two major problems, namely,

- Fourier analysis can not characterize the signals locally in time domain.
- Fourier expansion can approximate the stationary signals well, but can not do so for the non-stationary ones.

These drawbacks can be found in the following concrete examples. Consider two signals shown in Fig. 1.3.

According to the above analysis, it can be known that a complicated function $f(x)$ can be represented by a series of Fourier coefficients, say $\{a_0, a_1, b_1, a - 2, b_2, \dots\}$, i.e.

$$f(x) \implies \{a_0, a_1, b_1, \dots, a_i, b_i, \dots, a_j, b_j, \dots\}.$$

Where, $j > i$, and the Fourier coefficients a_j and b_j correspond to the components with higher frequency in signal $f(x)$. By contrast, the coefficients a_i and b_i correspond to the components with lower frequency in signal $f(x)$. For the given signals $f_1(x)$ and $f_2(x)$, we have

$$\begin{aligned} f_1(x) &\implies \{a_0^1, a_1^1, b_1^1, \dots, a_i^1, b_i^1, \dots, a_j^1, b_j^1, \dots\} \\ f_2(x) &\implies \{a_0^2, a_1^2, b_1^2, \dots, a_i^2, b_i^2, \dots, a_j^2, b_j^2, \dots\} \end{aligned} \tag{1.3}$$

It is easy to understand, from Figs. 1.3(c) and (d), that the values of the first several terms, which correspond to lower frequencies, in the Fourier coefficients are greater than those of the other terms, which are located in the tail of the series and have higher frequencies. The backer the coefficients are located, the closer to zero the values are. Thus, the energy of a signal is distributed on the front terms (lower frequencies) in the Fourier coefficients.

Consider the number of cycles in Figs. 1.3(c) and (d), it can be found that the number of cycles in Fig. 1.3(d) is greater than that in Figs. 1.3(c), which indicates that signal $f_2(x)$ has more nonzero components in the frequency domain than that $f_1(x)$ has. This situation is also described in

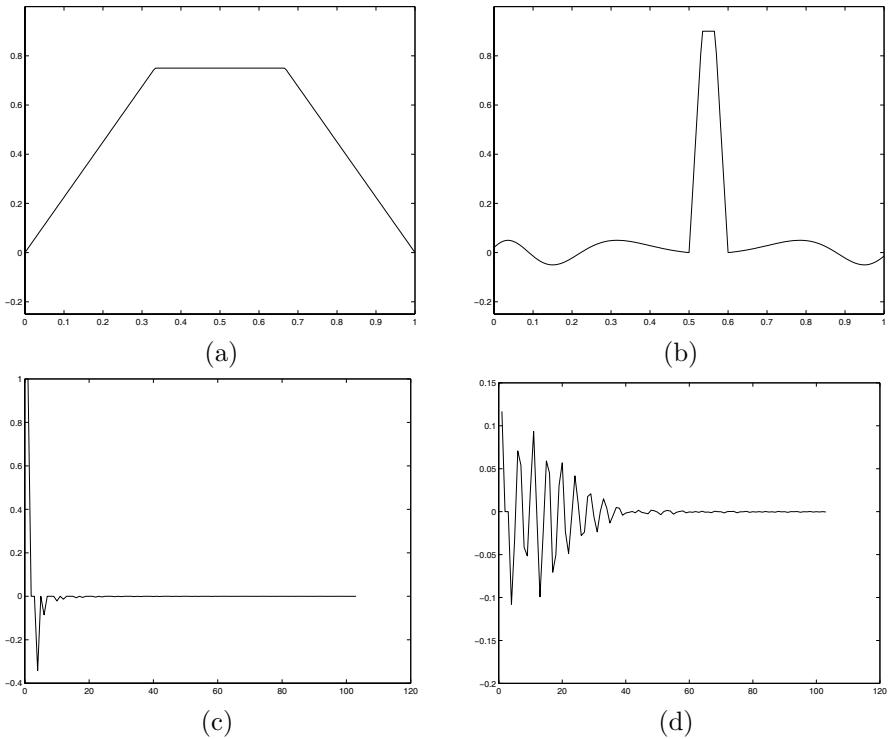


Fig. 1.3 The first row: the left side is the original signal $f_1(x)$, and the right side displays another original signal $f_2(x)$. The second row: the left side indicates the Fourier coefficients corresponding to $f_1(x)$, and the right side shows the Fourier coefficients corresponding to $f_2(x)$.

Eq. 1.3, where the number of nonzero terms in the Fourier expansion for signal $f_2(x)$ is more than that in the expansion for signal $f_1(x)$. This property arrests the time-frequency analysis using the Fourier theory. For instance, signal $f_2(x)$ possesses a transient component with very short interval. However, its corresponding Fourier expansion contains many terms which produce a long-duration vibration as illustrated in Fig. 1.3(d). Thus, a short-duration transient in the time domain corresponds to a long-duration vibration in the frequency domain. The characteristic of localization of the narrow transient signal in the time domain will disappear in the frequency domain. Therefore, it is impossible to locate and characterize the transient

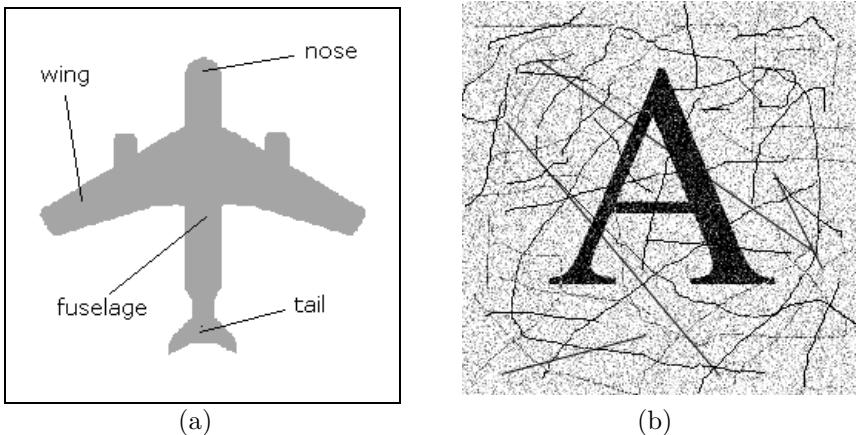


Fig. 1.4 The contours of objects, drawing lines and texts are transient components.

components in both the time and frequency domains using the Fourier coefficients. This makes the Fourier transforms less than optimal representations for analyzing signals, images and patterns containing transient or localized components. As a matter of fact, in pattern recognition, many important features are highly localized in spatial positions. The above weakness of the Fourier transform obstructs its application to pattern recognition.

Look at the following three examples, which deal with the field of pattern recognition, and are shown in Figs. 1.4 and 1.5. First, we consider an image of an aircraft with several drawing lines and character strings, which are displayed in Fig. 1.4(a). The contour of the aircraft is a localized image component, meanwhile, the drawing lines and character strings are also the transient image components. To recognize the aircraft from other objects, the contour of it is a useful feature, which is required to extract from the image. On the other hand, the drawing lines and character strings are the obstacles for the recognition, which have to be removed. For these two different image components, the Fourier analysis can not treat them efficiently.

The second example is illustrated in Fig. 1.4(b), which is a English letter "A" with noise. In pattern recognition, the boundary of the character should be extracted as a good feature, while, the noise has to be deleted. To do so, the transient component, boundary of letter "A", needs to be

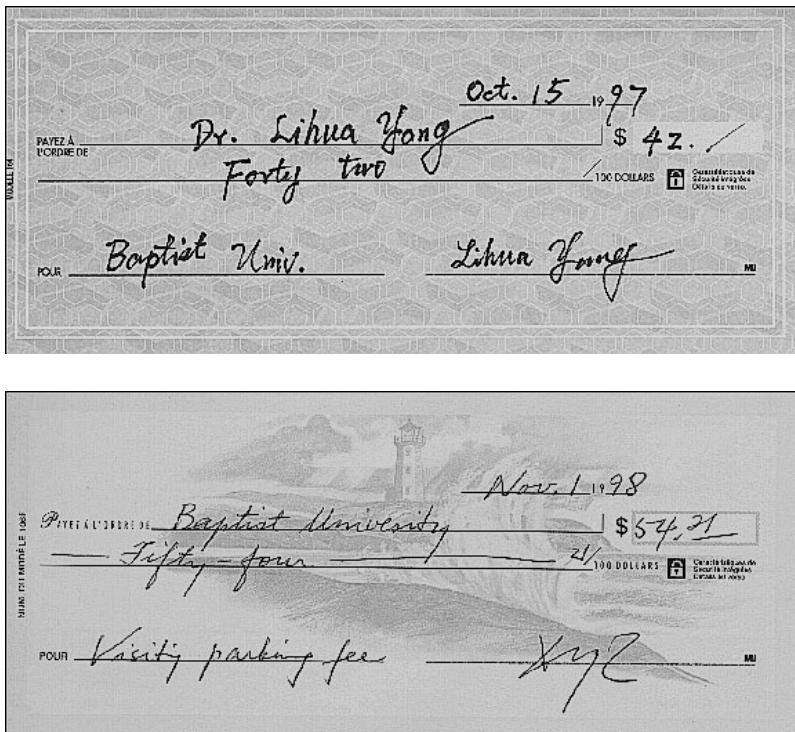


Fig. 1.5 The reference lines of bank check are transient components.

localized and processed. The deficiency of the Fourier analysis will obstruct this task.

The last example as shown in Fig. 1.5 deals with document analysis, which is an active branch in pattern recognition,. Bank check is very important and widely used in our daily life. In order to automatically process it, detecting reference lines plays a key role. However, these lines belong to the narrow transient components of the image. It is difficult to process them using Fourier analysis.

We, now, turn to discuss the approximation of signals $f_1(x)$ and $f_2(x)$ using basic wave function $\sin x$.

Fig. 1.6 consists of two columns. The left column describes the original non-transient signal $f_1(x)$ as shown in the top, and its approximations using

the partial sums of the Fourier expansions with $N = 4, 8, 15, 30$, from the top to the bottom, respectively. The right column presents the transient signal $f_2(x)$ and its approximations using the partial sums of the first N coefficients of Fourier expansions, $N = 4, 8, 15, 30$, from the top to the bottom, respectively.

The errors between the original signals and the partial sums of the Fourier expansions are illustrated in Fig. 1.7. The left column presents the original non-transient signal $f_1(x)$, and the errors. $f_1(x)$ is located on the top, the errors with different number of terms of Fourier expansions are described from the next-to-the-top to the bottom, which correspond to $n = 4, 8, 15$ and 30 respectively. The right column presents the original signal $f_2(x)$ and the errors with $n = 4, 8, 15, 30$, from the top to the bottom, respectively.

From the left columns in Figs. 1.6 and 1.7, it is clear that, for the non-transient signal $f_1(x)$, the fore 17 terms of the Fourier expansion can approximate the original signal very well. This corresponds to $n = 8$ in Eq. (1.3). It means that the series of $\{a_0, a_1, b_1, \dots, a_8, b_8\}$ contain the most of information in $f_1(x)$. Therefore, the representation of the non-transient signal $f_1(x)$ by the basic function $\sin x$ is extremely well-founded.

Unfortunately, the situation is not so good for the transient signal $f_2(x)$. It can be found, in the right columns in Figs. 1.6 and 1.7, that in the case of $n = 8$, the approximation error is still large. Although, the error is relatively small when $n = 30$, it is still visible. The reason for this occurrence is that for any transient or localized signals, like $f_2(x)$, the smooth basic wave $\sin x$ can not fit them perfectly.

To contest the above deficiencies, we should find other basic function $\psi(x)$ to replace the basic function $\sin x$. This new basic function should satisfy the following conditions:

- Similar to $\sin x$, any complicated signal $f(x)$ can be constructed by the linear combination of $\psi(jx - k)$ ($j, k \in \mathbb{Z}$), which are produced by the dilations and translations of the basic function $\psi(x)$.
- The expansion coefficients of a signal using the basic function $\psi(x)$ can reflect the locations of the transient or localized image components in the time domain.
- This new basic function $\psi(x)$ and its family can “fit” transient signal $f(x)$ much better than Fourier basic wave $\sin x$. In other words, they can minimize the error between the approximation of

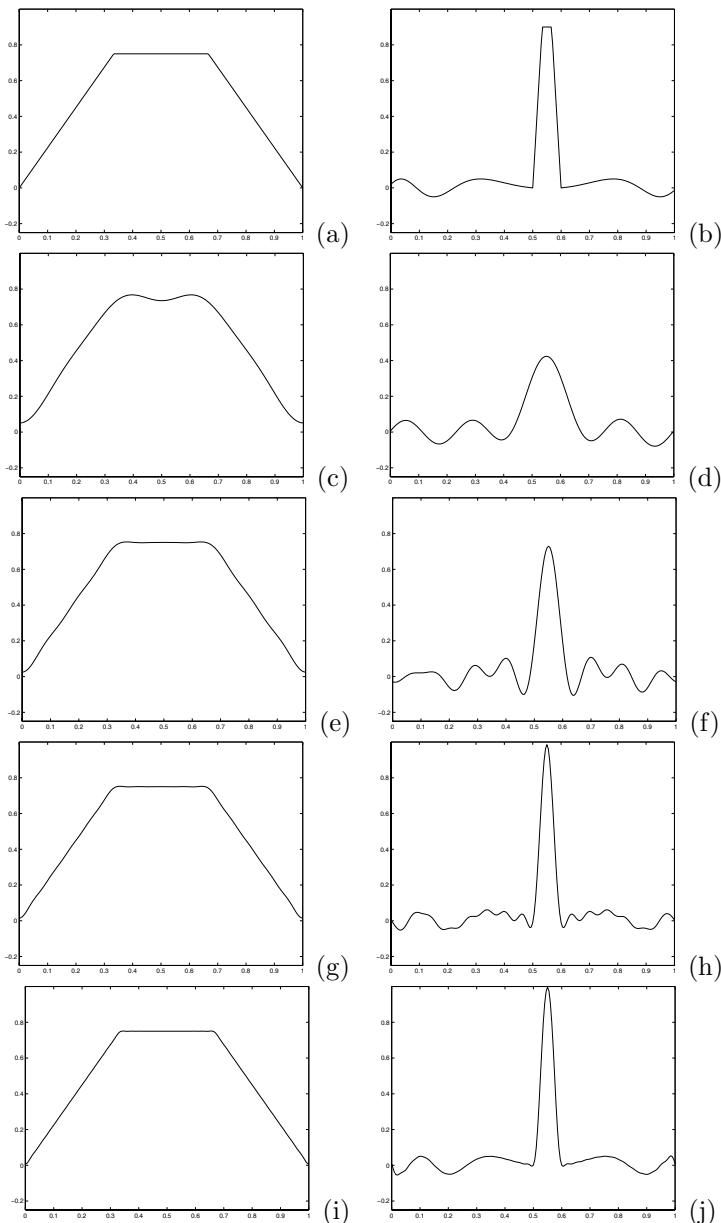


Fig. 1.6 The left column: original signal $f_1(x)$ and its partial sums of the Fourier expansion with $N=4, 8, 15, 30$ respectively. The right column: original signal $f_2(x)$ and its partial sums of the Fourier series of $f_2(x)$ with $N=4, 8, 15, 30$ respectively.

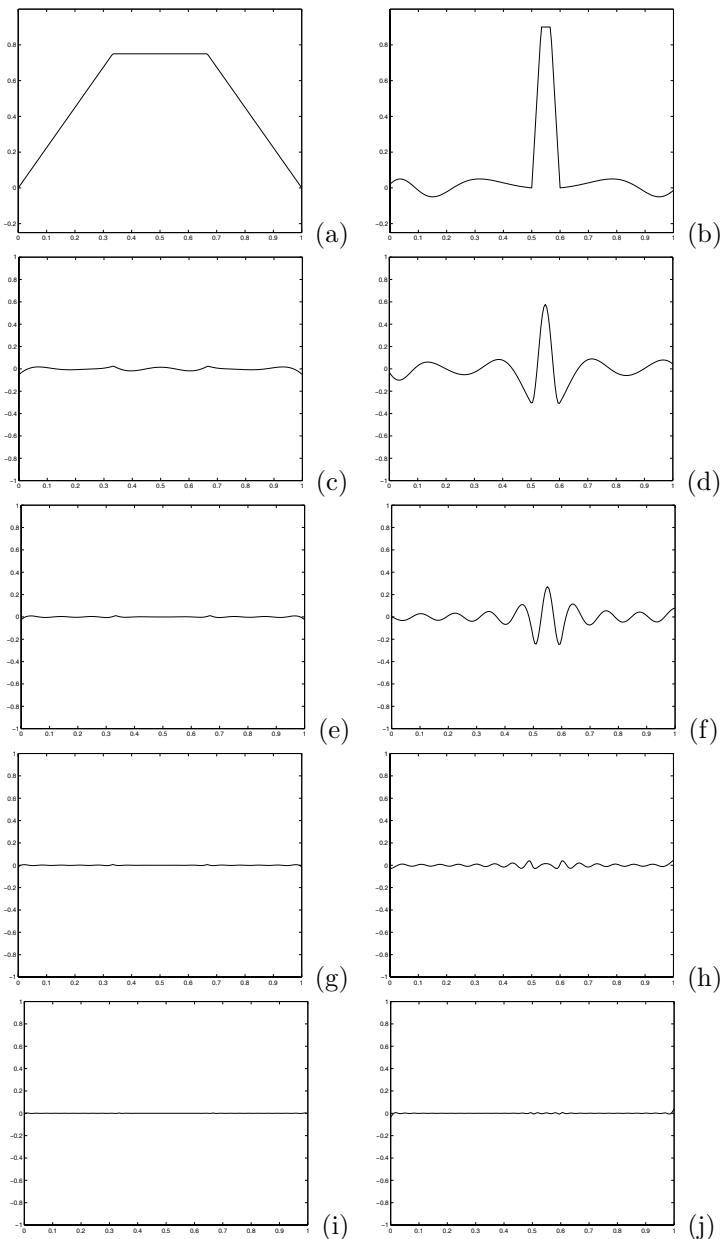


Fig. 1.7 The left column: original signal $f_1(x)$ and the errors with $N=4, 8, 15, 30$ respectively. The second column: original signal $f_2(x)$ and the errors with $N=4, 8, 15, 30$ respectively.

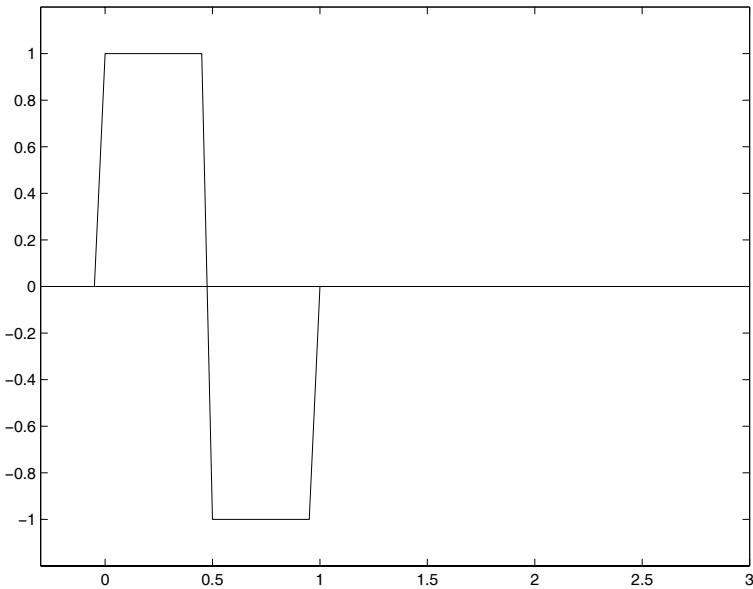


Fig. 1.8 Haar Wavelet.

the signal $f(x)$ and $f(x)$ itself.

In fact, we already found such a basic function $\psi(x)$ as early as 1910. It is the well-known square wave function – Haar function. This basic function $\psi(x)$ can be written in form of

$$\psi(x) = \begin{cases} 1 & x \in [0, \frac{1}{2}) \\ -1 & x \in [\frac{1}{2}, 1) \end{cases}$$

The graphic illustration of Haar function is presented in Fig. 1.8.

It has been mathematically proved that $\{\psi(2^j x - k) | j, k \in \mathbb{Z}\}$ can constitute an orthogonal basis of the finite energy signal space $L^2(\mathbb{R})$. They can also establish an orthonormal basis of $L^2(\mathbb{R})$ through a normalization:

$$\psi_{j,k}(x) := 2^{j/2} \psi(2^j x - k), \quad (j, k \in \mathbb{Z}). \quad (1.4)$$

Thus, any finite energy signal $f(x)$ can be represented by

$$f(x) = \sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} c_{j,k} \psi_{j,k}(x), \quad (1.5)$$

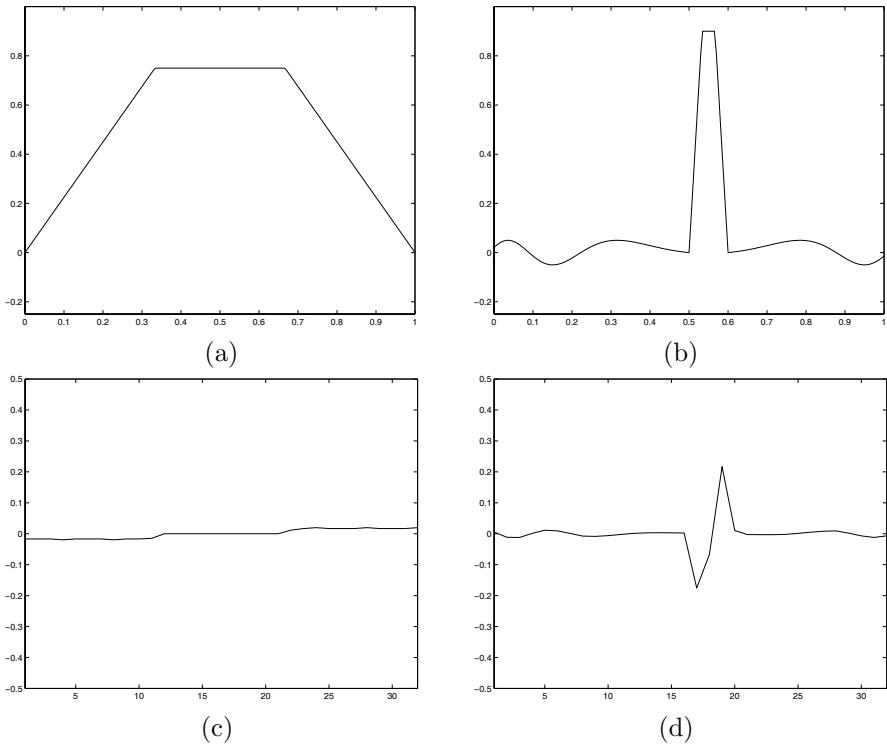


Fig. 1.9 The first row: the left side is the original signal $f_1(x)$, and the right side displays another original signal $f_2(x)$. The second row: the left side indicates the Haar coefficients corresponding to $f_1(x)$, and the right side shows the Haar coefficients corresponding to $f_2(x)$.

where $c_{j,k} := \int_{\mathbb{R}} f(x)\psi_{j,k}(x)dx$. Consequently, the first condition holds for Haar wavelet.

Now, we will expand signals $f_1(x)$ and $f_2(x)$ with Haar family. A part of the coefficients of the Haar wavelet expansion for $N = 5$, are shown in Fig. 1.9.

The original signals $f_1(x)$ (without transient components), and $f_2(x)$ (with transient components) are illustrated in Figs. 1.9(a) and (b) respectively. By applying the Haar transform to these signals, the Haar coefficients $\{c_{N,k}|k = 0, \dots, 2^N - 1\}$ of signals $f_1(x)$ and $f_2(x)$ are obtained and plotted in Figs. 1.9(c) and (d) respectively. The above coefficients can

successfully localize the signals. For the non-transient signal $f_1(x)$, nearly all Haar coefficients close to zero, which implies that no high frequencies are included in the original signal. By contrast, for the transient signal $f_2(x)$, only a few coefficients are affected, and produce a vibration in the frequency domain. Two peaks of the vibration in Fig. 1.9(d) correspond to just the positions of the transient components in signal $f_2(x)$ shown in Fig. 1.9(b). Meanwhile, each coefficient is just determined by the local action of the signal. In this way, Haar family can be employed to analyze the localization of signals in the time domain. Consequently, the second condition also holds for the Haar wavelet.

Next, we turn to the verification of third condition, i.e. we will check the question: can we apply the Haar wavelet to approximate signals with small error?

In Eq. 1.5, we denote

$$f_0(x) = \sum_{j=-\infty}^{-1} \sum_{k \in \mathbb{Z}} c_{j,k} \psi_{j,k}(x).$$

Similar to the Fourier expansion, by considering the energy, we have

$$\begin{aligned} & \int_{\mathbb{R}} |f(x) - [f_0 + \sum_{j=0}^N \sum_{k \in \mathbb{Z}} c_{j,k} \psi_{j,k}(x)]|^2 dx \\ &= \min_{d_0, d_1, \dots, d_N} \int_{\mathbb{R}} |f(x) - [f_0 + \sum_{j=0}^N \sum_{k \in \mathbb{Z}} d_{j,k} \psi_{j,k}(x)]|^2 dx \rightarrow 0 \quad (N \rightarrow \infty). \end{aligned}$$

where

$$f_0(x) + \sum_{j=0}^N \sum_{k \in \mathbb{Z}} c_{j,k} \psi_{j,k}(x)$$

is an approximation of $f(x)$.

We take signals $f_1(x)$ and $f_2(x)$ as examples again. The approximations of these signals using Haar family are presented graphically in Fig. 1.10. The original non-transient $f_1(x)$ and transient signals $f_2(x)$ are illustrated in Figs. 1.10(a) and (b) respectively. The approximation of $f_1(x)$ by the partial sums of the Haar expansion with $N=4$ (32 terms) and that with $N=5$ (64 terms) are shown in Figs. 1.10(c) and (g) respectively. The similar description of $f_2(x)$ by the partial sums of the Haar coefficients with $N=4$ and that with $N=5$ are displayed in Figs. 1.10(d) and (h) respectively. The

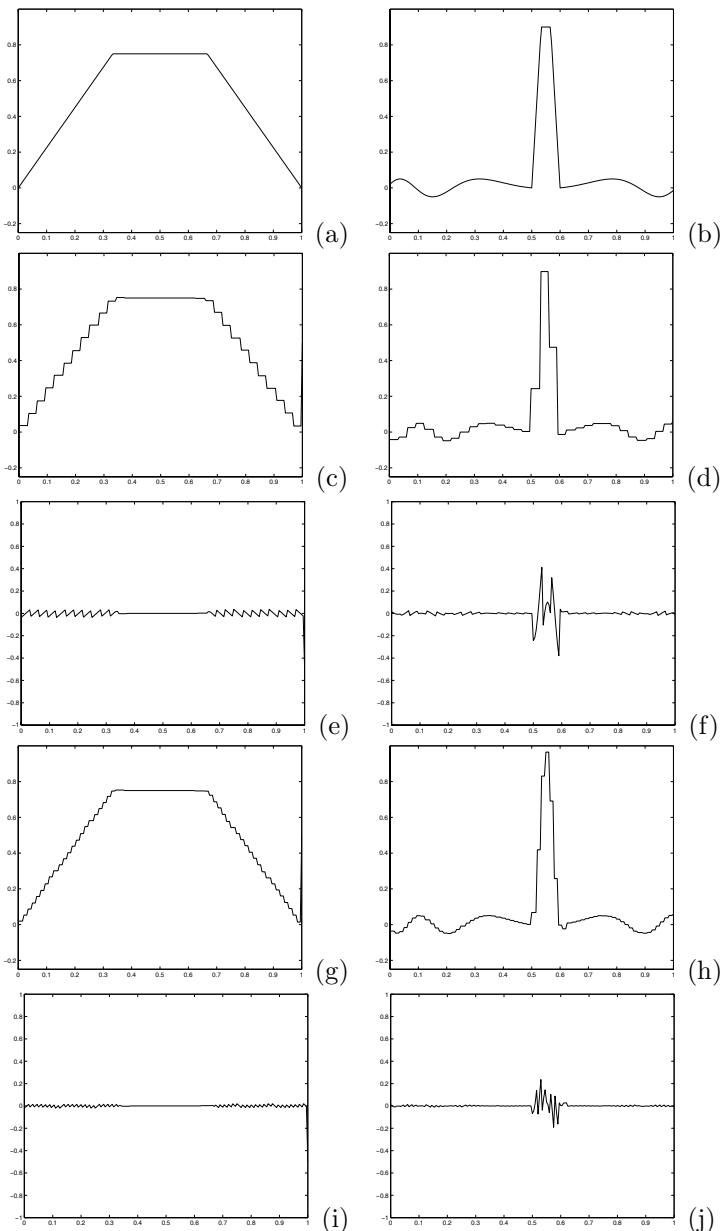


Fig. 1.10 Haar functions approximate signals: The left column: original non-transient signal $f_1(x)$, the partial sums of the Haar expansion with $N=4, 5$, and the approximating errors. The right column: original transient signal $f_2(x)$, the partial sums of the Haar expansion with $N=4, 5$, and the approximating errors.

errors between the original signal $f_1(x)$ and its approximations using the Haar function are presented in Figs. 1.10(e) and (i), which correspond to N=4 and 5 respectively. The errors between the original signal $f_2(x)$ and its approximations by the Haar coefficients are described graphically in Figs. 1.10(f) and (j) corresponding to N=4 and 5.

Now, we compare the result using $\sin x$ with that using the Haar function, in case of N=4. As shown in Figs. 1.6(d) and 1.10(h), when the Haar function is employed to approximate a transient signal, the result seems better than that when we use $\sin x$ due to the strong localization of the Haar family. Mathematically, it is said that the Haar function is compactly supported. Unfortunately, it can be found, from Fig. 1.10(g), that when we approximate a non-transient signal with Haar wavelet, the result is not so good as Fourier expansion. In this situation, saw-tooth-wave errors occur due to the discontinuity of Haar function. In summary, comparing with $\sin x$, Haar wavelet has much better localization and rather worse smoothness. That is why Haar wavelet did not attract enough attention for a long time. Before various wavelet bases were constructed in 1980s, it was only a dream to find out a basic wavelet $\psi(x)$, which has both the strong localization and smoothness. At the end of 1980s, wavelets achieved breakthrough development in both theory and application.

It has followed a long and tortuous course from Fourier analysis to wavelet analysis. In fact, as early as 1946, Gabor [Gabor, 1946] has found out that Fourier analysis is lacking to localize signals in the time domain. He applied Gaussian function as a “window” to improve the Fourier transform. It is referred to as Gabor transform and is applied to analyze the transient signals. Thereafter, Gabor transform led to the general window Fourier transform (or short-time Fourier transform) by the replacement of Gaussian function with other localized window functions. However, the size of the window in the window Fourier transform is fixed, it did not improve Fourier transformation thoroughly. Therefore, in this book, we will not discuss it in detail.

Today, wavelet analysis has become an international focus in many research fields and applications. Scientists and engineers in all over the world are working on wavelets widely and deeply in mathematics, engineering and many other disciplines.

The fundamental difference between the signal decompositions by wavelet transform and Fourier transform is depicted in Fig. 1.11. A singer sings a

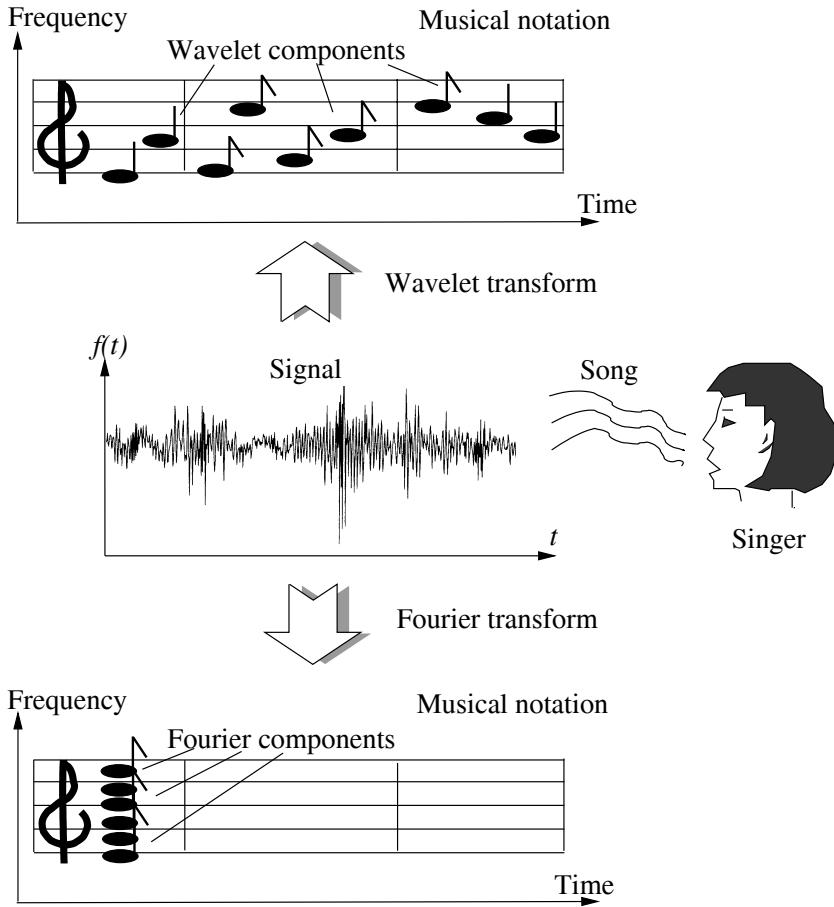


Fig. 1.11 Fundamental difference between the signal decompositions by wavelet transform and Fourier transform: A singer sings a song, which is the signal $f(t)$ of time t . It can be decomposed into the musical notes by wavelet transform due to the time-frequency localization. However, it would be failed if the Fourier transform will be utilized.

song, which is the signal $f(t)$ of time t . It can be decomposed into the musical notes by wavelet transform due to the property of the time-frequency localization of the wavelet transform. The musical notation can be viewed as depicting a 2D time-frequency space. Frequency (pitch) increases from the bottom of the scale to the top, while time (measured in beats) advances to the right. Each note on the sheet music corresponds to one wavelet com-

ponent, which appears in the recording of a performance of the song. If we were to analyze a recorded musical performance and write out the corresponding score, we would have a type of wavelet transform. Similarly, a recording of a musician's performance of a song can be viewed as an inverse wavelet transform, since it reconstructs the signal from a time-frequency representation. However, it would be failed if we use the Fourier transform to do so.

1.2 Brief Review of Pattern Recognition with Wavelet Theory

Wavelet theory is a versatile tool with very rich mathematical content and great applications. It has been employed in many fields and applications, such as signal processing, image analysis, communication systems, biomedical imaging, radar, air acoustics, theoretical mathematics, control system, and endless other areas. A lot of achievements have been made, for instance, many of them have been published [Auslander et al., 1990; Beylkin et al., 1991; Chui, 1992; Daubechies, 1990; Grossmann and Morlet, 1984; IEEE, 1993; Mallat, 1989b; SPIE, 1994; Daugman, 2003; Kumar and Kumar, 2005; Shankar et al., 2007; You and Tang, 2007; Li, 2008]. However, the research on applying the wavelets to pattern recognition is still too weak; not too many research projects deal with this topic at the present.

In this sub-section, a brief survey of pattern recognition with the wavelet theory is presented.

As for the applications of wavelet theory to pattern recognition, we can consider them to be two hands, namely, (1) system-component-oriented, and (2) application-task-oriented. Fig. 1.12 displays these two sides. On the left hand side, four components are enclosed in a recognition system. On the right hand side, the application-task-oriented ones can be categorized into the following groups:

- Iris pattern recognition
[Daugman, 2003],
- Face recognition using wavelet transform
[Kouzani and Ong, 2003; Lai et al., 1999; Yang et al., 2003b],
- Hand gestures classification
[Kumar and Kumar, 2005; Kumar et al., 2003; Sharnia et al., 2004],

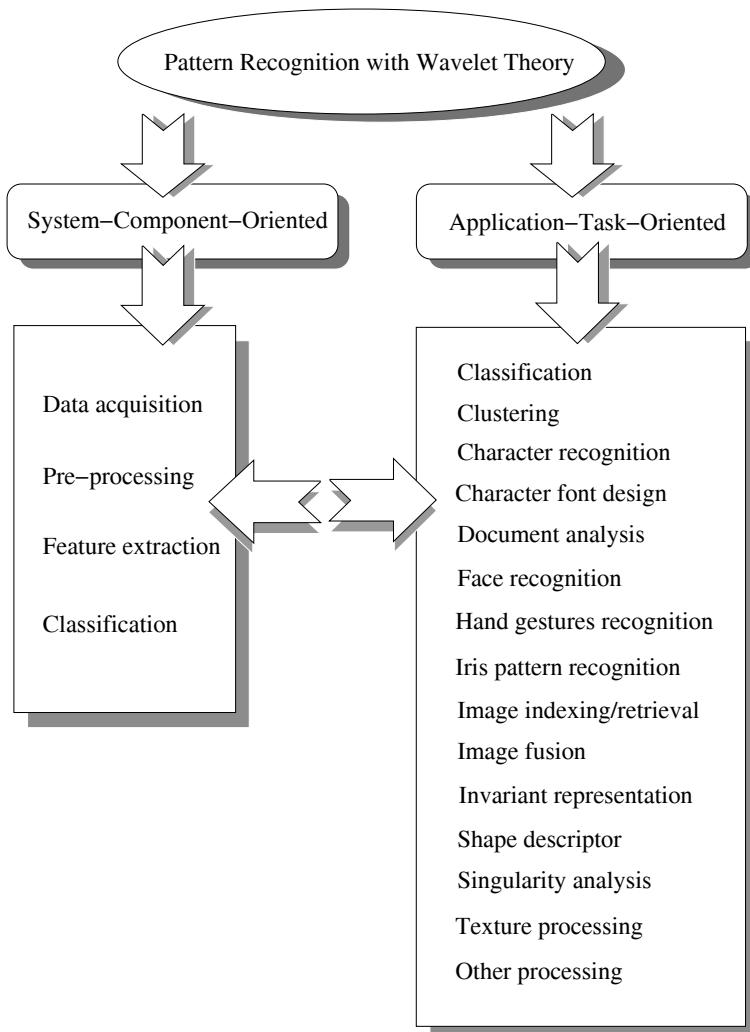


Fig. 1.12 Pattern recognition with wavelet theory.

- Classification and clustering
[Murtagh and Starck, 1998; Shankar et al., 2007; Tang and Ma, 2000],
- Document analysis with wavelets
[Liang et al., 1999; Tang et al., 1996a; Tang et al., 1997a; Tang

- et al., 1995a; Tang et al., 1997c],
- Analysis and detection of singularities with wavelets
[Mallat and Hwang, 1992; Young, 1993; Chen et al., 1995; Chen and Yang, 1995; Deng and Lyengar, 1996; Law et al., 1996; Tang et al., 1997c; Thune et al., 1997; Tieng and Boles, 1997a; Tang et al., 1998d; Tang et al., 2000; Tang and You, 2003],
- Wavelet descriptors for shapes of the objects
[Chuang and Kuo, 1996; Tang et al., 1998a; Tang et al., 1999; Tieng and Boles, 1997b; Wunsch and Laine, 1995; Yang et al., 2003a; You and Tang, 2007],
- Invariant representation of patterns
[Haley and Manjunath, 1999; Shen and Ip, 1999; Tang et al., 1998a; Yang et al., 2003a; Yoon et al., 1998],
- Handwritten and printed character recognition
[Kunte and Samuel, 2007; Lee et al., 1996; Tang et al., 1998a; Tang et al., 1999; Tang et al., 1996b; Tang et al., 1998d; Wunsch and Laine, 1995; You and Tang, 2007],
- Texture analysis and classification
[de Wouwer et al., 1999b; de Wouwer et al., 1999a; Haley and Manjunath, 1999; Liang and Tjahjadi, 2006; Muneeswaran et al., 2005],
- Image indexing and retrieval
[Jain and Merchant, 2004; Ksantini et al., 2006; Kubo et al., 2003; Moghaddam et al., 2005; Special-Issue-Digital-Library, 1996; Smeulders et al., 2000],
- Wavelet-based image fusion
[El-Khamy et al., 2006; Li, 2006; Li, 2008],
- Character processing with B-spline wavelet transform
[Yang et al., 1998],
- Others
[Chambolle et al., 1998; Chen and Yang, 1995; Combettes and Pesquet, 2004; Combettes, 1998; Liao and Tang, 2005; You et al., 2006]

It is clear that the two sides are related to each other. Each group of the right hand side relates to the component(s) on the left hand side. For instance, the tasks in the shape descriptor, character recognition, etc. concern all the components of the pattern recognition system; Singularity detector and invariant representation are related with the tasks, which are

carried out in the component of the feature extraction. In this sub-section, we discuss the applications of wavelet theory to pattern recognition by means of the right hand side, i.e. in the application-task-oriented point of view.

1.2.1 Iris Pattern Recognition

The highest density of biometric degrees-of-freedom which is both stable over time and easily measured, is to be found in the complex texture of the iris pattern of the eye [Daugman, 2003]. An example of the human iris pattern is presented in Fig. 1.13(a), which is imaged in near infrared light at a distance of 30 cm. It can be served as unique and reliable identifier.

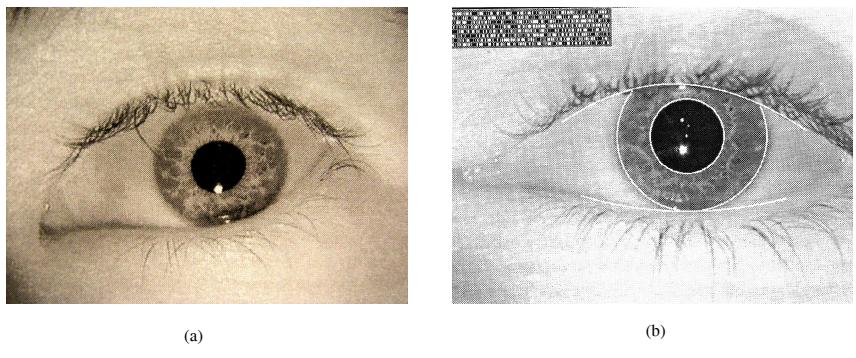


Fig. 1.13 (a) Example of a human iris pattern, (b) Isolation of an iris for encoding, and its resulting “IrisCode” [Daugman, 2003].

J. Daugman [Daugman, 2003] applies complex-valued 2D wavelet to iris pattern recognition. Let $\Psi(x, y)$ be a 2D mother wavelet, we can generate a complete self-similar family of parametrized daughter wavelets $\Psi_{muv\theta}(x, y)$ by

$$\Psi_{muv\theta}(x, y) = 2^{-2m} \Psi(x', y'),$$

$$\begin{aligned} x' &= 2^{-2m} [x \cos(\theta) + y \sin(\theta)] - u, \\ y' &= 2^{-2m} [-x \sin(\theta) + y \cos(\theta)] - v, \end{aligned}$$

where, the substituted variables x', y' incorporate dilation of the wavelet in size by 2^{-2m} , translations in position (u, v) , and rotations through angle θ .

In the complex-valued 2D wavelet, it is possible to use the real and imaginary parts of their convolution (\star) with an iris image $f(x, y)$ to extract a description of image structure in terms of local modulus and phase. Let $\Psi_{muv\theta}(x, y)$ be 2D daughter wavelet we used, the amplitude modulation function $A(x, y)$ and phase modulation function $\theta(x, y)$ can be presented respectively:

$$A(x, y) = \sqrt{(\Re\{\Psi_{muv\theta}(x, y) \star f(x, y)\})^2 + (\Im\{\Psi_{muv\theta}(x, y) \star f(x, y)\})^2},$$

and

$$\theta(x, y) = \tan^{-1} \frac{\Im\{\Psi_{muv\theta}(x, y) \star f(x, y)\}}{\Re\{\Psi_{muv\theta}(x, y) \star f(x, y)\}}.$$

The following operations are accomplished to localize precisely the inner and outer boundaries of the iris, and thereafter, to detect eyelids if they intrude, and exclude them:

$$\max_{r, x_0, y_0} \left| G_\sigma(r) \star \frac{\partial}{\partial r} \oint_{r, x_0, y_0} \frac{f(x, y)}{2\pi r} ds \right|,$$

where, contour integration parametrized for size and local coordinates r, x_0, y_0 at a scale of analysis σ set by some blurring function $G_\sigma(r)$ is performed over the iris pattern $f(x, y)$. The result of this optimization search is the determination of the circle parameters r, x_0, y_0 , which best fit the inner and outer boundaries of the iris. Fig. 1.13(b) shows the isolation of an iris for encoding, and its resulting “IrisCode” [Daugman, 2003].

1.2.2 Face Recognition Using Wavelet Transform

1. Wavelet-based PCA Approach

Principal component analysis (PCA) is one of the well-known methods for representing human face. [Sirovich and Kirby, 1987] first proposed to use Karhunen-Loeve transform to represent human faces. In their method, faces are represented by a linear combination of weighted eigenvector, known as eigenface features. In 1991, [Turk and Pentland, 1991] developed a face recognition system using PCA (K-L expansion). In 1993, [O’Toole et al., 1993] demonstrated that whereas the low-dimensional representation is optimal for identifying physical categories of face, such as gender and race, it is not optimal for recognizing a human face. In 1996, [Swets and Weng,

1996] combined the theories of K-L projection and multi-dimensional discriminant analysis to generate a set of most discriminating features for recognition.

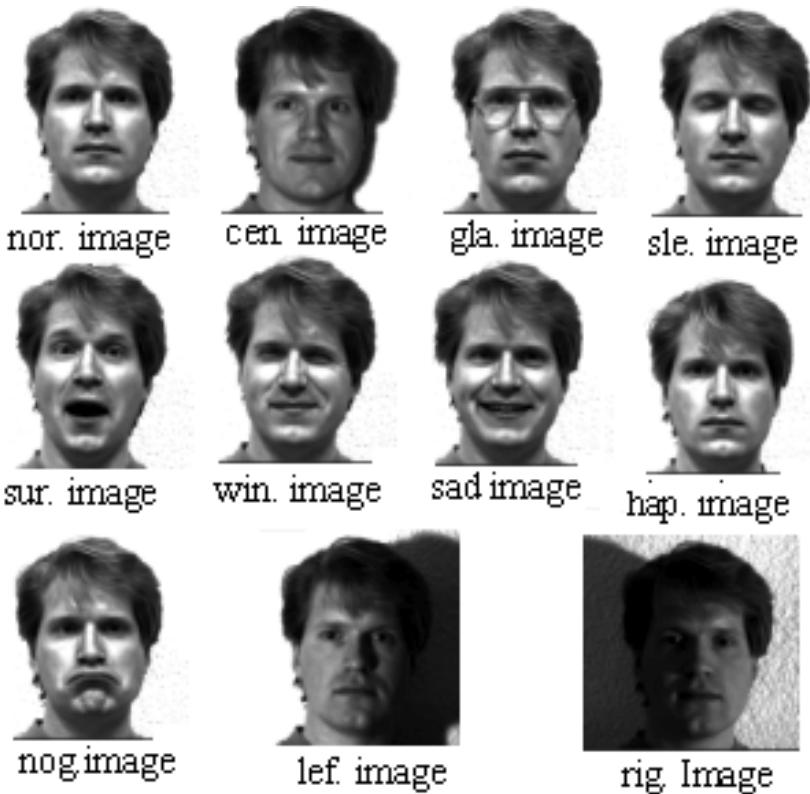


Fig. 1.14 Images of one person in Yale database.

Very good results are obtained when applying this approach to frontal-view head-&-shoulder images. However, if the face is with different orientations, facial expressions or occluded, the accuracy will be degraded dramatically. [Nastar and Ayache, 1996] investigated the relationship between variations in facial appearance and their deformation spectrum. They found that facial expressions and small occlusion affect the intensity manifold locally. Under frequency-based representation, only high frequency spectrum is affected, called high frequency phenomenon. Moreover, changes in pose

or scale of a face affect the intensity manifold globally, in which only their low frequency spectrum is affected, called low frequency phenomenon. Only a change in face will affect all frequency components.

The results of [Nastar and Ayache, 1996] shed light on how to solve the problem in facial appearance variations - use mid-range frequency for recognition. [Yuen et al., 1998] recently demonstrated that applying PCA method on Wavelet Transformed (WT) sub-image with mid-range frequency components gives better recognition accuracy than (1) applying PCA on WT sub-image with low frequency components only, or (2) applying PCA on the original image that contains all frequency components.

In view of these, [Lai et al., 1999] combines Fourier transform and wavelet transform for face recognition. Their method is described as follows. First, a compact support and orthogonal wavelet is applied to decompose the face image. The low-frequency subband of the decomposed is selected to represent his mother image. It is an optimal approximate image of mother image in lower dimension. Second, Fourier Transform (FT) is applied to the low-frequency subband images and represents them.

In [Lai et al., 1999], Yale and Olivetti databases are selected to evaluate their method. One person from each database is shown in Figs. 1.14 and 1.15. The Spectroface with different levels of wavelet decomposition of one face image from Yale database is presented in Fig. 1.16.

It is reported that

- The performance of Spectroface is better than that of Eigenface and Template matching, especially for facial expression and occlusion (by glasses) images.
- Spectroface approach is invariant to spatial translation. Both Eigenface and Template matching methods are sensitive to the spatial translation.
- The computation complexity of Spectrofaces is dependent on wavelet packet. Expanding N vectors $\{X_n \in R^d : n = 1, 2, \dots, N\}$ into wavelet packet coefficients is in $O(Nd \log d)$.

2. Nonlinear Wavelet Approximation Method

For a 2D orthonormal wavelet basis, let ψ^1 , ψ^2 and ψ^3 be three 2D wavelets, $\psi^0 := \phi$ be the corresponding scaling function, and

$$\psi_j^{k,e} := 2^{k/2}\psi^e(2^kx - j), \quad \forall e \in \Re_k,$$

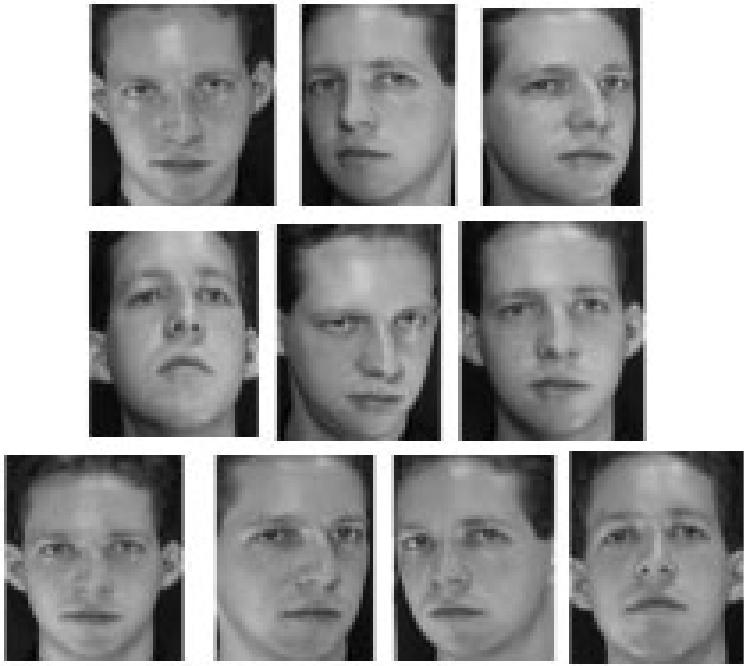


Fig. 1.15 Images of one person in Olivetti database.

where

$$\mathfrak{R}_k := \begin{cases} \{1, 2, 3\}, & k > 0, \\ \{0, 1, 2, 3\}, & k = 0. \end{cases}$$

Thus, any $f \in L^2(I)$ can be written by

$$f = \sum_{k \geq 0} \sum_j \sum_{e \in \mathfrak{R}_k} c_j^{k,e} \psi_j^{k,e}.$$

Let $X = \text{span}\{\psi_j^{k,e} : e \in \mathfrak{R}_k, 0 \leq k < K, j \in \mathbb{Z}^2\}$, then the approximation element

$$\tilde{f} = \sum_{0 \leq k < K} \sum_j \sum_{e \in \mathfrak{R}_k} c_j^{k,e} \psi_j^{k,e}$$

consists of

$$N = 3(2^{K-1} \cdot 2^{K-1} + 2^{K-2} \cdot 2^{K-2} + \dots + 2^1 \cdot 2^1 + 2^0 \cdot 2^0) = 2^{2K}$$

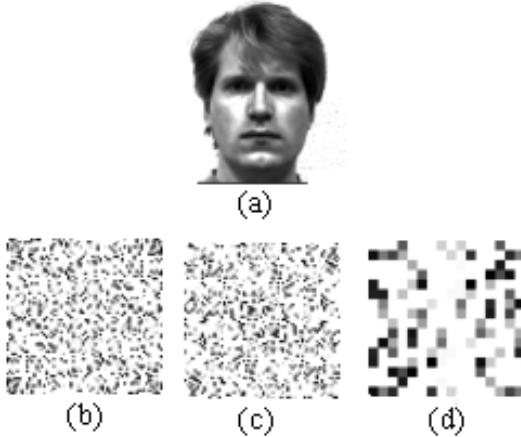


Fig. 1.16 (a) The original image; (b) The frequency representation of image (a); (c) Spectroface with 1-level wavelet decomposition; (d) Spectroface with 3-level wavelet decomposition.

terms in the sum. It is the unique best approximation element of f in X , and the linear approximation is presented by

$$\|f - \tilde{f}\|_{L_2(I)} \leq N^{-\alpha/2} \|f\|_{W^\alpha(L^2(I))}.$$

For any sequence $\{c_j^{k,e}\}$, let $\aleph(\{c_j^{k,e}\})$ be the number of nonzero items. We can denote

$$\sum_N := \left\{ \sum_{k \geq 0} \sum_j \sum_{e \in \mathcal{R}_k} c_j^{k,e} \psi_j^{k,e} \mid \aleph(\{c_j^{k,e}\}) \leq N \right\}.$$

For any

$$f = \sum_{k \geq 0} \sum_j \sum_{e \in \mathcal{R}_k} c_j^{k,e} \psi_j^{k,e},$$

we keep N coefficients with larger amplitudes and eliminating the small ones. Thus, an approximation \tilde{f} can be got, and $\tilde{f} \in \sum_N$. Finally, the nonlinear wavelet approximation can be arrived:

$$\|f - \tilde{f}\|_{L_2(I)} \leq N^{-\alpha/2} \|f\|_{B_\sigma^\alpha(L^\sigma(I))},$$

where, $B_\sigma^\alpha(L^\sigma(I))$ is a Besov space, $0 < \alpha < 1$ and $\sigma = 2/(1 + \alpha)$.

L. H. Yang et al. propose a method to face recognition using the nonlinear wavelet approximation [Yang et al., 2003b]. As mentioned by the authors, the linear wavelet approximation cannot reduce the approximation error by adding more coefficients. The nonlinear wavelet approximation has an important advantage over the linear one due to its ability to do so.

3. Wavelet Packets Transform for Classification of Facial Images

A. Z. Kouzani and S. H. Ong [Kouzani and Ong, 2003] utilize wavelet packets transform for lighting-effects classification in facial images. The wavelet packets transform is a generalization of the wavelet transform. In the wavelet transform, only the low-pass filter is iterated. It is assumed that lower frequencies contain more important information than higher frequencies. This assumption is not true for many images. The main difference between the wavelet packets transform and the wavelet transform is that, in the wavelet packets, the basic two-channel filter bank can be iterated either over the low-pass branch or the high-pass branch. This provides an arbitrary tree structure with each tree corresponding to a wavelet packets basis. It can offer a choice of optimal bases for the representation of specific signal. The basis can be selected to minimize the number of significantly nonzero coefficients in the transform [Kouzani and Ong, 2003]. Entropy is a suitable function for choice of the best basis. Shannon's equation for entropy enables searching for the smallest entropy expansion of a signal. It is presented below:

$$\begin{aligned} \epsilon^2(v; \{H_i, H_j\}) = & - \frac{\|v_+\|^2}{\|v\|^2} \ln \left(\frac{\|v_+\|^2}{\|v\|^2} \right) - \frac{\|v_-\|^2}{\|v\|^2} \ln \left(\frac{\|v_-\|^2}{\|v\|^2} \right) \\ & + \|v_+\|^2 \epsilon^2 \left(\frac{v_+}{\|v_+\|^2}, \{H_i\} \right) \\ & + \|v_-\|^2 \epsilon^2 \left(\frac{v_-}{\|v_-\|^2}, \{H_j\} \right), \end{aligned}$$

where, H is a Hilbert space, $v \in H$, $\|v\| = 1$ and $H = \oplus \sum H_i$ is an orthogonal decomposition of H .

The best basis algorithm [Kouzani and Ong, 2003] minimizes the cost function for the transform coefficients. It takes a complete decomposition according to the wavelet packets transform. In each node, which corresponds to a subspace of the image, the cost of the coefficient of the subspace is calculated and iterated.

1.2.3 Hand Gestures Classification

1. Visual Hand Gestures Classification Using Wavelet Transforms

Interaction is a common activity of our daily life. Hand actions play a very important role in the interaction. To improve human machine interaction and for helping disable people, it is desirable for machines to extract more information from human hand actions [Kumar et al., 2003]. Some examples of hand gestures are illustrated in Fig. 1.17. The meanings of these gestures are “hold”, “three”, “victory” and “point” from left to right.

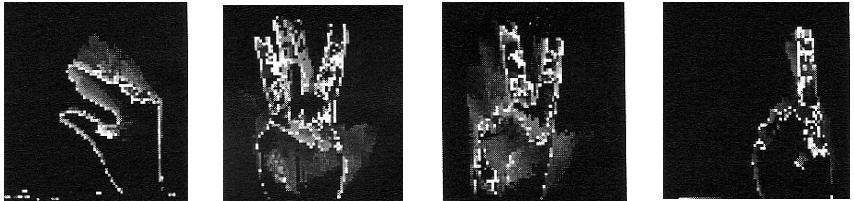


Fig. 1.17 Examples of human hand gestures [Kumar et al. 2003].

S. Kumar et al. present a novel technique for classifying human hand gestures based on stationary wavelet transform (SWT) [Kumar et al., 2003]. It uses view-based approach for representation of hand action, and artificial neural networks (ANN) for classification. This approach applies a cumulative image-difference method in the representation of action, which results in the construction of motion history image (MHI). Let $I(x, y, n)$ be an image sequence and let

$$D(x, y, n) = |I(x, y, n) - I(x, y, n - 1)|,$$

where, $I(x, y, n)$ is the intensity of each pixel at location (x, y) in the n th frame and $D(x, y, n)$ is the difference of consecutive frames representing regions of motion. Binarization of the difference image $D(x, y, n)$ over a threshold τ is $B(x, y, n)$

$$B(x, y, n) = \begin{cases} 1 & \text{if } D(x, y, n) > \tau, \\ 0 & \text{otherwise.} \end{cases}$$

The motion history image $MHI(H_N(x, y))$ is

$$MHI(H_N(x, y)) = \max \left(\bigcup_{n=1}^{N-1} B(x, y, n) \times n \right),$$

where, N represents the duration of the time window used to capture the motion.

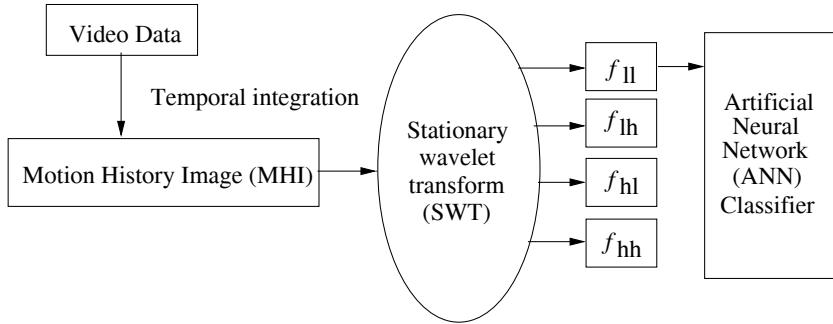


Fig. 1.18 The schematic diagram of the approach [Kumar et al. 2003].

These MHI's are decomposed into four sub-images (f_{ll} , f_{lh} , f_{hl} , f_{hh}) using stationary wavelet transform. The average image (f_{ll}) is fed as the global image descriptor to the artificial neural networks (ANN) for classification. The schematic diagram of the approach is illustrated in Fig. 1.18.

2. Hand Gestures Classification by Wavelet Transforms and Moment Based Features

S. Kumar and D. K. Kumar [Kumar and Kumar, 2005] propose a novel technique for classifying human hand gestures based on stationary wavelet transform (SWT) and geometric based moments. According to uniqueness theory of moments for a digital image of size (N, M) , the $(p + q)$ th-order moments m_{pq} are calculated by

$$m_{pq} = \frac{1}{NM} \sum_{x=1}^N \sum_{y=1}^M f(x, y)x^p y^q,$$

where $p, q = [0, 1, 2, \dots, n]$.

The schematic diagram of the wavelet-moment approach is illustrated in Fig. 1.19.

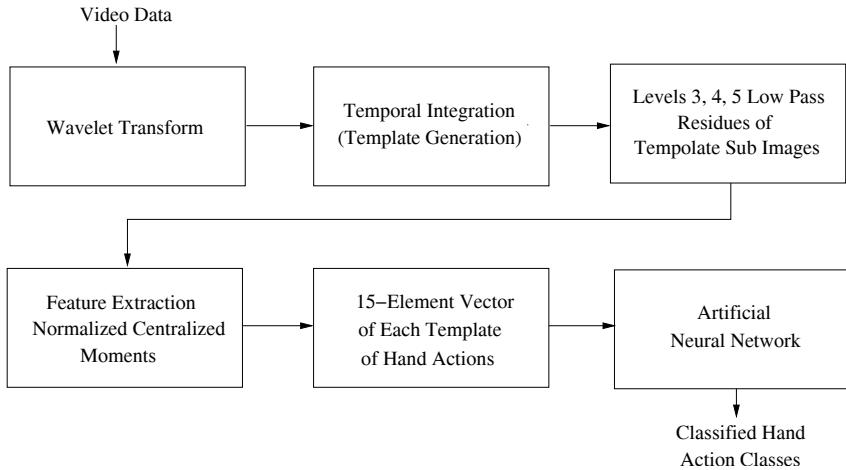


Fig. 1.19 The schematic diagram of the wavelet-moment approach [Kumar and Kumar, 2005].

3. Wavelet Directional Histograms of The Spatio-Temporal Templates of Human Gestures

A. Sharnia et al. [Sharnia et al., 2004] evaluate the efficacy of directional information of wavelet multi-resolution decomposition to enhance histogram-based classification of human gestures. The gestures are represented by spatio-temporal templates. This template collapses the spatial and temporal components of motion into a static gray scale image such that no explicit sequence matching or temporal analysis is required, and it reduces the dimensionality to represent motion. These templates are modified to be invariant to translation and scale. Two-dimensional, 3-level dyadic wavelet transforms are applied on the template resulting in one low-pass sub-image and nine highpass directional sub-images. Histograms of wavelet coefficients at different scales are used for classification purposes. The experiments demonstrate that while the statistical properties of the template provide high level of classification accuracy, the global detail activity available in highpass decompositions significantly improve the classification accuracy.

1.2.4 Classification and Clustering

Classification process can be categorized into two types:

- The classification with supervised learning: In this type of classification, there is a supervisor to teach the recognition system how to classify a known set of patterns, and thereafter, it let the system go ahead freely to classify other patterns. In this way, a priori information is needed to form the basis of the learning [Bow, 1992].
- The classification with non-supervised learning: The classification process is not depend on a priori information. Clustering is the non-supervised classification, which is the process of generating classes without any a priori knowledge about the patterns; neither can the proper training pattern sets be obtained.

As for the supervised classification with wavelet theory, there is one chapter in this book to discuss it. For the non-supervised classification, that is the clustering, this sub-section gives an example [Murtagh and Starck, 1998], where the wavelet theory is applied to clustering.

1. Classifier Design Based on Orthogonal Wavelet Series

Report [Tang and Ma, 2000] discusses the supervised classification. It provides an in-depth examination of the classifier design problem. First, it presents an overview of the fundamentals in pattern classifier design. In so doing, the emphasis is on *minimum average-loss classifier design* and *minimum error-probability classifier design*. Next, [Tang and Ma, 2000] specifically describes and discusses the use of orthogonal wavelet series in classifier design. It addresses the issue of how to derive a probability density estimate based on orthogonal wavelet series. From the multiresolution theory in wavelet analysis, it is known that

$$L^2(R) = \overline{\bigcup}_m V_m.$$

Let $p_m(x)$ denote the orthogonal project of $p(x)$ in space V_m . Thus, it can be found that

$$(L^2) \lim_{m \rightarrow \infty} p_m(x) = p(x), \quad p_m(x) = \sum_{n=-\infty}^{+\infty} a_{mn} 2^{m/2} \phi(2^m x - n).$$

The minimum mean square error estimator of $p_m(x)$ can be written as

$$\hat{p}_m(x) = \sum_{n=-\infty}^{+\infty} \left[\frac{1}{N} \sum_{i=1}^N 2^{m/2} \phi(2^m X_i - n) \right] 2^{m/2} \phi(2^m x - n).$$

A theorem is proved by [Tang and Ma, 2000], which indicates that when scaling function $\phi(x)$ and unknown density function $p(x)$ satisfy certain specific properties, orthogonal wavelet series density estimator $\hat{p}_m(x)$ will converge to $p(x)$.

Let scaling function $\phi(x) \in S_r$, and for a certain $\lambda \geq 1$, it satisfies property Z_λ . Let X be a continuous bounded density function random variable, and X_1, X_2, \dots, X_N be N independent identically distributed samples of X . Thus, if

$$p(x) \in H^\alpha, \quad \alpha > \lambda + \frac{1}{2}, \quad m \approx \lg N / (2\lambda + 1) \lg 2,$$

then

$$E|\hat{p}_m(x) - p(x)|^2 \leq O(2^{-2m\lambda}).$$

From the discussions in [Tang and Ma, 2000], we can note that the orthogonal wavelet series estimator differs from the kernel estimator and the traditional orthogonal series density estimator. Its basic idea shares some similarities to that of the traditional orthogonal series density estimator. However, it also satisfies several key properties of kernel estimator and exhibits some additional features. Generally speaking, the orthogonal wavelet series density estimator represents a new non-parametric way of estimating density functions, which has a great potential for practical applications. For instance, in pattern classifier design, sometimes, the probability density function, $p(x)$, of a certain feature vector may not be available. In such a case, we can readily replace $p(x)$ with $\hat{p}_m(x)$ using the above-described orthogonal wavelet series density estimator, and thus, effectively design the classifiers.

2. Neuro-Wavelet Classifier for Multispectral Remote Sensing Images

A neuro-wavelet supervised classifier is proposed in [Shankar et al., 2007] for land cover classification of multispectral remote sensing images. Features extracted from the original pixels information using wavelet transform (WT) are fed as input to a feed forward multi-layer neural network (MLP).

The WT basically provides the spatial and spectral features of a pixel along with its neighbors and these features are used for improved classification. For testing the performance of the proposed method, Two IRS-1A satellite images and one SPOT satellite image are used. Results are compared with those of the original spectral feature based classifiers and found to be consistently better. Simulation study revealed that Biorthogonal 3.3 (Bior3.3) wavelet in combination with MLP performed better compared to all other wavelets. Results are evaluated visually and quantitatively with two measurements, β index of homogeneity and Davies-Bouldin (DB) index for compactness and separability of classes. [Shankar et al., 2007] suggested a modified β index in accessing the percentage of accuracy (PA_β) of the classified images also.

3. Pattern Clustering Based on Noise Modeling in Wavelet Space

Point pattern clustering has constituted one of major strands in Cluster analysis. [Murtagh and Starck, 1998] describes an effective approach to object or feature detection in point patterns via noise modeling. Two advantages arrive at in this work, namely: (1)a multiscale approach with wavelet transform is computationally very efficient; (2) a direct treatment of noise and clutter, which leads to improve cluster detection. In this approach, the noise modeling is based on a Poisson process, and a non-pyramidal (or redundant) wavelet transform is applied.

Given a planar point pattern, a 2-D image is created by the following:

- Producing the tuple $(x, y, 1)$ when a point at (x, y) with value one;
- Projecting onto a plane by using a regular discrete grid (image) and assigning the contribution of points to the image pixels by an interpolation function, used by a wavelet transform referred to as à trous algorithm with a cubic B-spline.
- The à trous algorithm is then employed to the resulting image. The significant structures are extracted at each wavelet decomposition level, according to the noise model for the original image, $(x, y, 1)$.

The detailed description of the non-pyramidal (or redundant) wavelet transform can be found in [Shensa, 1992]. A summary of the “à trous” wavelet transform is presented below:

Step-1 To initialize i to 0, starting with an image $c_i(k)$, i.e. $c_0(k)$, which is the input image. The index k ranges over all pixels in the image.

Step-2 To increment i , and thereafter, carry out a discrete convolution of the image with a filter h to obtain $c_{i-1}(k)$. The distance between a central pixel and adjacent ones is 2^{i-1} . Note that the filter h is based on a cubic B-spline (5×5 filter).

Step-3 To obtain the discrete wavelet transform, $w_i(k) = c_{i-1}(k) - c_i(k)$.

Step-4 To return to step 2 if i is less than the number of resolution levels wanted (let p be the number of resolution levels).

As the result, the set

$$W = \{w_0, w_1, \dots, w_p, c_p\}$$

is produced, which represents the wavelet transform of the image, where c_p denotes a residual. Therefore, the following additive decomposition can be applied to the input image, $c_0(k)$:

$$c_0(k) = c_p + \sum_{i=1}^p w_i(k).$$

Fig. 1.20 shows a point pattern set, which is the simulated Gaussian cluster with 300 and 259 points, and background Poisson noise with 300 points. Fig. 1.21 presents the corresponding wavelet transform. Wavelet scales 1-6 are shown in sequence, left to right, starting at the upper left corner.

Images and sets of point patterns generally contain noise. Thus, the wavelet coefficients are noisy too. In most applications, it is necessary to know if a wavelet coefficient is produced from signal (i.e. it is significant) or due to the noise. [Murtagh and Starck, 1998] develops a statistical significance test to treat this problem. It defines that

$$P = \text{Prob}(|w_j| < \tau),$$

were τ denotes the detection threshold, which is defined for each scale. Given an estimation threshold, ϵ , if $P < \epsilon$, the wavelet coefficient value cannot be due to the noise along, and a significant wavelet coefficient can be detected.

The multiresolution support can be obtained by detecting the significant coefficients at each scale level. The multiresolution support is defined as

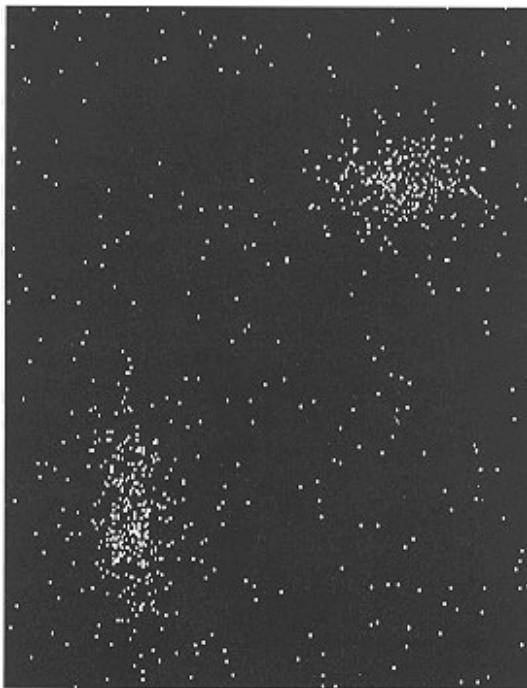


Fig. 1.20 An example of the point pattern [Murtagh and Starck, 1998].

follows [Starck et al., 1995]:

$$M(j, x, y) = \begin{cases} 1 & \text{if } w_j(x, y) \text{ is significant,} \\ 0 & \text{if } w_j(x, y) \text{ is not significant.} \end{cases}$$

The algorithm to create the multiresolution support is presented below:

- Step-1** To compute the wavelet transform of the image;
- Step-2** To estimate the noise standard deviation at each scale, and thereafter, deduce the statistically significant coefficients at each scale level;
- Step-3** To booleanize each scale, which can lead to the multiresolution support;
- Step-4** To modify using a priori knowledge if desired.

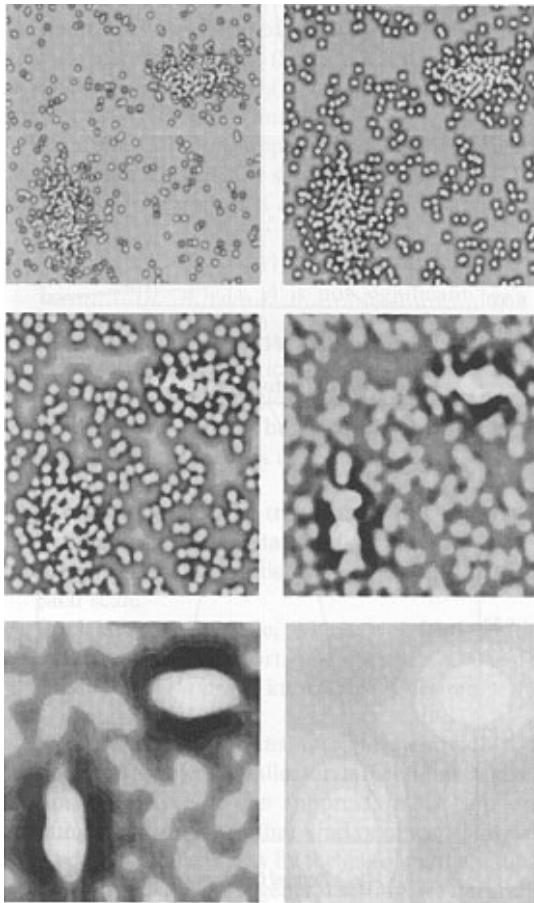


Fig. 1.21 Wavelet transform of the point pattern with scales of 1-6 [Murtagh and Starck, 1998].

In order to visualize the multiresolution support, we can create an image S defined by

$$S(x, y) = \sum_{j=1}^p 2^j M(j, x, y). \quad (1.6)$$

Fig. 1.22 is related to the multiresolution support image of the 5th scale image in Fig. 1.21. [Murtagh and Starck, 1998] uses this method to carry



Fig. 1.22 The multiresolution support image at the 5th scale level of the wavelet transform of the point pattern shown in Fig. 1.20 [Murtagh and Starck, 1998].

out three examples: (1) excellent recovery of Gaussian clusters, (2) diffuse rectangular cluster, and (3) diffuse rectangle and fuzzier Gaussian clusters. The results show that the computational complexity is $O(1)$.

1.2.5 Document Analysis with Wavelets

Document processing is one of the most active branches in the area of pattern recognition. Documents contain knowledge. Precisely, they are medium for transferring knowledge. In fact, much knowledge is acquired from documents such as technical reports, government files, newspapers, books, journals, magazines, letters, bank cheques, to name a few. The acquisition of knowledge from such documents by an information system can involve an extensive amount of hand-crafting. Such hand-crafting is time-consuming and can severely limit the application of information systems. Actually, it is a bottleneck of information systems. Thus, automatic

knowledge acquisition from documents has become an important subject. Since the 1960's, much research on document processing has been done [Tang et al., 1994]. Recently, the wavelet theory has been employed in this research [Liang et al., 1999; Tang et al., 1995a; Tang et al., 1996a; Tang et al., 1997a; Tang et al., 1997c].

In this book, we have a chapter to give a detailed presentation of wavelet-based document processing. In this sub-section, another achievement is presented below:

1. Form-Document Analysis by Reference Line Detection with 2-D Wavelet Transform

The major characteristics of forms are analyzed in [Tang et al., 1997a]:

- In general, a form consists of straight lines, which are oriented mostly in horizontal and vertical directions. These lines are referred to as *reference lines*.
- The reference lines are pre-printed to guide the users to complete the form.
- The information that should be entered to computer and processed is usually the filled data.
- In order to indicate the filling position, the reference lines can be used and the filled information usually appears either above, beneath, or beside these reference lines. Thus, in form processing, the reference lines have to be detected first, then we can find the useful information from a form based on them, and thereafter, enter to the computers.

In [Tang et al., 1997a], a novel wavelet-based method is presented. In this method, two-dimensional multiresolution analysis (MRA), wavelet decomposition algorithm, and compactly supported orthonormal wavelets are used to transform a document image into several sub-images. Based on these sub-images, the reference lines of a complex-background document can be extracted, and knowledge about the geometric structure of the document can be acquired. Particularly, this approach appears to be more efficient in processing form documents with multi-grey level background.

A document image can be transformed into four sub-images by applying the Mallat algorithm, namely, (1) LL sub-image, (2) LH sub-image, (3) HL sub-image, and (4) HH sub-image. We are interested in the LH and HL sub-images. The LH sub-image is achieved from a filter which allows

lower frequency components to reach across along the horizontal direction, as well as the higher frequencies along the vertical direction. That is an “enhancing” effect on the vertical, and “smoothing” effect on the horizontal. As a result, only horizontal lines remain in the LH sub-image. The situation of the HL sub-image is opposite to that of the LH one. In this way, the horizontal direction of the filter opens for the higher frequencies, and the vertical direction for lower frequency components. That is an “enhancing” effect on the horizontal, and “smoothing” effect on the vertical. Thus, only vertical lines remain in the HL sub-image.

2. Multiresolution Hadamard Representation and Its Application to Document Image Analysis

A novel class of wavelet transform referred to as the multiresolution Hadamard representation (MHR) is proposed by [Liang et al., 1999] for document image analysis.

The multiresolution Hadamard representation (MHR) is a 2-D dyadic wavelet representation which employs two Hadamard coefficients $[1, 1, 1, 1]$ and $[1, -1, -1, 1]$ as shown in Fig. 1.23. These coefficients are further normalized with respect of l^1 norm [Mallat, 1989c].

The 2-D dyadic wavelet representation is produced by applying the 1-D filters $h(\cdot)$ and $g(\cdot)$ to the 2-D image in both horizontal and vertical directions. This representation comprises four channels, namely, low-passed L , horizontal H , vertical V and diagonal D , at each level of transform. These channels can be defined by the following iterative formulas:

$$\begin{aligned} L_{u,v,j+1} &= \int_x \int_y L_{u,v,j} h(x-2u) h(y-2v), \\ H_{u,v,j+1} &= \int_x \int_y L_{u,v,j} h(x-2u) g(y-2v), \\ V_{u,v,j+1} &= \int_x \int_y L_{u,v,j} g(x-2u) h(y-2v), \\ D_{u,v,j+1} &= \int_x \int_y L_{u,v,j} g(x-2u) g(y-2v), \end{aligned}$$

where x and u are horizontal coordinates, y and v are vertical ones. Note that when $j = 0$, $L_{u,v,j}$ denotes the original image.

The $L_{(j+1)}$ channel is obtained by the convolution of L_j with the 2-D filter $h(x)h(y)$ shown in Fig. 1.24. The $H_{(j+1)}(V_{(j+1)})$ channel is produced by the convolution of L_j with the 2-D filter $h(x)g(y)(h(y)g(x))$. Since the

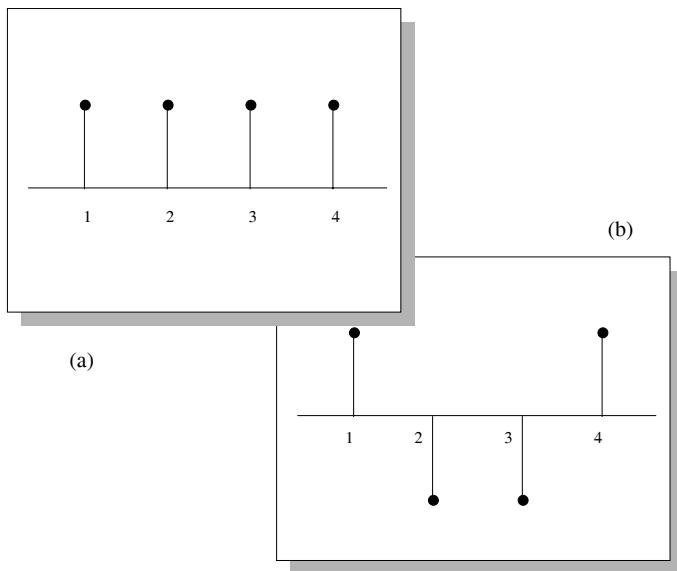


Fig. 1.23 The basic functions of the multiresolution Hadamard representation (MHR): (a) The low-pass filter $h(\cdot)$; (b) The high-pass filter $g(\cdot)$ [Liang et al., 1999].

shape of $h(x)g(y)(h(y)g(x))$ is similar to a horizontal (vertical) bar, this 2-D filter serves as a detector for horizontal (vertical) bars on L_j , which can be graphically illustrated in Fig. 1.25. The $D_{(j+1)}$ channel is obtained by the convolution of L_j with the 2-D filter $g(x)g(y)$. It is displayed graphically in Fig. 1.26.

In [Liang et al., 1999], the multiresolution Hadamard representation (MHR) is applied to document image analysis including the following processes:

- the exactation of half-tone picture,
- segmentation of document image into text blocks, and
- determination of character scales for each text block.

The transformed values of the vertical strokes of characters are very positive in some of the V channels, while, the horizontal strokes of characters react strongly in the H channels. The diagonal strokes of characters, on the other hand, react in both H and V channels. The pictures in newspapers are generally produced by half tones, where the tone of the pictures is

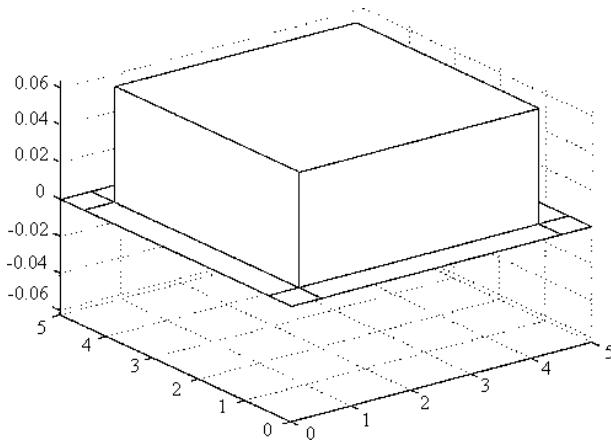


Fig. 1.24 The L channels of the 2-D filters $h(\cdot)g(\cdot)$ [Liang et al., 1999].

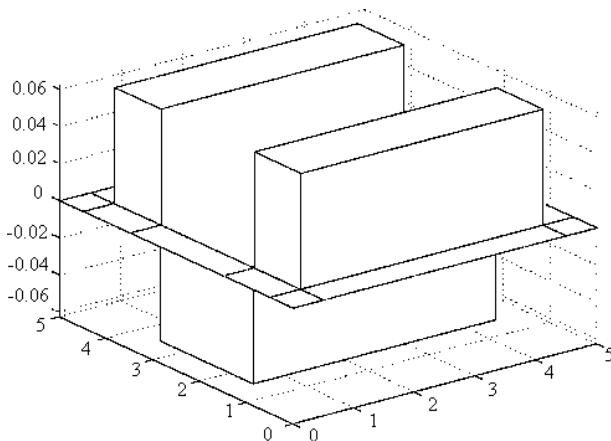


Fig. 1.25 The H and V channels of the 2-D filters $h(\cdot)g(\cdot)$ [Liang et al., 1999].

produced by small block dots with varying densities. While characters show up their strengths in the H and V channels, a half-tone picture reacts as a significant regular pattern in the D channel. This pattern is produced when the half-tone pictures are filtered by $(g(x)g(y))$.

In [Liang et al., 1999], the multiresolution Hadamard representation (MHR) is used to treat a portion of Chinese newspaper as shown in Fig. 1.27.

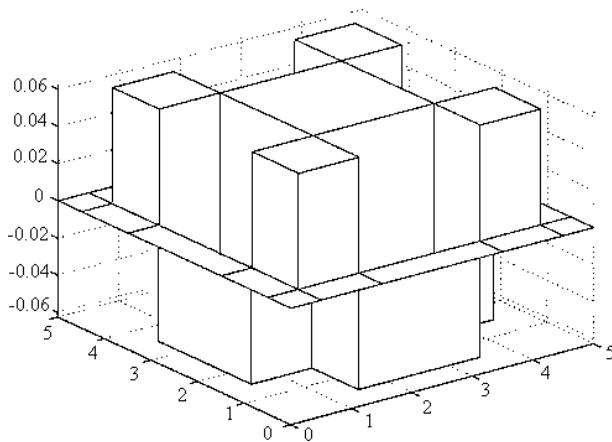


Fig. 1.26 The D channels of the 2-D filters $h(\cdot)g(\cdot)$ [Liang et al., 1999].



Fig. 1.27 A portion of Chinese newspaper [Liang et al., 1999].

It contains Chinese characters of three different sizes. The MHR developed in [Liang et al., 1999] picks up three text blocks with different scales from the original image, which are presented in Fig. 1.28. The multiresolution Hadamard representation of Fig. 1.27 is illustrated in Fig. 1.29, where the scale is $j = 1$.

The detailed description can be referred to [Liang et al., 1999].



Fig. 1.28 Three text blocks with different scales [Liang et al., 1999].

1.2.6 Analysis and Detection of Singularities with Wavelets

A significant application of wavelet theory is the analysis of singularities. Many methods based on wavelet theory have been developed to analyze the properties of the singularities and detect them from various signals/images, and some examples can be found in [Chen et al., 1995; Chen and Yang, 1995; Chuang and Kuo, 1996; Deng and Lyengar, 1996; IEEE, 1993; Law et al., 1996; SPIE, 1994; Tang et al., 1997c; Tang et al., 1998d; Thune et al., 1997; Tieng and Boles, 1997a; Young, 1993]. The edge is one class of singularities, which commonly appear in both the one-dimensional signals and two-dimensional images. The subject of wavelet transform is a remarkable mathematical tool to analyze the singularities including the edges, and further, to detect them effectively. A significant study related to this research topic has been done by Mallat, Hwang and Zhong, and published in the Special Issue on Wavelet Transforms and Multiresolution Signal Analysis of the IEEE Trans. on Information Theory [Mallat and Hwang, 1992] and IEEE Trans. on Pattern Analysis and Machine Intelligence [Mallat and Zhong, 1992].

Corners the special type of edges, which are very attractive features for many applications in pattern recognition and computer vision. In [Chen



Fig. 1.29 The multiresolution Hadamard representation of Fig. 1.27 at scale $j = 1$ [Liang et al., 1999].

et al., 1995], a new gray-level corner detection algorithm based on the wavelet transform is presented. The wavelet transform is used because the evolution across scales of its magnitudes and orientations can be used to characterize localized signals like edges including the corners. Most conventional corner detectors detect corners based on the edge detection information. However, these edge detectors perform poorly at corners, adversely affecting their overall performance. To overcome this drawback, [Chen et al., 1995] first proposes a new edge detector based on the ratio of the inter-scale wavelet transform modulus. This edge detector can correctly detect edges at the corner positions, making accurate corner detection possible. To reduce the number of points required to be processed, it applies the non-minima suppression scheme to the edge image and extract the minima image. Based on the orientation variance, these non-corner edge points are eliminated. In order to locate the corner points, it proposes a new corner indicator based on the scale invariant property of the corner orientations. By examining the corner indicator the corner points can be located accurately, as shown by experiments with the algorithm. In addition, since wavelet transform possesses the smoothing effect inherently, this algorithm is insensitive to noise contamination as well.

1. Edge Detection with Local Maximal Modulus of Wavelet Transform

A significant study related to this research topic is done by Mallat, Hwang and Zhong [Mallat and Hwang, 1992; Mallat and Zhong, 1992]. Many important contributions are made in their papers. They have proven that the maxima of the wavelet transform modulus can detect the locations of the irregular structures. Further, a numerical procedures to calculate their Lipshitz exponents is provided. It also numerically shows that one- and two-dimensional signals can be reconstructed, with a good approximation, from the local maxima of their wavelet transform modulus. The algorithm of the edge detection with local maximal modulus of wavelet transform is presented below:

Algorithm 1.1 Given an input digital signal $\{f(k, l) | k = 0, 1, \dots, K; l = 0, 1, \dots, L\}$

Step 1 To calculate the modulo of its wavelet transform

$$\{M_s f(k, l) | k = 0, 1, \dots, K; l = 0, 1, \dots, L\}$$

as well as the codes

$$\{CodeA_s f(k, l) | k = 0, 1, \dots, K; l = 0, 1, \dots, L\}$$

along the gradient directions;

Step 2 To take a threshold $T > 0$, for $k = 0, 1, \dots, K; l = 0, 1, \dots, L$, if

- (1) $|M_s f(k, l)| \geq T$,
- (2) $|M_s f(k, l)|$ reaches its local maximum along the gradient direction represented by $CodeA_s f(k, l)$,

then, (k, l) is an edge pixel.

2. Detection of Step-Structure Edges by Scale-Independent Algorithm and MASW Wavelet Transform

The local maxima modulus of the wavelet transform can provide enough information for analyzing the singularities, and can detect all singularities. However, it may not identify different structures of singularities [Mallat and Hwang, 1992; Mallat and Zhong, 1992].

In [Tang et al., 1998c], an important property is proved that the modulus of wavelet transform at each point of the step edge is a non-zero constant

which is independent on both the gradient direction and the scale of the wavelet transform. Thus, a novel algorithm called scale-independent algorithm is developed.

Algorithm 1.2 Given two-dimensional signal $f(x, y)$

Step 1: To take different scales s_1, \dots, s_J , and calculate $W_{s_j} f(x, y)$, ($1 \leq j \leq J$) based on

$$W_s^1 f(n, m) = \sum_{k,l} f(n - 1 - k, m - 1 - l) \psi_{k,l}^{s,1},$$

$$W_s^2 f(n, m) = \sum_{k,l} f(n - 1 - k, m - 1 - l) \psi_{k,l}^{s,2}.$$

Step 2: To select peak-threshold T , such that

$$|\nabla W_{s_j} f(x, y)| \geq T.$$

Step 3: To select proportional threshold R , such that

$$\frac{1}{R} \leq \frac{|\nabla W_{s_j} f(x, y)|}{|\nabla W_{s_l} f(x, y)|} \leq R, \quad (1 \leq j \leq J).$$

This method possesses an important property, i.e. the wavelet transform of a step-structure edge is scale-independent. It can improve the method proposed in [Mallat and Hwang, 1992; Mallat and Zhong, 1992], where the modulus-angle-separated-wavelet (MASW) is used. The precise definition of the MASW can be found in [Tang et al., 1998c; Tang et al., 1998d]. After applying the scale-independent algorithm to the images, only the contour of the aircraft is extracted, while all other edges including drawing lines and texts are eliminated.

In this book, we have a specific chapter to give a detailed example of this application.

1.2.7 Wavelet Descriptors for Shapes of the Objects

Shape of a pattern is one of the most important features in pattern recognition. The description of such a shape plays a key role in shape analysis. Many research projects [Chuang and Kuo, 1996; Hsieh et al., 1995; Tieng and Boles, 1997b; Wunsch and Laine, 1995] present some shape descriptors, which can represent digitized patterns. These descriptors are

derived from the wavelet transform of the contours of a pattern and particularly well-suited for the recognition of two-dimensional objects, such as handprinted characters. Three examples are presented below:

1. Wavelet Descriptor of Planar Curves: Theory and Applications

By using the wavelet transform, [Chuang and Kuo, 1996] develops a hierarchical planar curve descriptor that decomposes a curve into components of different scales so that the coarsest scale components carry the global approximation information while the finer scale components contain the local detailed information. It shows that the wavelet descriptor has many desirable properties such as multiresolution representation, invariance, uniqueness, stability, and spatial localization. A deformable wavelet descriptor is also proposed by interpreting the wavelet coefficients as random variables. The applications of the wavelet descriptor to character recognition and model-based contour extraction from low SNR images are examined. Numerical experiments are performed to illustrate the performance of the wavelet descriptor.

2. Wavelet Descriptors for Multiresolution Recognition Of Handprinted Characters

Paper [Wunsch and Laine, 1995] presents a novel set of shape descriptors that represents a digitized pattern in concise way and that is particularly well-suited for the recognition of handprinted characters. The descriptor set is derived from the wavelet transform of a pattern's contour. The approach is closely related to feature extraction methods by Fourier Series expansion. The motivation to use an orthonormal wavelet basis rather than the Fourier basis is that wavelet coefficients provide localized frequency information, and that wavelets allow us to decompose a function into a multiresolution hierarchy of localized frequency bands. This paper describes a character recognition system that relies upon wavelet descriptors to simultaneously analyze character shape at multiple levels of resolution. The system was trained and tested on a large database of more than 6000 samples of handprinted alphanumeric characters. The results show that wavelet descriptors are an efficient representation that can provide for reliable recognition in problems with large input variability.

3. Wavelet-Based Shape form Shading

Paper [Hsieh et al., 1995] proposes a wavelet-based approach for solving the shape from shading (SFS) problem. The proposed method takes

advantage of the nature of wavelet theory, which can be applied to efficiently and accurately represent “things,” to develop a faster algorithm for reconstructing better surfaces. To derive the algorithm, the formulation of Horn and Brooks ((Eds.) *Shape from Shading*, MIT Press, Cambridge, MA,1989), which combines several constraints into an objective function, is adopted. In order to improve the robustness of the algorithm, two new constraints are introduced into the objective function to strengthen the relation between an estimated surface and its counterpart in the original image. Thus, solving the SFS problem becomes a constrained optimization process. Instead of solving the problem directly by using Euler equation or numerical techniques, the objective function is first converted into the wavelet format. Due to this format, the set of differential operators of different orders, which is involved in the whole process, can be approximated with connection coefficients of Daubechies bases. In each iteration of the optimization process, and appropriate stem size, which can result in maximum decrease of the objective function, is determined. After finding correct iterative schemes, the solution of the SFS problem can finally be decided. Compared with conventional algorithms, the proposed scheme is a great improvement in the accuracy as well as the convergence speed of the SFS problem. Experimental results, using both synthetic and real images, prove that the proposed method is indeed better than traditional methods.

4. Representation of 2-D Pattern by 1-D Wavelet Sub-patterns

Tang [Tang et al., 1998a] presents an approach to represent a 2-D shape by several 1-D wavelet sub-patterns. In this way, first, a 2-D pattern is converted into an 1-D curve by the dimensionality reduction [Tang et al., 1991]. Thereafter, according to the wavelet orthonormal decomposition, the 1-D curve can be decomposed orthogonally into several high-frequency sub-curves and low-frequency ones using the wavelet transform.

5. Wavelet Descriptors for Multiresolution Recognition Of Handprinted Characters

Paper [Wunsch and Laine, 1995] presents a novel set of shape descriptors that represents a pattern in concise way and that is particularly well-suited for the recognition of handprinted characters. The descriptors are derived from a contour of a pattern using wavelet transform. This method is closely related to the feature extraction by Fourier series expansion. The motivation to use an orthonormal wavelet basis rather than the Fourier one is

that wavelet coefficients provide localized frequency information, and that wavelets allow us to decompose a function into a multiresolution hierarchy of localized frequency bands. This paper describes a character recognition system where the wavelet descriptors are used to analyze character shape at multiple levels of resolution. A large set of samples of handprinted alphanumeric characters are used to train and test this system. The results show that wavelet descriptors can efficiently represent the shapes of the objects, and can provide for reliable recognition with large input variability.

1.2.8 *Invariant Representation of Patterns*

The invariance characteristic of size, orientation and translation is a significant subject in pattern recognition. With such capability, a pattern recognition system can find many applications in character recognition, computer vision, document processing, and many others. A lot of researchers paid attention to this subject, many methods have been developed. Recently, new methods in accordance with the wavelet-based approach have been made [Haley and Manjunath, 1999; Shen and Ip, 1999; Tieng and Boles, 1997b; Yoon et al., 1998]. In this sub-section, we introduce three papers, where the wavelet approach has been used [Shen and Ip, 1999; Tang et al., 1998a; Yoon et al., 1998].

1. Extraction of Rotation-Invariant Feature by Ring-projection-wavelet-fractal Method

In our previous work [Tang et al., 1998a], a novel approach to extract features with property of rotation-invariant in pattern recognition is presented, that utilizes ring-projection-wavelet-fractal signatures (RPWFS). In particular, this approach reduces the dimensionality of a two-dimensional pattern by way of a ring-projection method, and thereafter, performs Daubechies' wavelet transform on the derived one-dimensional pattern to generate a set of wavelet transformed sub-patterns, namely, curves that are non-self-intersecting. Further from the resulting non-self-intersecting curves, the divider dimensions are readily computed. These divider dimensions constitute a new feature vector for the original two-dimensional pattern, defined over the curves' fractal dimensions.

An overall description of this approach can be illustrated by a diagram shown in Fig. 1.30. The detailed description of this method can be found in Chapter 9 in this book.

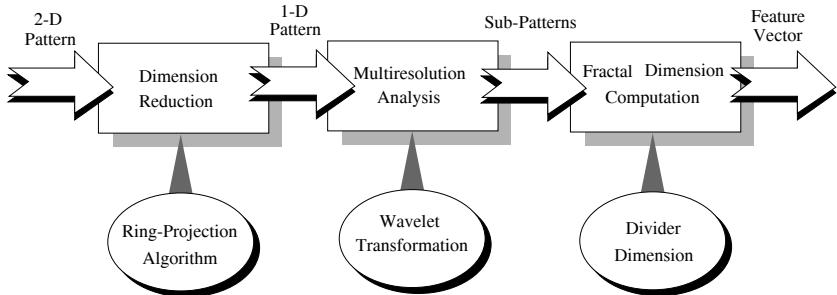


Fig. 1.30 Diagram of Ring-projection-wavelet-fractal method.

2. Wavelet Rotation-Invariant Shape Descriptors for Recognition of 2-D Pattern

In a pattern recognition system, typically, a set of numerical features are extracted from an image. The selection of discriminative features is a crucial step in the system. The use of moment invariants as features for identification of 2D shape has received much attention. [Shen and Ip, 1999] investigates a set of wavelet rotation invariant moments presented for capturing global and local information from the objects of interest, together with a discriminative feature extraction method, for the classification of seemingly similar 2D objects with subtle differences.

To achieve rotation invariant moments, typically, a generalized expression is used, which can be written by

$$F_{pq} = \int \int f(r, \theta) g_p(r) e^{jq\theta} r dr d\theta, \quad (1.7)$$

where F_{pq} denotes the pq -order moment, and $g_p(r)$ stands for a function of radial variable r . In addition, p and q are integers. It has been proved that

- The value of $\|F_{pq}\|$ is rotation invariant, where $\|F_{pq}\| = \sqrt{F_{pq} \cdot F_{pq}^*}$, and the symbol $*$ denotes conjugate of complex number.
- The combined moments, such as $F_{p_i q} \cdot F_{p_j q}^*$ are also rotation invariant.

To facilitate the analysis, an expression of 1D sequence is substituted for that of the 2D image. Thus, Eq. 1.7 becomes

$$F_{pq} = \int S_q(r) \cdot g_p(r) r dr, \quad (1.8)$$

where $S_q(r) = f(r, \theta)e^{jq\theta}d\theta$. [Shen and Ip, 1999] treats $\{g_p(r)\}$ in Eq. 1.8 as wavelet basis functions:

$$\psi^{a,b}(r) = \frac{1}{\sqrt{a}}\left(\frac{r-b}{a}\right).$$

That indicates that the basis functions $\{g_p(r)\}$ are replaced by wavelet basis functions $\{\psi^{a,b}(r)\}$. The mother wavelet used in this work is a cubic B-spline in Gaussian approximation form (Fig. 1.31):

$$\begin{aligned}\psi(r) &= \frac{4a^{n-1}}{\sqrt{2\pi(n+1)}}\sigma_w \cos(2\pi f_0(2r-1)) \\ &\times \exp\left(-\frac{(2r-1)^2}{2\sigma_w^2(n+1)}\right),\end{aligned}$$

where $n = 3$, $a = 0.697066$, $f_0 = 0.409177$ and $\sigma_w^2 = 0.561145$. The details of this mother wavelet can be fund in [Unser et al., 1992]. Let $a = 0.5^m$, $m = 0, 1, 2, 3$, and $b = 0.5n0.5^m$, $n = 0, 1, \dots, 2^{m+1}$. Thereafter, the wavelet defined along a radial axis in any orientation is denoted by

$$\psi_{m,n}(r) = 2^{m^2}\psi(2^mr - 0.5n)$$

Consequently, a set of wavelet moment invariants for classifying objects

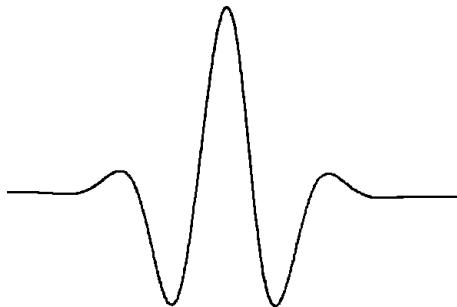


Fig. 1.31 The cubic B-spline mother wavelet [Shen and Ip, 1999].

can be defined as follows:

$$\|F_{m,n,q}^{wavelet}\| = \left\| \int S_q(r) \cdot \psi_{m,n}(r) r dr \right\|, \quad (1.9)$$

where $\psi_{m,n}(r)$ replaces $g_g(r)$ in Eq. 1.8. The parameters are: $m = 0, 1, 2, 3$, $n = 0, 1, \dots, 2^{m+1}$, and $q = 0, 1, 2, 3$. It can be fund that $\|F_{m,n,q}^{\text{wavelet}}\|$ is actually a wavelet transform of $S_q(r)r$. $\|F_{m,n,q}^{\text{wavelet}}\|$ can also be considered to be the first moment of $S_q(r)$ at the m th scale level with shift index n . It has been proved that the wavelet moment invariants, $\|F_{m,n,q}^{\text{wavelet}}\|$, are invariant to rotation of a object.

In [Shen and Ip, 1999], the wavelet moment invariants along with a minimum-distance classifier are used, and a high classification rate for four different sets of patterns are achieved. For instance, 100% recognition rate is obtained for a test set consisting of 26 upper cased English letters with different scales and orientations.

3. Scale-Invariant Object Recognition Based on the Multiresolution Approximation

Boundary-based pattern matching is one of the most useful approaches in pattern recognition. [Yoon et al., 1998] presents a multiresolution approximation approach to obtaining boundary representation for object recognition. The summary of this approach is presented below:

It establishes the theory of the continuous multiresolution approximation (CMA), and implements a fast algorithm for the continuous wavelet transform (CWT). The CMT is a vehicle for scale-invariant matching or coarse-to-fine matching. In addition, the fast algorithm for the CWT enables us to quickly compute a representation. [Yoon et al., 1998] proposes good representations for boundary-based object matching. The representations are obtained using the CMA. These representations allow us to recognize objects in the presence of noise, occlusion, scale variation, rotation, and translation. It tests various types of objects such as tools, guns, maps, etc., with occlusion and scale variations. It also tests those objects by adding various types of noise. The test results show the proposed representations are reliable and consistent.

Object matching in the presence of noise, occlusion, and scale variations is considered the most difficult problem in the area of boundary-based object matching. [Yoon et al., 1998] call this scale-invariant matching. To solve this problem, the scaling effect of an object by using the CWT is modeled. The model enables us to use the CWT for scale-invariant matching. Further, a scale-invariant representation is proposed to handle the scale-invariant matching.

Firstly, the modeling scaling effect of an object by using the wavelet

transform is introduced below:

When an analog signal is converted to a digital one, the analog signal is pre-filtered by a lowpass filter to reduce aliasing and it is then sampled as shown in Fig. 1.32. An object in a picture may have different scales as the distance of the camera from the target object changes. By means of the wavelet transform, the scaling effect of an object can be modeled, which can lead to generating a scale-invariant representation for the scaling effect. The mathematical description can be presented as follows:

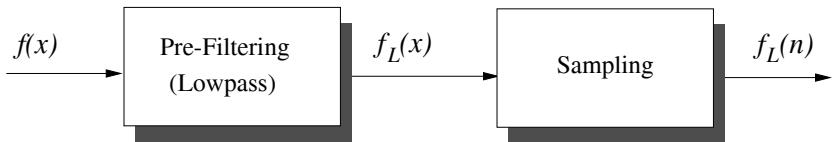


Fig. 1.32 Digitization of analog signal as pre-filtering and sampling [Yoon et al., 1998].

Let $h(x)$ be the lowpass filter, thus, the output, $f_L(x)$, of the filtering of an original signal $f(x)$ is

$$f_L(x) = f(x) \cdot h(x) = \int f(\alpha)h(\alpha - x)d\alpha.$$

The filtered output of the scaled signal is

$$f_L(sx) = \frac{1}{s}f(sx) \cdot h(x) = \int \frac{1}{s}f(s\alpha)h(\alpha - x)d\alpha. \quad (1.10)$$

Setting $\alpha' = s\alpha$, Eq. 1.10 becomes

$$f_L(sx) = \int \frac{1}{s}f(\alpha') \cdot h\left(\frac{\alpha'}{s} - x\right) \frac{d\alpha'}{s} = f(x) \cdot \frac{1}{s}h\left(\frac{x}{s}\right). \quad (1.11)$$

Eq. 1.11 is the same formulation as the wavelet transform at the scale s , the filter $h(x)$ is considered to be a continuous scaling function. This formulation implies that the scaling effect of an object in the process of digitization of an analog input image can be represented by the wavelet transform. This idea is the motivation for using the wavelet transform for a scale-invariant representation.

Secondly, the scale-invariant representation is established by means of the continuous multiresolution approximation (CMA). The CMA has two

important features to receive significant representations for pattern recognition, namely, (1) the CMA provides approximations of objects at various scales, hence, scale-invariant representations can be constructed by using the approximations; (2) the representations can efficiently computed by using the fast algorithm.

The approximation, $\{R_s(k), Q_s(k)\}$ of a boundary at the scale 2^s is mathematically described as follows:

$$R_s(k) = [r(t) \cdot \beta_s^3(-t)]_{t=pk}, \quad (1.12)$$

$$Q_s(k) = [q(t) \cdot \beta_s^3(-t)]_{t=pk}, \quad (1.13)$$

where p denotes the sampling period in the CWT. Applying Eq. 1.13 to an original boundary $\{r(t), q(t)\}$ produces a pair of functions for an approximation of a boundary $\{R_s(k), Q_s(k)\}$ at each scale. The example of an approximation is illustrated in Fig. 1.33. An original boundary is shown in Fig. 1.33(a), and an approximation of the original boundary at one-half of the scale is presented in Fig. 1.33(b). The curvature function of the gun is displayed in Fig. 1.33(c). The curvature function is invariant under scaling, rotation, and translation of a curve. Moreover, zero crossings of the curvature function are important features for shape analysis or recognition. Therefore, we can use the zero crossing of the curvature functions of the approximations using the CWT to construct the proposed representations.

The detailed procedure to receive the scale invariant representation can be presented by the following algorithm:

Step-1 Interpolation: To interpolate an original boundary with the linear interpolation filter $h_1(n)$ given by

$$M(j, x, y) = \begin{cases} 1 - \frac{|n|}{L} & |n| < L, \\ 0 & \text{otherwise.} \end{cases}$$

where L is an interpolation length between the original samples.

Step-2 Initialization: To apply the dilated cubic B-spline filter for initialization to obtain low-resolution boundaries at scales of $0.5 < s < 1$. The low-resolution boundaries of the interpolated signal is given by

$$R_s(k) = [r(t) \cdot 2^s \beta^3(-2^s t)]_{t=2^{-s}k}.$$

Step-3 Decomposition: To obtain the low-resolution boundaries by use

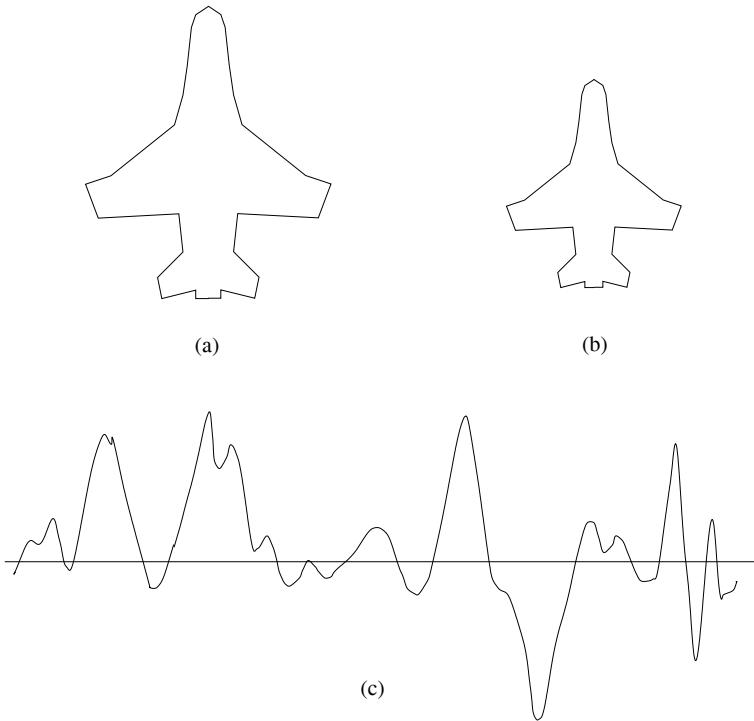


Fig. 1.33 The example of an approximation of a gun boundary.

of the DWT, and is presented below:

$$R_s(k) = [R_{s-1} \cdot \beta_2^3(k)]_{\downarrow 2}$$

where the discrete filter $\beta_2^3(k)$ is a binomial one.

Step-4 Construction of scale-invariant representation: To compute the curvature functions and find zero crossings of the curvature functions over the desired scales. Moreover, to construct the scale-invariant representation by making zero crossing on the $t - k$ plane.

Experiments are conducted in [Yoon et al., 1998]. Several images, such as guns, tools, maps, etc. are used. Some objects are taken from a camera, and some, for example, guns are from the X-ray. Maps are taken from a

commercial image database. The results have shown the proposed representations are reliable and consistent.

1.2.9 *Handwritten and Printed Character Recognition*

Character recognition including the identification of handwritten and printed characters is a major branch in the field of pattern recognition. A quite a number of articles, which deal with this branch, have been published. However, only a few have used wavelets [Lee et al., 1996; Tang et al., 1996b; Tang et al., 1998d; Tang et al., 1998a; Wunsch and Laine, 1995]. In this sub-section, two publications [Lee et al., 1996; Wunsch and Laine, 1995] are briefly introduced. In addition, a specific chapter in this book is available to provide a detailed description of the character recognition with the wavelet theory.

1. Wavelet Descriptors for Recognition of Handprinted Characters

Paper [Wunsch and Laine, 1995] describes a character recognition system that relies upon wavelet descriptors to simultaneously analyze character shape at multiple levels of resolution. The system was trained and tested on a large database of more than 6000 samples of handprinted alphanumeric characters. The results show that wavelet descriptors are an efficient representation that can provide for reliable recognition in problems with large input variability.

2. Extracting Multiresolution Features in Recognition of Handwritten Numerals with 2-D Haar Wavelet

The well-known Haar wavelet is adequate for local detection of line segments and global detection of line structures with fast computation. [Lee et al., 1996] develops a method based on the Haar wavelet. It enables us to have an invariant interpretation of the character image at different resolutions and presents a multiresolution analysis in the form of coefficient matrices. Since the details of character image at different resolutions generally characterize different physical structures of the character, and the coefficients obtained from wavelet transform are very useful in recognizing unconstrained handwritten numerals. Therefore, in [Lee et al., 1996], wavelet transform with a set of Haar wavelets is used for multiresolution feature extraction in handwriting recognition.

In this way, we take

$$1 = \cos^2 \frac{\omega}{2} + \sin^2 \frac{\omega}{2}.$$

Let us write

$$\begin{aligned} |H(\omega)|^2 &= \cos^2 \frac{\omega}{2} = \frac{1 + \cos \omega}{2} \\ &= \frac{1}{4} [1 + 2 \cos \omega + \cos^2 \omega + \sin^2 \omega] \\ &= \frac{1}{4} [(1 + \cos \omega)^2 + \sin^2 \omega] \\ &= \left| \frac{1 + \cos \omega - i \sin \omega}{2} \right|^2 = \left| \frac{1 + e^{-i\omega}}{2} \right|^2, \end{aligned}$$

so that

$$|H(\omega + \pi)|^2 = \sin^2 \frac{\omega}{2} = \frac{1 - \cos \omega}{2} = \left| \frac{1 - e^{-i\omega}}{2} \right|^2.$$

Thus, we can take

$$H(\omega) = \frac{1 + e^{-i\omega}}{2} = \sum_{k=0}^1 \frac{1}{\sqrt{2}} h_k e^{-i\omega k},$$

where

$$h_0 = \frac{1}{\sqrt{2}}, \quad h_1 = \frac{1}{\sqrt{2}}.$$

The scaling function $\varphi(x)$ and wavelet function $\psi(x)$ can be represented by

$$\varphi(x) = \begin{cases} 1 & \text{if } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases} \quad (1.14)$$

and

$$\begin{aligned} \psi(x) &= c_1 \varphi(2x) - c_0 \varphi(2x - 1) \\ &= \varphi(2x) - \varphi(2x - 1) \\ &= \begin{cases} 1 & \text{if } 0 < x < \frac{1}{2} \\ -1 & \text{if } \frac{1}{2} \leq x < 1 \\ 0 & \text{otherwise} \end{cases} \end{aligned} \quad (1.15)$$

An image of the handwritten character can be decomposed into its wavelet coefficients by using Mallat's pyramid algorithm. By using Haar wavelets, an image F is decomposed as follows:

$$F = \begin{bmatrix} & & \\ w & x & \\ y & z & \end{bmatrix} \Rightarrow \begin{bmatrix} p & q & & b \\ r & a & & \\ & c & & d \\ & & & \end{bmatrix} \Rightarrow \begin{bmatrix} s & t & & b \\ u & v & & \\ & c & & d \\ & & & \end{bmatrix}$$

$$\begin{aligned} a &= 1/4(w + x + y + z), & b &= 1/4(w - x + y - z) \\ c &= 1/4(w + x - y - z), & d &= 1/4(w - x - y + z) \\ s &= 1/4(p + q + r + a), & t &= 1/4(p - q + r - a) \\ u &= 1/4(p + q - r - a), & v &= 1/4(p - q - r + a) \end{aligned} \quad (1.16)$$

In Eq. (1.16), the image $\{w, x, y, z\}$ is decomposed into image $\{a\}$, $\{b\}$, $\{c\}$, and $\{d\}$ at resolution 2^{-1} . The image $\{a\}$ corresponds to the lowest frequencies (D_1), $\{b\}$ gives the vertical high frequencies (D_2), $\{c\}$ the horizontal high frequencies (D_3), and $\{d\}$ the high frequencies in horizontal and vertical directions (D_4). Likewise, the image $\{p, q, r, a\}$ at resolution 2^{-1} is decomposed into image $\{s\}$, $\{t\}$, $\{u\}$, and $\{v\}$ at resolution 2^{-2} . This decomposition can be archived by convolving the 2×2 image array with the Haar masks as follows:

$$D_k^j = \frac{1}{4} I^{j+1} \otimes H_k \quad k = 1, 2, 3, 4 \quad (1.17)$$

where, \otimes is the convolution operator, D_k^j are decomposed images at resolution 2^j , I^{j+1} is 2×2 image at resolution 2^{j+1} , and H_k are Haar masks defined in Fig. 1.34.

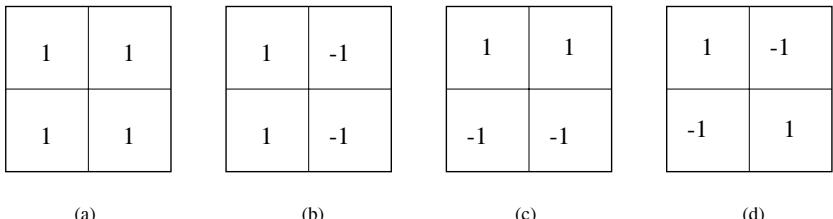


Fig. 1.34 Haar masks (a) Lowest frequencies, (b) Vertical high frequencies, (c) Horizontal high frequencies, (d) High frequencies in horizontal and vertical directions.

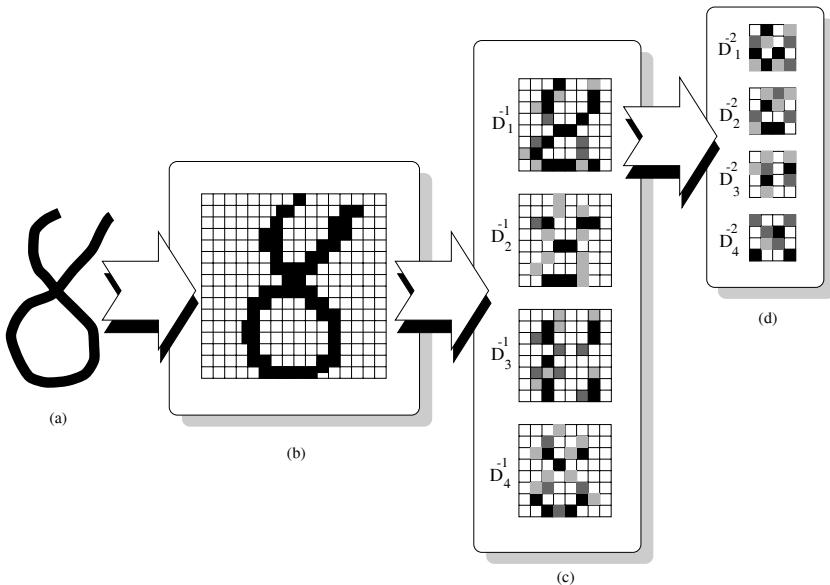


Fig. 1.35 Overview of multiresolution feature extraction.

In this application, the decomposed results at resolution 2^{-1} and 2^{-2} are used as multiresolution features . Fig. 1.35 shows the process of multiresolution feature extraction. Fig. 1.35(a) gives an input image of the handwritten number “8”, its digitized image is shown in Fig. 1.35(b). The decomposed feature vector at resolution 2^{-1} is illustrated in Fig. 1.35(c), while the decomposed feature vector at resolution 2^{-2} is in Fig. 1.35(d).

These features are used to recognize the unconstrained handwritten numerals from the database of Concordia University of Canada (Fig. 1.36), Electro-Technical Laboratory of Japan (Fig. 1.37), and Electronics and Telecommunications Research Institute of Korea (Fig. 1.38). The error rates are 3.20%, 0.83%, and 0.75%, respectively. These results are shown that the proposed scheme is very robust in terms of various writing styles and sizes. The detailed description of this application can be referred in [Lee et al., 1996].

3. Wavelet Descriptors for Recognition of Printed Kannada Text in Indian languages

An OCR system for Indian languages, especially for Kannada, a popular

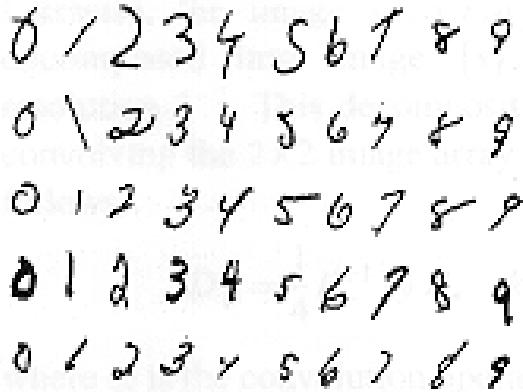


Fig. 1.36 Handwritten numeral database of Concordia University of Canada.

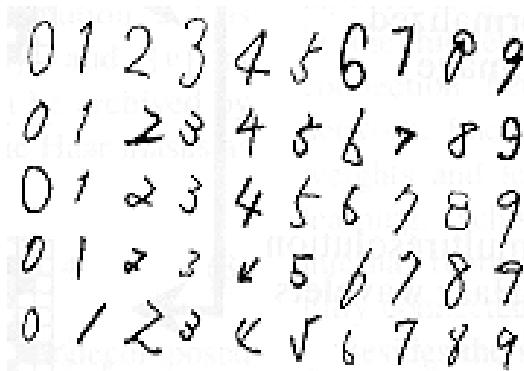


Fig. 1.37 Handwritten numeral database of Electro-Technical Laboratory of Japan.

South Indian language, is developed in [Kunte and Samuel, 2007]. Some examples of Kannada language is presented in Fig. 1.39.

1D discrete wavelet transform (DWT) is used for feature extraction. Daubechies wavelet from the family of orthonormal wavelets is considered. A DWT when applied to a sequence of coordinates from the character contour returns a set of approximation coefficients and a set of detailed coefficients. The approximation coefficients correspond to the basic shape of the contour (low frequency components) and the detailed coefficients

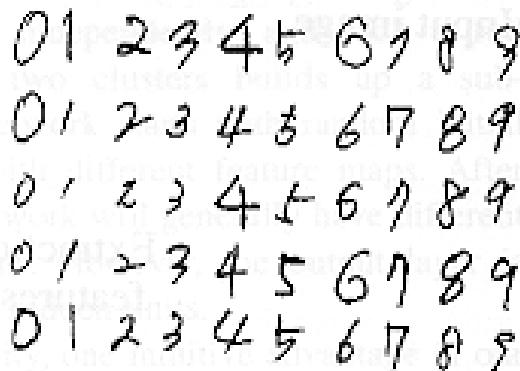


Fig. 1.38 Handwritten numeral database of Electronics and Telecommunications Research Institute of Korea.

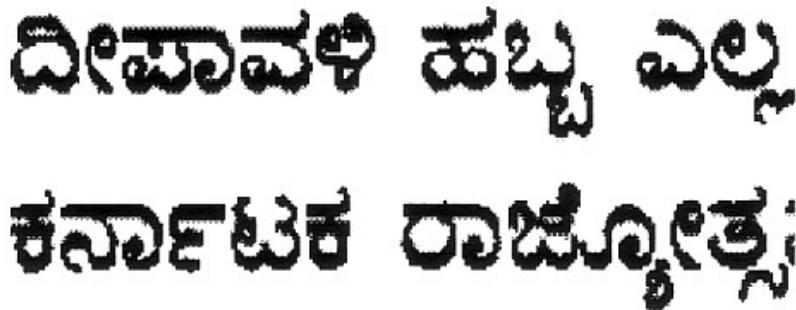


Fig. 1.39 Some examples of Kannada text in Indian languages [Kunte and Samuel, 2007].

correspond to the details of the contour (high frequency components), which reflect the contour direction, curvature, etc. Neural classifiers are effectively used for the classification of characters based on wavelet features. The system methodology can be extended for the recognition of other south Indian languages, especially for Telugu.

1.2.10 *Texture Analysis and Classification*

Texture is a specific kind of pattern, the texture analysis is one of the most important techniques used in the analysis and classification of images

where repetition or quasi-repetition of fundamental elements occurs. So far, there is no precise definition of texture. Three principal approaches are used in texture analysis, namely, statistical, spectral and structural. In this sub-section, several methods based on the wavelet theory [Haley and Manjunath, 1999; de Wouwer et al., 1999a; de Wouwer et al., 1999b] are introduced:

1. Wavelet correlation signatures for color texture characterization

In the last decade, multiscale techniques for gray-level texture analysis have been intensively used. [de Wouwer et al., 1999a] aims to extend these techniques to color images. It introduces wavelet energy-correlation signatures and derives the transformation of these signatures upon linear color space transformations. Experiments are conducted on a set of 30 natural colored texture images in which color and gray-level texture classification performances are compared. It is demonstrated that the wavelet correlation features contain more information than the intensity or the energy features of each color plane separately. The influence of image representation in color space is evaluated.

2. Rotation-Invariant Texture Classification Using a Complete Space-Frequency Model

A method of rotation-invariant texture classification based on a complete space -frequency model is introduced in [Haley and Manjunath, 1999]. A polar, analytic form of a two-dimensional (2-d) Gabor wavelet is developed, and a multiresolution family of these wavelets is used to compute information-conserving micro-features. From these micro-features, a micro-model, which characterizes spatially localized amplitude, frequency, and directional behavior of the texture, is formed. The essential characteristics of a texture sample, and its macro-features, are derived from the estimated selected parameters of the micro-model. Classification of texture samples is based on the macro-model derived from a rotation invariant subset of macro features. In experiments, comparatively high correct classification rates were obtained using large sample sets.

3. Statistical Texture Characterization from Discrete Wavelet Representation

Texture analysis plays an important role in many tasks of image processing, pattern recognition, robot vision, computer vision, etc. [de Wouwer

et al., 1999b] conjectures that the texture can be characterized by the statistics of the wavelet detail coefficients and therefore introduces two feature sets:

- The wavelet histogram signatures, which capture all first-order statistics using a model based approach;
- The wavelet co-occurrence signatures, which reflect the coefficients' second-order statistics.

The (average) best results are obtained by combining both feature sets. The introduced features are very promising for many image processing tasks such as texture recognition, segmentation, and indexing image databases.

In this sub-section, we would like to briefly introduce the basic idea of [de Wouwer et al., 1999b]. In this work, the authors combine the statistical and multiscale view on texture. They conjecture that texture can be completely characterized from the statistical properties of its multiscale representation. The first-order statistical information is derived from the detail image histogram. The detail histograms of natural textured images can be modeled by a family of exponential functions. Introducing the parameters of this model as texture features completely describes the wavelet coefficients' first-order statistics. Further, improvement in texture description is obtained from the coefficients' second-order statistics, which can be described using the detailed image co-occurrence matrices. The most complete description is obtained by combining both first- and second-order statistical information.

The 2-D discrete wavelet transform is applied in this work, which is a separable filterbank:

$$L_n(b_i, b_j) = [H_x * [H_y * L_{n-1}]_{\downarrow 2,1}]_{\downarrow 1,2}(b_i, b_j) \quad (1.18)$$

$$D_{n1}(b_i, b_j) = [H_x * [G_y * L_{n-1}]_{\downarrow 2,1}]_{\downarrow 1,2}(b_i, b_j) \quad (1.19)$$

$$D_{n2}(b_i, b_j) = [G_x * [H_y * L_{n-1}]_{\downarrow 2,1}]_{\downarrow 1,2}(b_i, b_j) \quad (1.20)$$

$$D_{n3}(b_i, b_j) = [G_x * [G_y * L_{n-1}]_{\downarrow 2,1}]_{\downarrow 1,2}(b_i, b_j) \quad (1.21)$$

where $*$ denotes the convolution operator, $\downarrow 2, 1(\downarrow 1, 2)$ subsampling along the rows (columns), and $L_0 = I(\vec{x})$ is the original image. H and G are a low and bandpass filter, respectively.

The histogram of the wavelet detail coefficients are noted as $h_{ni}(u)$; thus $h_{ni}(u)du$ is the probability that a wavelet coefficient $D_{ni}(\vec{b})$ has a value between u and $u + du$. The detail histograms of natural textured

images can be modeled by a family of exponential [Mallat, 1989b]:

$$h(u) = K e^{-(|u|/\alpha)^\beta}. \quad (1.22)$$

In this model, α and β are wavelet histogram signatures, which are easily interpreted as specific, independent characteristics of the detail histogram. They contain all first-order information present in the detail histogram. Eq. 1.22 can be employed for many texture analysis tasks.

When the features based on the first-order statistics do not suffice, the second-order statistics can improve the texture discrimination. The wavelet co-occurrence signatures can reflect the coefficients' second-order statistics. The element (j, k) of the co-occurrence matrix $C_{ni}^{\delta\theta}$ is defined as the joint probability that a wavelet coefficient $\tilde{D}_{ni} = j$ co-occurs with a coefficient $\tilde{D}_{ni} = k$ on a distance δ in direction θ . Formulas for eight common co-occurrence features are provided in [de Wouwer et al., 1999b]. These features extracted from the detail images are referred to as the wavelet co-occurrence signatures.

A database consisting of 30 real-world 512×512 images from different natural scenes is used for the experiments, which is presented in Fig. 1.42. This database is available at

<http://www.white.media.mit.edu/vismod/imagery/VisionTexture>,

which is provided by MIT Media Lab. Each image region is transformed to an overcomplete wavelet representation of depth four using a biorthogonal spline wavelet of order two [Unser et al., 1993]. Four different feature sets are generated, namely:

- 12 wavelet energy signatures,
- 24 wavelet histogram signatures,
- 96 wavelet co-occurrence signatures
- All feature from 2 and 3.

The results of the experiments can be found in [de Wouwer et al., 1999b].

4. Adaptive Scale Fixing for Multiscale Texture Segmentation

K. H. Liang and T. Tjahjadi [Liang and Tjahjadi, 2006] address two challenging issues in unsupervised multiscale texture segmentation: determining adequate spatial and feature resolutions for different regions of the image, and utilizing information across different scales/resolutions. The

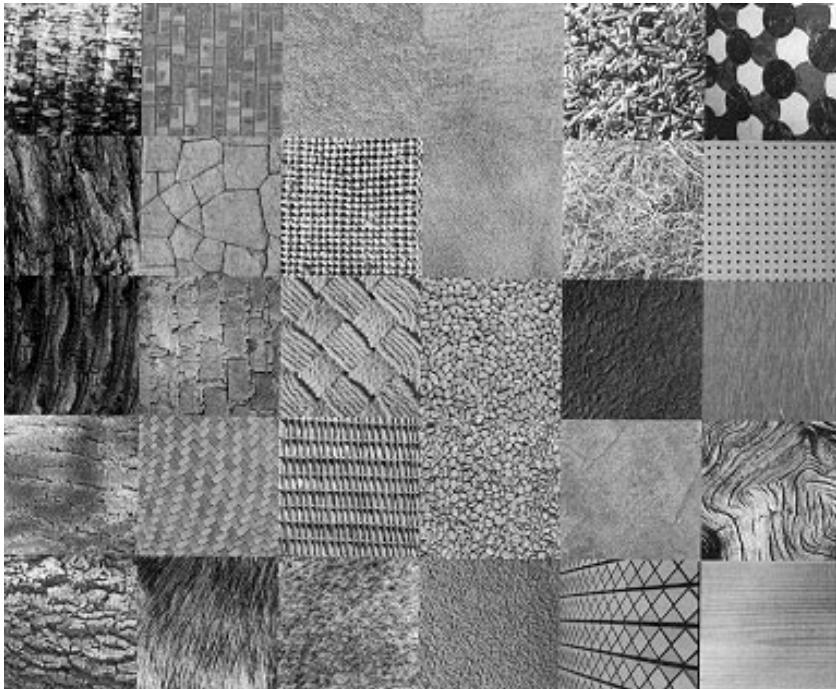


Fig. 1.40 Selection of images from the VisTex-database used in [de Wouwer et al., 1999b].

center of a homogeneous texture is analyzed using coarse spatial resolution, and its border is detected using fine spatial resolution so as to locate the boundary accurately. The extraction of texture features is achieved via a multiresolution pyramid. The feature values are integrated across scales/resolutions adaptively. The number of textures is determined automatically using the variance ratio criterion. Experimental results on synthetic and real images demonstrate the improvement in performance of the proposed multiscale scheme over single scale approaches.

Adaptive scale fixing is a method which determines the optimum Δw_s at a site s for segmentation. It also facilitates the utilization of information across resolutions. The aim is to maximize the precisions of both texture estimation and boundary localization, overcoming the limitation of single-resolution approaches. This method generates a multiscale segmentation map as shown in Fig. 1.41. Multiscale representation of a two-texture

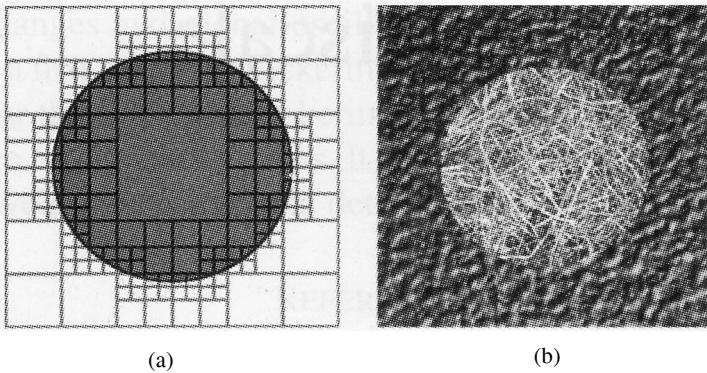


Fig. 1.41 Adaptive Scale Fixing [Liang and Tjahjadi, 2006].

image is presented in Fig. 1.41(a), where two homogeneous textures are depicted as grey and white. The central regions of homogeneous textures are represented using large windows, while the border regions of homogeneous textures are represented using small windows. This representation guarantees that only one texture exists in each window, thus reducing erroneous texture characterization. An image with textures of metal and straw is illustrated in Fig. 1.41(b).

5. Combination of Gabor Wavelet Transforms and Moments for Texture Segmentation

K. Muneeswaran et al. propose a combination approach [Muneeswaran et al., 2005]. It tries to incorporate into itself the better of the two methodologies, namely, Gabor wavelet and general moments. This method involves forming the combinational feature vector for a texture image from the extracted features by both approaches. This combinational approach has been evolved since both measures tend to look for features of a texture image in different views. The following equation is used to construct the feature vector:

$$F_c = \partial(F_M \oplus F_G),$$

where, F_c is the feature vector representing a pixel, constructed by the combinational approach, F_M is the feature vector obtained for the same pixel by applying moments and F_G is the feature vector using Gabor wavelet. ∂

is a stabilizing function which is defined below:

$$\partial(X) = [\Delta x_i, \forall i \in \{1, \dots, m\}],$$

where, X is a vector and m is the vector length and the weight value Δ is calculated by

$$\Delta = \left[\sum_{i=1}^m x_i^2 \right]^{-1/2}.$$

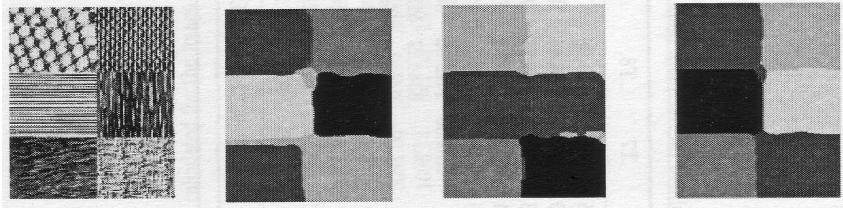


Fig. 1.42 Results of segmentation: (1) left most - input image containing various number of textures; (2) middle-left - segmentation result by Gabor wavelet; (3) middle-right - segmentation result by moments; (4) right most - result by the combination method [Muneeswaran et al., 2005].

1.2.11 *Image Indexing and Retrieval*

Digital image libraries and other multimedia databases have been dramatically expanded in recent years. Storage and retrieval of images in such libraries become a real demand in industrial, medical, and other applications. A solution for this problem is content-based image indexing and retrieval (CBIR), in which some features are extracted from every picture, and stored as an index vector. Thereafter, the index is compared in retrieval phase to find some similar pictures to the query image [Special-Issue-Digital-Library, 1996; Smeulders et al., 2000]. Two major approaches can be used in CBIR, namely, spatial domain-method and transform domain-method.

1. Wavelet Correlogram

An approach called “wavelet correlogram” is proposed by H. A. Moghaddam et al. [Moghaddam et al., 2005], which can take advantage of both spatial and transform domain information. There are three steps in this way.

(1) Wavelet coefficients are computed to decompose space-frequency information of the image; (2) A quantization step is then utilized before computing directional autocorrelograms of the wavelet coefficients; (3) Finally, index vectors are constructed using these wavelet correlograms. Fig. 1.43 shows the diagram of the wavelet correlogram indexing method.

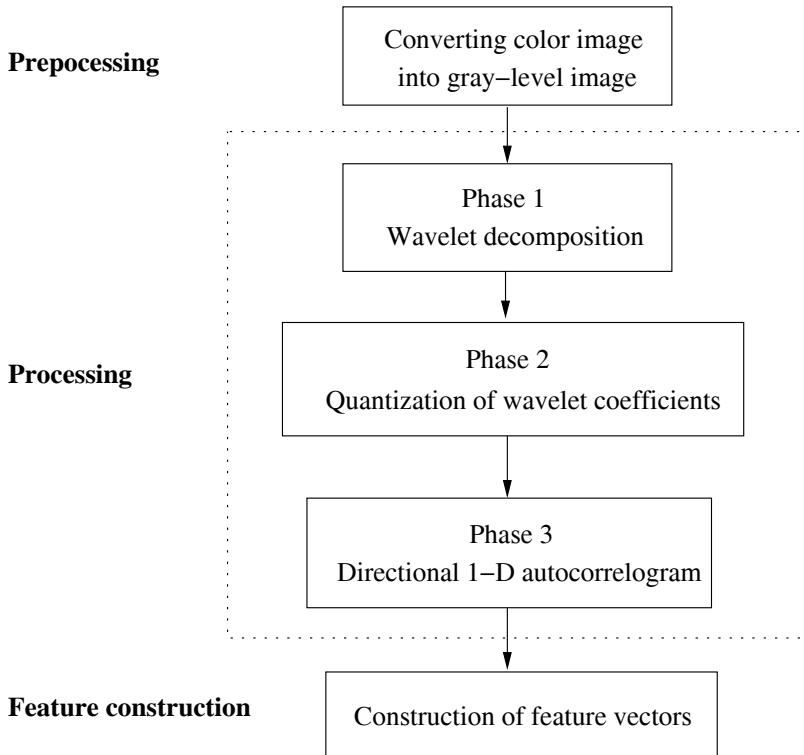


Fig. 1.43 Diagram of the wavelet correlogram indexing method [Moghaddam et al., 2005].

The wavelet autocorrelogram computing equations are

$$\Gamma^{m,n}(i, i, k) = \{(p_1, p_2) | p_1, p_2 \in W_{l_i}^{m,n}, |p_1 - p_2|_n = k\},$$

$$\alpha^{m,n}(i, k) = \frac{|\Gamma^{m,n}(i, i, k)|}{2h_{l_i}(W^{m,n})},$$

where, $|\Gamma^{m,n}(i, i, k)|$ represents the size of the wavelet autocorrelogram for each pair of (m, n) ; $hl_i(W^{m,n})$ is the total number of pixels of level l_i .

For evaluation of the proposed approach, some query images were selected randomly from a 1000 image subset of the COREL database

<http://wang.ist.psu.edu/docs/related>.

The experimental results indicate that a total average of 71% matched retrieved images is achievable using the wavelet correlogram indexing retrieval algorithm [Moghaddam et al., 2005].

2. Region Separation and Multiresolution Analysis

A simple and fast querying method for content-based image retrieval is developed by R. Ksantini et al. [Ksantini et al., 2006]. Using the multispectral gradient, a color image is split into two disjoint parts that are the homogeneous color regions and the edge regions.

(1) The homogeneous regions are represented by the traditional color histograms.

$$h_k^h(c) = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \delta(I_k(i, j) - c) \chi_{[0, \eta]}(\lambda_{max}(i, j)),$$

where, λ_{max} stands for the largest eigenvalue, $\chi_{[0, \eta]}$ is the characteristic function, for each $c \in \{0, \dots, 255\}$ and $k = a, b$.

(2) The edge regions are represented by the multispectral gradient module mean histograms.

$$h_k^e h(c) = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \delta(I_k(i, j) - c) \chi_{[\eta, +\infty]}(\lambda_{max}(i, j)) \lambda_{max}(i, j).$$

In order to measure the similarity degree between two color images both quickly and effectively, R. Ksantini et al. use a one-dimensional pseudo-metric, which makes use of the one-dimensional wavelet decomposition and compression of the extracted histograms. This querying method is invariant to the query color image object translations and color intensities.

Apart from the above methods, a content-based image retrieval technique is presented by M. Kubo et al. [Kubo et al., 2003]. which uses wavelet-based shift and brightness invariant edge features. P. Jain and S. N. Merchant [Jain and Merchant, 2004] propose a method using wavelet-based multiresolution histogram for fast image retrieval.

1.2.12 Wavelet-Based Image Fusion

Image fusion is one of the most important techniques to enhance image information. The fused image is more suitable for image processing. Image fusion is a process by which, information from different observation images are incorporated into a single image. The importance of image fusion lies in the fact that each observation image contains complementary information. When this complementary information is integrated with that of another observation, an image with the maximum amount of information is obtained.

Different approaches have been adopted for multi-sensor or multiple observation image fusion from the simple image averaging approach to the wavelet transform image fusion approach. In [El-Khamy et al., 2006], a super-resolution Linear Minimum Mean Square Error (LMMSE) algorithm for image fusion is proposed by S. E. El-Khamy et al. The schematic diagram of the proposed super-resolution LMMSE algorithm can be found in Fig. 1.44.

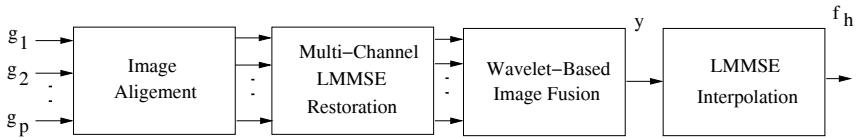


Fig. 1.44 The schematic diagram of the proposed super-resolution algorithm [El-Khamy et al., 2006].

$$f_h = R_{f_n} \begin{bmatrix} D_1 H_1 M_1 \\ \vdots \\ D_p H_p M_p \end{bmatrix}^t \left[\begin{bmatrix} D_1 H_1 M_1 \\ \vdots \\ D_p H_p M_p \end{bmatrix} R_{f_n} \begin{bmatrix} D_1 H_1 M_1 \\ \vdots \\ D_p H_p M_p \end{bmatrix}^t + R_n \right]^{-1}$$

$$\begin{bmatrix} g_1 \\ \cdot \\ \cdot \\ \cdot \\ g_p \end{bmatrix},$$

where, g_i , $i = 1, 2, \dots, p$ stand for the observed images, R_{f_n} and R_n are the high resolution images and noise correlation matrices respectively, D_i , $i = 1, 2, \dots, p$ is the uniform downsample matrix, H_i , $i = 1, 2, \dots, p$ is the blur matrix, and M_i , $i = 1, 2, \dots, p$ is the registration shift matrix.

El-Khamy [El-Khamy et al., 2006] adopts the wavelet transform image fusion approach to integrate the data from the multiple outputs of the LMMSE image restoration step. This is due to the fact that the multiple outputs of the LMMSE restoration step are correctly registered and aligned. The registration of the multiple inputs to the wavelet fusion step is a very important pre-requisite to the success of the fusion step. In the application of the simple wavelet image fusion scheme, the wavelet packet decomposition is calculated for each observation to obtain the multiresolution levels of the images to be fused. In the transform domain, the coefficients in all resolution levels whose absolute values are larger are chosen between the available observations. This rule is known as the maximum frequency rule. Using this method, fusion takes place in all resolution levels and dominant features at each scale are preserved. Another alternative to the maximum frequency rule is the area-based selection rule. This rule is called the local variance or the standard deviation rule. The local variance of the wavelet coefficients is calculated as a measure of local activity levels associated with each wavelet coefficient. If the measures of activity of the wavelet coefficients in each of the two images to be fused are close together, the average of the two wavelet coefficients is taken; otherwise, the coefficient with maximum absolute is chosen. Generally, the process of wavelet packet image fusion can be summarized in the following steps: (1) The available images are first registered (This is achieved for the outputs of the LMMSE restoration step). (2) The wavelet packet decomposition of the observations is calculated using a suitable basis function and decomposition level. (3) A suitable fusion rule is used to select the wavelet coefficients from the source observations. (4) A decision map is created according to the fusion rule. (5) A wavelet packet reconstruction is performed on the combinations created coefficients.

A novel pixel level image fusion schemes are presented by H. Li [Li, 2006] based on multi-scale decomposition. In this way, the wavelet coefficients of image are chosen according to the image fusion operators and different fusion rules. Experiments of multi-spectral image and high-resolution panchromatic images are given. It shows that the wavelet-based human visual system method can achieve better fusion performance than others.

An approach to multisensor remote sensing image fusion using stationary wavelet transform is proposed in [Li, 2008]. In this paper, the authors investigate the effects of orthogonal/biorthogonal filters and decomposition depth on using stationary wavelet analysis for fusion. Spectral discrepancy and spatial distortion are used as quality measures. Empirical results lead to some recommendations on the wavelet filter parameters for use in remote sensing image fusion applications.

1.2.13 Others

Apart from the above sections, there are many other applications of wavelet theory to pattern recognition. For instance, [Chambolle et al., 1998] examines the relationship between wavelet-based image processing algorithms and variational problems. Algorithms are derived as exact or approximate minimizers of variational problems; in particular, it shows that wavelet shrinkage can be considered the exact minimizer of the following problem: Given an image F defined on a square I . minimize over all g in the Besov space $B_1^1(L_1(I))$ the functional $\|F - g\|_{L_2(I)}^2 + \lambda \|g\|_{B_1^1(L_1(I))}$. It uses the theory of nonlinear wavelet image compression in $L_2(I)$ to derive accurate error bounds for noise removal through wavelet shrinkage applied to images corrupted with i.i.d., mean zero, Gaussian noise. A new signal-to-noise ratio (SNR), which claim more accurately reflects the visual perception of noise in images, arises in this derivation. It presents extensive computations that support the hypothesis that near-optimal shrinkage parameters can be derived if one knows (or can estimate) only two parameters about an image F : the largest α for which $F \in B_q^\alpha(L_q(I))$, $1/q = \alpha/2 + 1/2$, and the norm $\|F\|_{B_q^\alpha(L_q(I))}$. Both theoretical and experimental results indicate that this choice of shrinkage parameters yields uniformly better results than Donoho and Johnstone's VisuShrink procedure.

A standard wavelet multiresolution analysis can be defined via a sequence of projectors onto a monotone sequence of closed vector subspaces possessing certain properties. [Combettes, 1998] proposes a nonlinear ex-

tension of this framework in which the vector subspaces are replaced by convex subsets. These sets are chosen so as to provide a recursive, monotone approximation scheme that allows for various signal and image features to be investigated. Several classes of convex multiresolution analysis are discussed and numerical applications to signal and image-processing problems are demonstrated.

Image restoration problems can naturally be cast as constrained convex programming problems in which the constraints arise from a priori information and the observation of signals physically related to the image to be recovered. In [Combettes and Pesquet, 2004], the focus is placed on the construction of constraints based on wavelet representations. Using a mix of statistical and convex-analytical tools, P. L. Combettes and J. -C. Pesquet propose a general framework to construct wavelet-based constraints. The resulting optimization problem is then solved with a block-iterative parallel algorithm which offers great flexibility in terms of implementation. Numerical results illustrate an application of the proposed framework.

A novel approach for thinning character using modulus minima of wavelet transform is developed in [You et al., 2006]. A method for signal denoising using wavelets and block hidden Markov model is presented by Z. Liao [Liao and Tang, 2005].

This page intentionally left blank

Chapter 2

Continuous Wavelet Transforms

Similar to Fourier transform, the wavelet theory consists of two parts, namely, the wavelet transform and the wavelet basis. In this chapter, we will focus on the former and organize it into three sections: (1) the general theory of wavelet transform, (2) the filtering properties of the wavelet transforms, and (3) the characterization of Lipschitz regularity of signals by wavelet transforms.

2.1 General Theory of Continuous Wavelet Transforms

In Fourier transform

$$\mathcal{F}f(\xi) = \int_{\mathbb{R}} f(\xi - x) \overline{e^{i\xi x}} dx,$$

if we replace $e^{i\xi x}$, the dilation of the basic wave function e^{ix} , with $\psi(\frac{t-b}{a})$, the translation and dilation of the basic wavelet $\psi(x)$, then, the transform is referred as a continuous wavelet transform (also called an integral wavelet transform), which is defined as follows:

Definition 2.1 A function $\psi \in L^2(\mathbb{R})$ is called an admissible wavelet or a basic wavelet if it satisfies the following “admissibility ” condition:

$$C_\psi := \int_{\mathbb{R}} \frac{|\hat{\psi}(\xi)|^2}{|\xi|} d\xi < \infty. \quad (2.1)$$

The continuous (or integrable) wavelet transform with kernel ψ is defined

by

$$(W_\psi f)(a, b) := |a|^{-\frac{1}{2}} \int_{\mathbb{R}} f(t) \overline{\psi(\frac{t-b}{a})} dt, \quad f \in L^2(\mathbb{R}), \quad (2.2)$$

where $a, b \in \mathbb{R}$ and $a \neq 0$ are the dilation parameter and the translation parameter respectively.

Before showing the functions of the admissibility condition (2.1), we would like to discuss some basic properties of the continuous wavelet transform at first. According to the admissibility condition (2.1), if $\psi \in L^1(\mathbb{R})$, it can be mathematically inferred that

$$\hat{\psi}(0) = 0$$

and

$$\hat{\psi}(\xi) \rightarrow 0 \quad (|\xi| \rightarrow \infty).$$

This indicates that the function ψ is a bandpass filter.

The characteristics of the localized components of a signal $f(t)$ can be described by the continuous wavelet transform.

- Because of the damp of $\psi(x)$ at infinity, the localized characteristic of f near $x = b$ is described by (2.2). It will be clearer if we extremely assume that the $\psi(x)$ always be zero out of $[-1, 1]$. Then, for all $x \notin [b - |a|, b + |a|]$, we have

$$\psi\left(\frac{t-b}{a}\right) = 0.$$

Therefore,

$$\begin{aligned} (W_\psi f)(a, b) &= |a|^{-\frac{1}{2}} \int_{\mathbb{R}} f(t) \overline{\psi\left(\frac{t-b}{a}\right)} dt \\ &= |a|^{-\frac{1}{2}} \int_{b-|a|}^{b+|a|} f(t) \overline{\psi\left(\frac{t-b}{a}\right)} dt. \end{aligned}$$

This means that $(W_\psi f)(a, b)$ is completely determined by the behaviors of f in $[b - |a|, b + |a|]$ with the center b (Fig. 2.1). It is said that $(W_\psi f)(a, b)$ describes only the localized characteristic of f in $[b - |a|, b + |a|]$. The smaller a , the better the localized characteristic of f , which can also be found in Fig. 2.1.

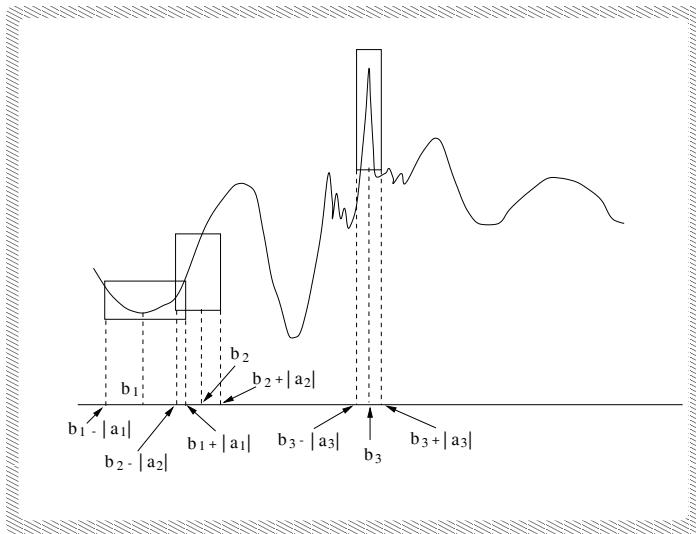


Fig. 2.1 Time-frequency localization.

- On the other hand, by the basic properties of Fourier transform, we have:

$$\begin{aligned}
 (W_\psi f)(a, b) &= \frac{|a|^{-\frac{1}{2}}}{2\pi} \int_{\mathbb{R}} \hat{f}(\xi) (\psi(\frac{\psi - b}{a})) \hat{\psi}(\xi) d\xi \\
 &= \frac{|a|^{-\frac{1}{2}}}{2\pi} \int_{\mathbb{R}} \hat{f}(\xi) |a| e^{-ib\xi} \hat{\psi}(a\xi) d\xi \\
 &= \frac{|a|^{\frac{1}{2}}}{2\pi} \int_{\mathbb{R}} \hat{f}(\xi) e^{-ib\xi} \hat{\psi}(a\xi) d\xi.
 \end{aligned}$$

There is only a phase difference between $e^{-ib\xi} \hat{\psi}(a\xi)$ and $\hat{\psi}(a\xi)$. The bandpass property of ψ ensures that the energy of $\hat{\psi}(a\xi)$ concentrates on two bands, while the bandwidth depends on the scale parameter a with positive ratio. Hence, the localization of \hat{f} is illustrated by $(W_\psi f)(a, b)$, and \hat{f} gets worse localization with a smaller a .

ψ can be viewed as a basic window function. Through this window, we can not observe the integrated f or \hat{f} , whereas the local performances of f and \hat{f} are very clear. Moving the translation parameter b , each part of f and \hat{f} can be traveled. Meanwhile, by adjusting a , we can observe f and \hat{f}

with different localization. The later is the so-called focus property of the wavelet.

For the convenience of our discussion, under the meaning of statistics, we define the center of basic window ψ as

$$x_\psi^* := \frac{1}{\|\psi\|_2^2} \int_{\mathbb{R}} t |\psi(t)|^2 dt. \quad (2.3)$$

Moreover, we define the radius of the window , which is the statistic width of the energy of ψ to its center, as

$$\Delta_\psi := \frac{1}{\|\psi\|_2} \left[\int_{\mathbb{R}} (t - x^*)^2 |\psi(t)|^2 dt \right]^{1/2}, \quad (2.4)$$

where $\|\psi\|_2$ is the norm of ψ , i.e.,

$$\|\psi\|_2 := \left(\int_{\mathbb{R}} |\psi(x)|^2 dx \right)^{1/2}.$$

The window of ψ (also called the time window of ψ in general) is

$$[x_\psi^* - \Delta_\psi, x_\psi^* + \Delta_\psi].$$

In the same way, the window of $\hat{\psi}$ (also called the frequency window of ψ) is

$$[x_{\hat{\psi}}^* - \Delta_{\hat{\psi}}, x_{\hat{\psi}}^* + \Delta_{\hat{\psi}}].$$

The following square region

$$[x_\psi^* - \Delta_\psi, x_\psi^* + \Delta_\psi] \times [x_{\hat{\psi}}^* - \Delta_{\hat{\psi}}, x_{\hat{\psi}}^* + \Delta_{\hat{\psi}}]$$

is referred to the time-frequency window of ψ . For $a, b \in \mathbb{R}, a \neq 0$, a set of wavelets can be generated by the basic wavelet ψ as follows:

$$\psi_{a,b}(t) := |a|^{-\frac{1}{2}} \psi\left(\frac{t-b}{a}\right). \quad (2.5)$$

Thus, the time-frequency window of $\psi_{a,b}$ is

$$[b + ax_\psi^* - |a|\Delta_\psi, b + ax_\psi^* + |a|\Delta_\psi] \times \left[\frac{x_{\hat{\psi}}^*}{a} - \frac{1}{|a|}\Delta_{\hat{\psi}}, \frac{x_{\hat{\psi}}^*}{a} + \frac{1}{|a|}\Delta_{\hat{\psi}} \right]. \quad (2.6)$$

The procedure of reasoning is below:

$$\begin{aligned}
 x_{\psi_{a,b}}^* &= \frac{1}{\|\psi\|_2^2} \int_{\mathbb{R}} t|a|^{-1} |\psi(\frac{t-b}{a})|^2 dt \\
 &= \frac{|a|^{-1}}{\|\psi\|_2^2} \int_{\mathbb{R}} (ax + b) |\psi(x)|^2 |a| dx \\
 &= \frac{1}{\|\psi\|_2^2} \left[a \int_{\mathbb{R}} x |\psi(x)|^2 dx + b \|\psi\|_2^2 \right] \\
 &= ax_{\psi}^* + b;
 \end{aligned}$$

$$\begin{aligned}
 \Delta_{\psi_{a,b}} &= \frac{1}{\|\psi\|_2} \left[\int_{\mathbb{R}} (t - ax_{\psi}^* - b)^2 |a|^{-1} |\psi(\frac{t-b}{a})|^2 dt \right]^{1/2} \\
 &= \frac{1}{\|\psi\|_2} \left[\int_{\mathbb{R}} (ax - ax_{\psi}^*)^2 |\psi(x)|^2 |a|^{-1} |a| dx \right]^{1/2} \\
 &= \frac{|a|}{\|\psi\|_2} \left[\int_{\mathbb{R}} (x - x_{\psi}^*)^2 |\psi(x)|^2 dx \right]^{1/2} \\
 &= |a| \Delta_{\psi};
 \end{aligned}$$

$$\hat{\psi}_{a,b}(\xi) = \int_{\mathbb{R}} |a|^{-\frac{1}{2}} \psi(\frac{t-b}{a}) e^{-i\xi t} dt = |a|^{\frac{1}{2}} e^{-ib\xi} \hat{\psi}(a\xi);$$

$$\|\hat{\psi}_{a,b}\|_2^2 = |a| \int_{\mathbb{R}} |\hat{\psi}(a\xi)|^2 d\xi = \|\hat{\psi}\|_2^2;$$

$$\begin{aligned}
 x_{\hat{\psi}_{a,b}}^* &= \frac{1}{\|\hat{\psi}_{a,b}\|_2^2} \int_{\mathbb{R}} \xi |a| |\hat{\psi}(a\xi)|^2 d\xi \\
 &= \frac{1}{a} \frac{1}{\|\hat{\psi}_{a,b}\|_2^2} \int_{\mathbb{R}} \xi |\hat{\psi}(\xi)|^2 d\xi \\
 &= \frac{1}{a} x_{\hat{\psi}}^*;
 \end{aligned}$$

$$\begin{aligned}
 \Delta_{\hat{\psi}_{a,b}} &= \frac{1}{\|\hat{\psi}\|_2} \left[\int_{\mathbb{R}} (\xi - \frac{1}{|a|} x_{\hat{\psi}}^*)^2 |a| |\hat{\psi}(a\xi)|^2 d\xi \right]^{1/2} \\
 &= \frac{1}{|a|} \frac{1}{\|\hat{\psi}\|_2} \left[\int_{\mathbb{R}} (|a|\xi - x_{\hat{\psi}}^*)^2 |\hat{\psi}(a\xi)|^2 |a| d\xi \right]^{1/2}
 \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{|a|} \frac{1}{\|\hat{\psi}\|_2} \left[\int_{\mathbb{R}} (\xi - x_{\hat{\psi}}^*)^2 |\hat{\psi}(\xi)|^2 d\xi \right]^{1/2} \\
&= \frac{1}{|a|} \Delta_{\hat{\psi}}.
\end{aligned}$$

From (2.6), we can easily find out, when $|a|$ changes to be smaller, the time window of ψ becomes more narrow, whereas the frequency window becomes wider. The area of its time-frequency window is a constant, which is irrelevant to a and b .

$$(2\Delta_{\psi_{a,b}})(2\Delta_{\hat{\psi}_{a,b}}) = 4\Delta_{\psi}\Delta_{\hat{\psi}}.$$

From the above discussion, we find an intrinsic fact: for a basic wavelet ψ , it is impossible to achieve perfect localization in both the time domain and the frequency domain simultaneously. Is there any ψ could make the area of the time-frequency to be small enough? Unfortunately, the following theorem, the famous Heisenberg uncertainty principle, gave a negative answer to this question.

Theorem 2.1 *Let $\psi \in L^2(\mathbb{R})$ satisfy $x\psi(x)$ and $\xi\hat{\psi}(\xi) \in L^2(\mathbb{R})$. Then*

$$\Delta_{\psi}\Delta_{\hat{\psi}} \geq \frac{1}{2}.$$

Furthermore, the equality in the above equation holds if and only if

$$\psi(x) = ce^{iax}g_{\alpha}(x - b),$$

where $c \neq 0, \alpha > 0, a, b \in \mathbb{R}$ and $g_{\alpha}(x)$ is the Gaussian function defined by

$$g_{\alpha}(x) := \frac{1}{2\sqrt{\pi\alpha}} e^{-\frac{x^2}{4\alpha}}. \quad (2.7)$$

Proof. We assume that the window centers of ψ and $\hat{\psi}$ are 0 without losing generality. Based on the basic knowledge of Fourier analysis, we have

$$\begin{aligned}
(\Delta_{\psi}\Delta_{\hat{\psi}})^2 &= \frac{\left(\int_{\mathbb{R}} t^2 |\psi(t)|^2 dt \right) \left(\int_{\mathbb{R}} \xi^2 |\hat{\psi}(\xi)|^2 d\xi \right)}{\|\psi\|_2^2 \|\hat{\psi}\|_2^2} \\
&= \frac{\left(\int_{\mathbb{R}} t^2 |\psi(t)|^2 dt \right) \left(\int_{\mathbb{R}} |\hat{\psi}'(\xi)|^2 d\xi \right)}{\|\psi\|_2^2 \|\hat{\psi}\|_2^2}
\end{aligned}$$

$$\begin{aligned}
&= \frac{\|x\psi(x)\|_2^2 \|\hat{\psi}'(x)\|_2^2}{\|\psi\|_2^2 \|\hat{\psi}\|_2^2} \\
&= \frac{\|x\psi(x)\|_2^2 2\pi \|\psi'(x)\|_2^2}{\|\psi\|_2^2 2\pi \|\psi\|_2^2} \\
&\geq \frac{\|x\psi(x)\psi'(x)\|_2^2}{\|\psi\|_2^4} \\
&\geq \frac{1}{\|\psi\|_2^4} \left| \operatorname{Re} \int_{\mathbb{R}} x\psi(x) \overline{\psi'(x)} dx \right|^2 \\
&= \frac{1}{\|\psi\|_2^4} \left| \frac{1}{2} \int_{\mathbb{R}} x \frac{d}{dx} |\psi(x)|^2 dx \right|^2 \\
&= \frac{1}{\|\psi\|_2^4} \left(\frac{1}{2} \int_{\mathbb{R}} |\psi(x)|^2 dx \right)^2 \\
&= \frac{1}{4}.
\end{aligned}$$

Thus, we have

$$\Delta_\psi \Delta_{\hat{\psi}} \geq \frac{1}{2}.$$

Further, by the conditions, which preserve the equal-sign in the Holder inequality, we know that the equalities in above reasoning will be tenable, if and only if there is a constant α such that

$$\begin{cases} -\operatorname{Re}(x\psi(x)\overline{\psi'(x)}) = |x\psi(x)\overline{\psi'(x)}|, \\ |x\psi(x)| = 2\alpha|\psi'(x)|, \\ \|\psi\|_2 \neq 0 \end{cases}.$$

First, by the second equality, we have

$$x\psi(x) = 2\alpha\psi'(x)e^{i\theta(x)},$$

where $\theta(x)$ is a real-valued function. Second, by the first equality, we have

$$-x\psi(x)\overline{\psi'(x)} \geq 0.$$

Then,

$$-2\alpha|\psi'(x)|^2 e^{i\theta(x)} \geq 0.$$

It infers $e^{i\theta(x)} = -1$. Thus, we have

$$x\psi(x) = -2\alpha\psi'(x).$$

By resolving this ordinary differential equation, we can obtain

$$\psi(x) = ce^{-\frac{x^2}{4\alpha}}.$$

Finally, by the third equality, we know that $c \neq 0$.

It should be mentioned, at the beginning of our proof, we have assumed that the window centers of ψ and $\hat{\psi}$ are 0. Otherwise, consider

$$\tilde{\psi}(x) := e^{-iax}\psi(x+b),$$

where $a := x_{\hat{\psi}}^*, b := x_{\psi}^*$. The centers of the time window and the frequency window of $\tilde{\psi}(x)$ are 0. It means that the above reasoning is available to $\tilde{\psi}(x)$. It is easy to know that

$$\Delta_{\psi} = \Delta_{\tilde{\psi}}, \quad \Delta_{\hat{\psi}} = \Delta_{\hat{\tilde{\psi}}},$$

Hence, $\Delta_{\psi}\Delta_{\hat{\psi}} \geq \frac{1}{2}$. The equality holds if and only if

$$\tilde{\psi}(x) = ce^{-\frac{x^2}{4\alpha}},$$

i.e.

$$\psi(x) = ce^{i(x-b)a}e^{-\frac{(x-b)^2}{4\alpha}},$$

where $c \neq 0$. Replacing $\frac{ce^{iab}}{2\sqrt{2\pi\alpha}}$ with c , we have

$$\psi(x) = ce^{i\alpha x}g_{\alpha}(x-b),$$

where $c \neq 0, \alpha > 0, a, b \in \mathbb{R}$ and $g_{\alpha}(x)$ is the Gaussian function defined by (2.7). This establishes the theorem. ■

An important fact is supported by this theorem: no matter what wavelet ψ we choose, it can not achieve perfect localization in both the time domain and the frequency domain simultaneously. The time window and the frequency window just like the length and width of a rectangle with a fixed area. When the time window is narrow, the frequency window must be wide. By the contrary, when the time window is wide, the frequency window must be narrow (see Fig. 2.2). In fact, the narrow time window and the wide frequency window are good for the high frequency part of a signal, whereas the wide time window and the narrow frequency window are good

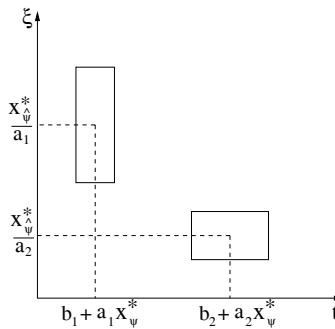


Fig. 2.2 Time-frequency windows, $a_1 < a_2$.

for the low frequency part of a signal. Therefore, this property of wavelet just meets the demand of signal analysis. It can be applied for self-adaptive signal analysis.

The area of the time-frequency window of Gaussian is the smallest, which is $4\Delta_\psi\Delta_{\hat{\psi}} = 2$. Therefore, for the localized time-frequency analysis, Gaussian function is the best. It has been a traditional tool for signal analysis.

As the same as Fourier transform, wavelet transform is invertible. The inverse wavelet transform can be viewed as the reconstruction of the original signal. First of all, we need to prove the following theorem, which specifies that wavelet transform keeps the law of energy-conservation.

Theorem 2.2 *Let ψ be a basic wavelet. Then*

$$\int_{\mathbb{R}} \int_{\mathbb{R}} W_\psi f(a, b) \overline{W_\psi g(a, b)} \frac{dadb}{a^2} = C_\psi \langle f, g \rangle$$

holds for any $f, g \in L^2(\mathbb{R})$, where C_ψ is a constant defined by (2.1), $\langle f, g \rangle$ is the inner product of f and g . Particularly, for $g = f$ we have

$$\int_{\mathbb{R}} \int_{\mathbb{R}} |W_\psi f(a, b)|^2 \frac{dadb}{a^2} = C_\psi \|f\|_2^2.$$

Proof. Obviously, the second equality is a specific instance of the first equality at $g = f$. Thus, only the first equality needs to be proved. It is easy to know that

$$(\psi(\frac{t-b}{a}))^\wedge(\xi) = |a| e^{-ib\xi} \hat{\psi}(a\xi).$$

Denote that

$$\begin{cases} F(x) := \hat{f}(x)\overline{\hat{\psi}(ax)} \\ G(x) := \hat{g}(x)\overline{\hat{\psi}(ax)} \end{cases}$$

Then,

$$\begin{aligned} (W_\psi f)(a, b) &= \langle f(x), |a|^{-\frac{1}{2}}\psi\left(\frac{x-b}{a}\right) \rangle \\ &= \frac{1}{2\pi} \langle \hat{f}(\xi), |a|^{\frac{1}{2}}e^{-b\xi}\hat{\psi}(a\xi) \rangle \\ &= \frac{1}{2\pi} |a|^{\frac{1}{2}} \int_{\mathbb{R}} \hat{f}(\xi) \overline{\hat{\psi}(a\xi)} e^{b\xi} d\xi \\ &= \frac{1}{2\pi} |a|^{\frac{1}{2}} \int_{\mathbb{R}} F(\xi) e^{b\xi} d\xi \\ &= \frac{1}{2\pi} |a|^{\frac{1}{2}} \hat{F}(-b). \end{aligned}$$

In the same way,

$$(W_\psi g)(a, b) = \frac{1}{2\pi} |a|^{\frac{1}{2}} \hat{G}(-b).$$

Therefore,

$$\begin{aligned} &\int_{\mathbb{R}} (W_\psi f)(a, b) \overline{(W_\psi g)(a, b)} db \\ &= (\frac{1}{2\pi})^2 |a| \int_{\mathbb{R}} \hat{F}(-b) \overline{\hat{G}(-b)} db \\ &= (\frac{1}{2\pi})^2 |a| \int_{\mathbb{R}} \hat{F}(b) \overline{\hat{G}(b)} db \\ &= \frac{1}{2\pi} |a| \int_{\mathbb{R}} F(x) \overline{G(x)} dx. \end{aligned}$$

Furthermore,

$$\begin{aligned} &\int_{\mathbb{R}} \int_{\mathbb{R}} (W_\psi f(a, b) \overline{W_\psi g(a, b)}) \frac{da db}{a^2} \\ &= \int_{\mathbb{R}} \frac{1}{2\pi} |a| \int_{\mathbb{R}} F(x) \overline{G(x)} dx \frac{da}{a^2} \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{|\psi(ax)|^2}{|a|} \hat{f}(x) \overline{\hat{g}(x)} dx da \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2\pi} \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \frac{|\psi(ax)|^2}{|ax|} d(ax) \right) \hat{f}(x) \overline{\hat{g}(x)} dx \\
&= \frac{1}{2\pi} \left(\int_{\mathbb{R}} \frac{|\psi(\xi)|^2}{|\xi|} d\xi \right) \int_{\mathbb{R}} \hat{f}(x) \overline{\hat{g}(x)} dx \\
&= \frac{1}{2\pi} C_{\psi} \langle \hat{f}, \hat{g} \rangle \\
&= C_{\psi} \langle f, g \rangle.
\end{aligned}$$

This completes our proof. ■

According to the above theorem, the inverse wavelet transform, which is also called the reconstruction formula, can be formally inferred. We consider

$$\overline{W_{\psi}g(a,b)} = \langle |a|^{-\frac{1}{2}} \psi\left(\frac{x-b}{a}\right), g \rangle.$$

If we give up the exactness in mathematics temporarily, we have

$$\begin{aligned}
C_{\psi} \langle f, g \rangle &= \int_{\mathbb{R}} \int_{\mathbb{R}} W_{\psi}f(a,b) \langle |a|^{-\frac{1}{2}} \psi\left(\frac{x-b}{a}\right), g \rangle \frac{dadb}{a^2} \\
&= \langle \int_{\mathbb{R}} \int_{\mathbb{R}} W_{\psi}f(a,b) |a|^{-\frac{1}{2}} \psi\left(\frac{x-b}{a}\right) \frac{dadb}{a^2}, g \rangle.
\end{aligned}$$

Because that g can be any function in $L^2(\mathbb{R})$, we can write

$$f(x) = C_{\psi}^{-1} \int_{\mathbb{R}} \int_{\mathbb{R}} W_{\psi}f(a,b) |a|^{-\frac{1}{2}} \psi\left(\frac{x-b}{a}\right) \frac{dadb}{a^2}. \quad (2.8)$$

This is the formal inverse wavelet transform, by which we can reconstruct the original signal from the wavelet transform $W_{\psi}f(a,b)$. The above reasoning is not exact in mathematics. The exact inverse formula is studied as follows.

Theorem 2.3 *Let ψ be a basic wavelet. Then, for any $f \in L^2(\mathbb{R})$, the inverse wavelet transform (2.8) holds in the sense of $L^2(\mathbb{R})$ -norm, namely:*

$$\left\| f(x) - C_{\psi}^{-1} \int_{|a| \geq A} da \int_{|b| \leq B} (W_{\psi}f)(a,b) |a|^{-\frac{1}{2}} \psi\left(\frac{x-b}{a}\right) \frac{dadb}{a^2} \right\|_2 \rightarrow 0,$$

as $A \rightarrow 0$, $B \rightarrow \infty$, where C_{ψ} is a constant defined by (2.1).

Proof. For any $A, B > 0$, we prove the following equality at first:

$$\begin{aligned} & \left\langle \int \int_{|a| \geq A, |b| \leq B} (W_\psi f)(a, b) |a|^{-\frac{1}{2}} \psi\left(\frac{x-b}{a}\right) \frac{dadb}{a^2}, g(x) \right\rangle \\ &= \int \int_{|a| \geq A, |b| \leq B} (W_\psi f)(a, b) \overline{(W_\psi g)(a, b)} \frac{dadb}{a^2}. \end{aligned}$$

In fact,

$$\begin{aligned} & \int_{\mathbb{R}} \int \int_{|a| \geq A, |b| \leq B} \left| (W_\psi f)(a, b) |a|^{-\frac{1}{2}} \psi\left(\frac{x-b}{a}\right) g(x) \right| \frac{dadb}{a^2} dx \\ & \leq \int \int_{|a| \geq A, |b| \leq B} |(W_\psi f)(a, b)| |a|^{-\frac{1}{2}} \left(\int_{\mathbb{R}} |\psi\left(\frac{x-b}{a}\right)|^2 dx \right)^{\frac{1}{2}} \|g\|_2 \frac{dadb}{a^2} dx \\ & = \|\psi\|_2 \|g\|_2 \int \int_{|a| \geq A, |b| \leq B} |(W_\psi f)(a, b)| \frac{dadb}{a^2} \\ & \leq \|\psi\|_2 \|g\|_2 \left(\int \int_{|a| \geq A, |b| \leq B} |(W_\psi f)(a, b)|^2 \frac{dadb}{a^2} \right)^{1/2} \\ & \quad \left(\int \int_{|a| \geq A, |b| \leq B} \frac{dadb}{a^2} \right)^{1/2} \\ & = \|\psi\|_2 \|g\|_2 C_\psi^{1/2} \|f\|_2 \left(4B \int_A^\infty \frac{da}{a^2} \right)^{1/2} \\ & = \|\psi\|_2 \|g\|_2 \|f\|_2 \left(\frac{4B}{A} C_\psi \right)^{1/2} < \infty, \end{aligned}$$

where C_ψ is the constant defined by (2.1). By the Fubini theorem in real-analysis [Rudin, 1974], we have

$$\begin{aligned} & \left\langle \int \int_{|a| \geq A, |b| \leq B} (W_\psi f)(a, b) |a|^{-\frac{1}{2}} \psi\left(\frac{x-b}{a}\right) \frac{dadb}{a^2}, g(x) \right\rangle \\ &= \int_{\mathbb{R}} \int \int_{|a| \geq A, |b| \leq B} (W_\psi f)(a, b) |a|^{-\frac{1}{2}} \psi\left(\frac{x-b}{a}\right) \overline{g(x)} \frac{dadb}{a^2} dx \\ &= \int \int_{|a| \geq A, |b| \leq B} (W_\psi f)(a, b) |a|^{-\frac{1}{2}} \int_{\mathbb{R}} \psi\left(\frac{x-b}{a}\right) \overline{g(x)} dx \frac{dadb}{a^2} \\ &= \int \int_{|a| \geq A, |b| \leq B} (W_\psi f)(a, b) \overline{(W_\psi g)(a, b)} \frac{dadb}{a^2}. \end{aligned}$$

This is the equality that we intend to prove. Therefore,

$$\begin{aligned}
& \left\| f(x) - C_\psi^{-1} \int_{|a| \geq A} da \int_{|b| \leq B} (W_\psi f)(a, b) |a|^{-\frac{1}{2}} \psi\left(\frac{x-b}{a}\right) \frac{dadb}{a^2} \right\|_2 \\
&= C_\psi^{-1} \sup_{\|g\|_2=1} |C_\psi \langle f, g \rangle - \\
&\quad \left\langle \int \int_{|a| \geq A, |b| \leq B} (W_\psi f)(a, b) |a|^{-\frac{1}{2}} \psi\left(\frac{x-b}{a}\right) \frac{dadb}{a^2}, g(x) \right\rangle| \\
&= C_\psi^{-1} \sup_{\|g\|_2=1} \left| C_\psi \langle f, g \rangle - \int \int_{|a| \geq A, |b| \leq B} (W_\psi f)(a, b) \overline{(W_\psi g)(a, b)} \frac{dadb}{a^2} \right| \\
&= C_\psi^{-1} \sup_{\|g\|_2=1} \left| \int \int_{|a| < A \text{ or } |b| > B} (W_\psi f)(a, b) \overline{(W_\psi g)(a, b)} \frac{dadb}{a^2} \right| \\
&\leq C_\psi^{-1} \sup_{\|g\|_2=1} \left(\int \int_{\mathbb{R}^2} |(W_\psi g)(a, b)|^2 \frac{dadb}{a^2} \right)^{1/2} \\
&\quad \left(\int \int_{|a| < A \text{ or } |b| > B} |(W_\psi f)(a, b)|^2 \frac{dadb}{a^2} \right)^{1/2} \\
&= C_\psi^{-1} \sup_{\|g\|_2=1} C_\psi^{1/2} \|g\|_2 \left(\int \int_{|a| < A \text{ or } |b| > B} |(W_\psi f)(a, b)|^2 \frac{dadb}{a^2} \right)^{1/2} \\
&= C_\psi^{-1/2} \left(\int \int_{|a| < A \text{ or } |b| > B} |(W_\psi f)(a, b)|^2 \frac{dadb}{a^2} \right)^{1/2} \\
&\rightarrow 0, \quad (A \rightarrow 0, B \rightarrow \infty).
\end{aligned}$$

The proof of the theorem is complete. ■

Note: The result of Theorem 2.3 is not so clear as formula (2.8). It is important to determine whether formula (2.8) is tenable.

By the result of Theorem 2.3, the following conclusion can be mathematically inferred: as $A \rightarrow 0, B \rightarrow \infty$,

$$f_{A,B}(x) := C_\psi^{-1} \int_{|a| \geq A} da \int_{|b| \leq B} (W_\psi f)(a, b) |a|^{-\frac{1}{2}} \psi\left(\frac{x-b}{a}\right) \frac{dadb}{a^2}$$

converges to $f(x)$ in measure ([Rudin, 1974]). According to Riesz theorem ([Rudin, 1974]), there are two sequences $A_k \rightarrow 0+$ and $B_k \rightarrow +\infty$, such

that $f_{A_k, B_k}(x) \rightarrow f(x)$, ($k \rightarrow \infty$), a.e. $x \in \mathbb{R}$. That is

$$f(x) = \lim_{k \rightarrow \infty} C_\psi^{-1} \int_{|a| \geq A_k} da \int_{|b| \leq B_k} (W_\psi f)(a, b) |a|^{-\frac{1}{2}} \psi\left(\frac{x-b}{a}\right) \frac{dadb}{a^2}.$$

If the right integral of (2.8) exists, (2.8) holds a.e. $x \in \mathbb{R}$. Based on this conclusion, we have

Theorem 2.4 *Let ψ be a basic wavelet. Then for any $f \in L^2(\mathbb{R})$, (2.8) holds a.e. $x \in \mathbb{R}$ if*

$$\int_{\mathbb{R}} \int_{\mathbb{R}} \left| (W_\psi f)(a, b) |a|^{-\frac{1}{2}} \psi\left(\frac{x-b}{a}\right) \right| \frac{dadb}{a^2} < \infty, \quad (2.9)$$

where C_ψ is a constant, which can be defined by (2.1).

The following theorem gives a sufficient condition such that (2.9) holds.

Theorem 2.5 *Let $\psi \in L^1(\mathbb{R})$ be a basic wavelet. For any $f \in L^2(\mathbb{R})$, if there exists a non-negative measurable function $F(a)$, such that*

$$|W_\psi f(a, b)| \leq F(a), \quad \text{and} \quad \int_{\mathbb{R}} \frac{1}{|a|^{3/2}} F(a) da < \infty,$$

then, (2.9) holds. Consequently, the inverse wavelet transform (2.8) holds a.e. $x \in \mathbb{R}$. Furthermore, if $\psi(x)$ is continuous, the right part of (2.8) is a continuous function on \mathbb{R} .

Proof. It is known that

$$\begin{aligned} & \int_{\mathbb{R}} \int_{\mathbb{R}} \left| (W_\psi f)(a, b) |a|^{-\frac{1}{2}} \psi\left(\frac{x-b}{a}\right) \right| \frac{dadb}{a^2} \\ & \leq \int_{\mathbb{R}} \int_{\mathbb{R}} F(a) \left| \psi\left(\frac{x-b}{a}\right) \right| \frac{1}{|a|^{5/2}} dadb \\ & = \int_{\mathbb{R}} F(a) \frac{1}{|a|^{5/2}} \left(\int_{\mathbb{R}} \left| \psi\left(\frac{x-b}{a}\right) \right| db \right) da \\ & = \|\psi\|_1 \int_{\mathbb{R}} \frac{1}{|a|^{3/2}} F(a) da \\ & < \infty. \end{aligned}$$

Thus, (2.9) holds. If $\psi(x)$ is continuous, we denote

$$w_x(a, b) := (W_\psi f)(a, x-b) |a|^{-\frac{1}{2}} \psi\left(\frac{b}{a}\right) \frac{1}{a^2}.$$

Then

$$|w_x(a, b)| \leq F(a) \left| \psi\left(\frac{b}{a}\right) \right| \frac{1}{|a|^{5/2}} \in L^1(\mathbb{R}^2).$$

$\forall x_n, x \in \mathbb{R}, x_n \rightarrow x (n \rightarrow \infty)$, by the Lebesgue dominated convergence theorem (see [Rudin, 1974]), we have:

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}} \int_{\mathbb{R}} w_{x_n}(a, b) da db = \int_{\mathbb{R}} \int_{\mathbb{R}} w_x(a, b) da db.$$

Note that

$$\begin{aligned} & \lim_{n \rightarrow \infty} \int_{\mathbb{R}} \int_{\mathbb{R}} (W_\psi f)(a, b) |a|^{-\frac{1}{2}} \psi\left(\frac{x_n - b}{a}\right) \frac{da db}{a^2} \\ &= \lim_{n \rightarrow \infty} \int_{\mathbb{R}} \int_{\mathbb{R}} w_{x_n}(a, b) da db \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} w_x(a, b) da db, \end{aligned}$$

therefore, the right part of (2.8), $\int_{\mathbb{R}} \int_{\mathbb{R}} w_x(a, b) da db$, is a continuous function on \mathbb{R} . Our proof is complete. ■

Note: Let ψ be a basic wavelet, $\alpha > 0$ and $f \in L^2(\mathbb{R})$. If there exists constants $C, \delta > 0$, such that

$$|W_\psi f(a, b)| \leq C |a|^{\alpha + \frac{1}{2}}, \quad (\forall b \in \mathbb{R}),$$

for $|a| < \delta$, the conditions of the previous theorem are satisfied. In fact, it will be clear if we let

$$F(a) := \begin{cases} C |a|^{\alpha + \frac{1}{2}}, & |a| < \delta; \\ \|f\|_2 \|\psi\|_2, & |a| \geq \delta \end{cases}$$

The necessity of the admissibility condition (2.1) is very important. It ensures the existence of the inverse wavelet transform. As a conclusion, it is feasible to apply the wavelet transform to signal analysis, image processing and pattern recognition. It is easy to see that this condition is rather weak and is almost equivalent to $\hat{\psi}(0) = 0$, or

$$\int_{\mathbb{R}} \psi(x) dx = 0. \tag{2.10}$$

Actually, if $\psi(x) \in L^2(\mathbb{R})$ and there exists $\alpha > 0$, such that $(1 + |x|)^\alpha \psi(x) \in L^1(\mathbb{R})$, then, (2.1) must be equivalent to (2.10). This conclusion is proved as below:

By $(1 + |x|)^\alpha \psi(x) \in L^1(\mathbb{R})$, we know that $\psi(x) \in L^1(\mathbb{R})$. It means that $\hat{\psi}(\xi)$ is a continuous function in \mathbb{R} . If (2.1) is tenable, then $\hat{\psi}(0) = 0$. It indicates that (2.10) holds. Contrarily, if (2.10) holds, we assume $\alpha \leq 1$ without losing generality, then

$$\begin{aligned} |\hat{\psi}(\xi)| &= \left| \int_{\mathbb{R}} \psi(x) e^{-i\xi x} dx \right| \\ &= \left| \int_{\mathbb{R}} \psi(x) [e^{-i\xi x} - 1] dx \right| \\ &\leq \int_{\mathbb{R}} \left| \psi(x) 2 \sin \frac{\xi x}{2} \right| dx \\ &\leq |\xi|^\alpha \int_{|\xi x| \leq 1} |x^\alpha \psi(x)| dx + |\xi|^\alpha \int_{|\xi x| > 1} |x^\alpha \psi(x)| dx \\ &= |\xi|^\alpha \int_{\mathbb{R}} |x^\alpha \psi(x)| dx \\ &= C|\xi|^\alpha, \end{aligned}$$

where $C := \int_{\mathbb{R}} |x^\alpha \psi(x)| dx$. Therefore,

$$C_\psi := \int_{\mathbb{R}} \frac{|\hat{\psi}(\xi)|^2}{|\xi|} d\xi \leq C^2 \int_{|\xi| \leq 1} |\xi|^{2\alpha-1} d\xi + \int_{|\xi| > 1} |\hat{\psi}(\xi)|^2 d\xi < \infty.$$

It specifies that (2.1) holds. This finished our proof. ■

In general, the wavelets applied to practice have good damping and satisfy the condition $(1 + |x|)^\alpha \psi(x) \in L^1(\mathbb{R})$ ($\alpha > 0$). As a result, in engineering, the wavelet is often defined as the function in $L^2(\mathbb{R})$ which satisfies the condition (2.10). This definition is not exact in mathematics, however, it is harmless in the application of wavelet. The condition (2.10) is a objective specification of the vibration of ψ . The damping and the vibration are two basic characteristics of basic wavelets.

In signal analysis, only the positive frequency is need to be considered. That is $a > 0$. With this premise, we have another theorem.

Theorem 2.6 *Let ψ be a basic wavelet satisfying*

$$\int_0^\infty \frac{|\hat{\psi}(\xi)|^2}{\xi} d\xi = \int_0^\infty \frac{|\hat{\psi}(-\xi)|^2}{\xi} d\xi = \frac{1}{2} C_\psi < \infty.$$

Then, for any $f, g \in L^2(\mathbb{R})$, we have

$$\int_0^\infty \int_{\mathbb{R}} W_\psi f(a, b) \overline{W_\psi g(a, b)} \frac{dadb}{a^2} = \frac{1}{2} C_\psi \langle f, g \rangle.$$

In particular,

$$\int_0^\infty \int_{\mathbb{R}} |W_\psi f(a, b)|^2 \frac{dadb}{a^2} = \frac{1}{2} C_\psi \|f\|_2^2.$$

The proof of this theorem is similar to those of Theorem 2.2 and Theorem 2.3.

2.2 The Continuous Wavelet Transform as a Filter

Let ψ be a basic wavelet and we denote

$$\tilde{\psi}(x) := \overline{\psi(-x)}. \quad (2.11)$$

We define scale wavelet transform as follows:

$$W_s f(x) := W_s^{\tilde{\psi}} f(x) := (f * \tilde{\psi})(x). \quad (2.12)$$

By the definition of the wavelet transform, for $a > 0$, we have

$$\begin{aligned} (W_\psi f)(a, b) &:= \int_{\mathbb{R}} f(t) a^{-\frac{1}{2}} \overline{\psi\left(\frac{t-b}{a}\right)} dt \\ &= \int_{\mathbb{R}} f(t) a^{-\frac{1}{2}} \tilde{\psi}\left(\frac{b-t}{a}\right) dt \\ &= a^{\frac{1}{2}} (f * \tilde{\psi}_a)(b), \end{aligned}$$

where $\tilde{\psi}_a(x) := \frac{1}{a} \tilde{\psi}\left(\frac{x}{a}\right)$. It specifies that the wavelet transform is actually a convolution, which is also called a filter in engineering.

The design of filters is very important in the filter theory. There are two kinds of digital filters, namely, the finite impulse response (FIR) and the infinite impulse response (IIR). The former is easy to be realized and has good time localization. Therefore, in the above scale wavelet transforms, the basic wavelet ψ , which is the kernel of the transform, should be compactly supported. Although Heisenberg uncertainty principle has specified that the area of the time-frequency domain can not be arbitrarily small, the localization in the time-frequency domain still can be perfect, if ψ and

its Fourier transform $\hat{\psi}$ have compact support simultaneously. Unfortunately, from the following theorem, we can find that this condition cannot be satisfied.

Theorem 2.7 *If $f \in L^2(\mathbb{R})$ is a non-zero function, f and its Fourier transform \hat{f} can not be compactly supported simultaneously.*

Proof. If \hat{f} is compactly supported, $\text{supp } \hat{f} \subset [-B, B]$, then

$$f(z) := \frac{1}{2\pi} \int_{-B}^B \hat{f}(\xi) e^{i\xi z} d\xi$$

is an analytic function on complex plane \mathbb{C} . Because of the zero-isolation of non-zero analytic functions, f can not be compactly supported. This finishes our proof. ■

We have known that the Fourier transform $\hat{\psi}$ of a compactly supported basic wavelet ψ can not be compactly supported. Now, we wish its damping property would be good enough. It is essentially equivalent to that the smoothness of ψ would be good enough. The following theorem ensures this fact.

Theorem 2.8 *Let m be a non-negative integer and $H^m(\mathbb{R})$ be the Sobolev space of order m defined by*

$$H^m(\mathbb{R}) := \{f \mid f, f', \dots, f^{(m)} \in L^2(\mathbb{R})\}.$$

Then, $f \in H^m(\mathbb{R})$ if and only if

$$\int_{\mathbb{R}} (1 + |x|^m) |\hat{f}(x)| dx < \infty. \quad (2.13)$$

Proof. This is a fundamental fact in the theory of Sobolev spaces. The detailed proof can be found in [Gilbarg and Trudinger, 1977]. ■

According to the definition of $H^m(\mathbb{R})$, $f \in H^m(\mathbb{R})$ indicates that f has certainly differentiability and smoothness (or regularity). However, (2.13) indicates that \hat{f} has damping of order m . Therefore, when we choose a basic wavelet ψ as a filter function, in order to ensure that ψ is good for localized analysis in the time-frequency, we must consider both its damping (or compact support) and its regularity.

Another very important property of filters is the linear phase. With the linear phase or the generalized linear phase, a filter can avoid distortion.

The linear phase is mathematically defined as follows.

Definition 2.2 $f \in L^2(\mathbb{R})$ is said to have liner phase if its Fourier transform satisfies

$$\hat{f}(\xi) = \epsilon |\hat{f}(\xi)| e^{-ia\xi},$$

where $\epsilon = 1$ or -1 and a is a real constant. f is said to have generalized linear phase if there exists a real function $F(\xi)$ and two real constants a and b , such that

$$\hat{f}(\xi) = F(\xi) e^{-i(a\xi+b)}.$$

Obviously, if f has linear phase, it also has generalized linear phase. The generalized linear phase will degenerate to the linear phase if and only if $e^{-ib} = \pm 1$ and the real function $F(\xi)$ keeps its sign, i.e., identically positive or identically negative.

According to the above mathematical definition, the linear phase or the generalized linear phase of ψ is an attribute in Fourier transform domain. In the following, we will demonstrate that the generalized linear phase is equivalent to a symmetry of ψ itself.

Theorem 2.9 $f \in L^2(\mathbb{R})$ has generalized phase if and only if f is skew-symmetric at $a \in \mathbb{R}$, i.e., there exists a constant $b \in \mathbb{R}$ such that

$$e^{ib} f(a+x) = \overline{e^{ib} f(a-x)}, \quad x \in \mathbb{R}.$$

In particular, for a real function f , it has generalized linear phase if and only if it is symmetric or antisymmetric at $a \in \mathbb{R}$. More precisely,

$$f(a+x) = f(a-x), \quad x \in \mathbb{R},$$

or

$$f(a+x) = -f(a-x), \quad x \in \mathbb{R}.$$

Proof. It is easy to see that f has generalized phase, if and only if two constants a and $b \in \mathbb{R}$ exist, such that

$$\hat{f}(\xi) e^{i(a\xi+b)} = F(\xi)$$

is a real function. That is

$$\hat{f}(\xi) e^{i(a\xi+b)} = \overline{\hat{f}(\xi) e^{i(a\xi+b)}}, \quad (\xi \in \mathbb{R}),$$

which is equivalent to

$$\int_{\mathbb{R}} \hat{f}(\xi) e^{i(a\xi+b)} e^{i\xi x} d\xi = \overline{\int_{\mathbb{R}} \hat{f}(\xi) e^{i(a\xi+b)} e^{-i\xi x} d\xi}, \quad (x \in \mathbb{R}).$$

i.e.,

$$e^{ib} f(a+x) = \overline{e^{ib} f(a-x)}, \quad (x \in \mathbb{R}).$$

If f is a real function, e^{i2b} must be a real function. It means $e^{i2b} = 1$ or -1 . Thus

$$f(a+x) = f(a-x), \quad x \in \mathbb{R},$$

or

$$f(a+x) = -f(a-x), \quad x \in \mathbb{R}.$$

It specifies that f is symmetric or antisymmetric on a . Our proof is complete. ■

Summarily, as a filter, in order to be good for the localized analysis in the time domain, the basic wavelet ψ should have good damping or have compact support. On the other hand, ψ should have good regularity or smoothness, to obtain a good property for the localized analysis in the frequency domain. At last, ψ should be skew-symmetric to avoid the distortion.

2.3 Characterization of Lipschitz Regularity of Signal by Wavelet

In mathematics, a signal is actually a function. A stationary signal always corresponds to a smooth function, while a transient one refers to a singularity. In fact, the concepts of the stability and singularity of the signal are ambiguous without using the tool of mathematics. In this section, we will introduce an accurate description of the regularity of the signal mathematically employing the exponent of Lipschitz.

Definition 2.3 Let α satisfy $0 \leq \alpha \leq 1$. A function f is called uniformly Lipschitz α over the interval (a, b) , if there exists a positive constant K , such that

$$|f(x_1) - f(x_2)| \leq K|x_1 - x_2|^\alpha, \quad \forall x_1, x_2 \in (a, b),$$

and we denote $f \in C^\alpha(a, b)$. The constant α is called Lipschitz exponent.

Function f to be uniformly Lipschitz α over (a, b) is also equivalent to the following definition:

Definition 2.4 Let α satisfy $0 \leq \alpha \leq 1$. A function f is called uniformly Lipschitz α over the interval (a, b) if

$$K := \sup_{x_1, x_2 \in (a, b), x_1 \neq x_2} \frac{|f(x_1) - f(x_2)|}{|x_1 - x_2|} < \infty. \quad (2.14)$$

It shows that the variation of f depends on $|x_1 - x_2|^\alpha$ over (a, b) . Function f has neither the “fracture” points over (a, b) , nor sharp transient points. The shape of the f depends on α , and the sharpness of the shape can be decided by the constant K in (2.14).

In the above definition, the Lipschitz α is confined within $\alpha \leq 1$. Otherwise, when $\alpha > 1$, the above definition loses its meaning, because in this case, the function will be too smooth. Since the definition of the Lipschitz regularity with $\alpha \leq 1$ does not deal with the differentiation, we can extend the definition of the Lipschitz exponent to $\alpha > 1$.

Definition 2.5 Let $\alpha > 1$, and n be the largest number which is less than α . A function f is uniformly Lipschitz α over (a, b) , if f is n th differentiable and $f^{(n)} \in C^{\alpha-n}(a, b)$. It is symbolized by $f \in C^\alpha(a, b)$.

The concept of uniformly Lipschitz α can be extended to $\alpha < 0$, and the definition can be modified as:

Definition 2.6 Let α be $-1 \leq \alpha < 0$. If the primitive function of f

$$F(x) := \int_a^x f(t) dt$$

is uniformly Lipschitz $(\alpha + 1)$ over (a, b) , f is said to be uniformly Lipschitz α over (a, b) , and denoted by $f \in C^\alpha(a, b)$.

In the following, we will prove an important property, that is the Lipschitz regularity closely correlates with the decay property of wavelet transform in scale.

Theorem 2.10 Let $\alpha > 0$, and n be the largest integer which is less than α , $\psi \in L^2(\mathbb{R})$ satisfy $(1 + |x|)^\alpha \psi(x) \in L^1(\mathbb{R})$ and have vanishing moments

of order n , i.e.,

$$\int_{\mathbb{R}} t^k \psi(t) dt = 0, \quad (k = 0, 1, \dots, n).$$

Then, there exists a constant $C > 0$, such that $\forall f \in L^2(\mathbb{R}) \cap C^\alpha(\mathbb{R})$, the following holds

$$|(W_\psi f)(a, b)| \leq C|a|^{\alpha+\frac{1}{2}}, \quad (\forall a, b \in \mathbb{R}).$$

Proof. $\forall a, b \in \mathbb{R}$, we have

$$\begin{aligned} |(W_\psi f)(a, b)| &= \left| \int_{\mathbb{R}} f(t) \frac{1}{|a|^{1/2}} \psi\left(\frac{t-x}{a}\right) dt \right| \\ &= \left| \int_{\mathbb{R}} \left[f(t) - f(x) - f'(x)(t-x) - \dots - \frac{f^{(n)}(x)}{n!}(t-x)^n \right] \right. \\ &\quad \cdot \left. \frac{1}{|a|^{1/2}} \psi\left(\frac{t-x}{a}\right) dt \right| \\ &= \left| \int_{\mathbb{R}} \left[\frac{f^{(n)}(\xi)}{n!}(t-x)^n - \frac{f^{(n)}(x)}{n!}(t-x)^n \right] \right. \\ &\quad \cdot \left. \frac{1}{|a|^{1/2}} \psi\left(\frac{t-x}{a}\right) dt \right| \quad (\xi \text{ is between } t \text{ and } x) \\ &\leq C \int_{\mathbb{R}} \frac{|t-x|^{\alpha-n}}{n!} |t-x|^n \frac{1}{|a|^{1/2}} \left| \psi\left(\frac{t-x}{a}\right) \right| dt \\ &= \frac{C}{n!} |a|^{\alpha+\frac{1}{2}} \int_{\mathbb{R}} |t|^\alpha |\psi(t)| dt \\ &\leq C|a|^{\alpha+\frac{1}{2}}. \end{aligned}$$

where C denotes a positive constant. The proof is complete. \blacksquare

Theorem 2.11 Let $0 < \alpha < 1$, and $\psi \in L^2(\mathbb{R})$ have vanishing moment of order 0 and satisfy:

$$(1+|x|)^\alpha \psi(x) \in L^1(\mathbb{R}), \quad |\psi'(x)| = O\left(\frac{1}{1+|x|^\sigma}\right), \quad (\exists \sigma > 1).$$

Then,

$$f \in C^\alpha(\mathbb{R}) \iff \exists C > 0 : |(W_\psi f)(a, b)| \leq C|a|^{\alpha+\frac{1}{2}}, \quad (\forall a, b \in \mathbb{R}).$$

Proof. Here we give only the proof of the sufficiency part. According to the note in Theorem 2.5, we have $f(x) \in C(\mathbb{R})$, and the following inverse

wavelet transform holds:

$$f(x) = C_\psi^{-1} \int_{\mathbb{R}} \int_{\mathbb{R}} (W_\psi f)(a, b) \psi_{a,b}(x) \frac{dadb}{a^2}, \quad (\forall x \in \mathbb{R}).$$

Therefore

$$\begin{aligned} & C_\psi |f(x) - f(y)| \\ & \leq \int_{\mathbb{R}} \int_{\mathbb{R}} |(W_\psi f)(a, b)| \frac{1}{|a|^{1/2}} \left| \psi\left(\frac{x-b}{a}\right) - \psi\left(\frac{y-b}{a}\right) \right| \frac{dadb}{a^2} \\ & \leq C \int_{|a| \leq \delta} \int_{\mathbb{R}} |a|^{\alpha+1/2} \frac{1}{|a|^{1/2}} \left(\left| \psi\left(\frac{x-b}{a}\right) \right| + \left| \psi'\left(\frac{y-b}{a}\right) \right| \right) \frac{dadb}{a^2} + \\ & \quad + C \int_{|a| > \delta} \int_{\mathbb{R}} |a|^{\alpha+1/2} \frac{1}{|a|^{1/2}} |\psi'(\xi)| \frac{|x-y|}{|a|} \frac{dadb}{a^2} \\ & \quad \quad \quad \text{(where } \xi \text{ is between } \frac{x-b}{a} \text{ and } \frac{y-b}{a} \text{)} \\ & \leq C \int_{|a| \leq \delta} |a|^{\alpha-1} da \|\psi\|_1 + C \left(\int_{|a| > \delta} \int_{\mathbb{R}} |a|^{\alpha 3} \frac{1}{1+|\xi|^\sigma} dadb \right) |x-y| \end{aligned}$$

Since ξ is between $\frac{x-b}{a}$ and $\frac{y-b}{a}$, we have that (please refer to the note followed by this proof)

$$\left| \frac{x-b}{a} \right| \leq |\xi| + \left| \frac{x-b}{a} - \frac{y-b}{a} \right| = |\xi| + \left| \frac{x-y}{a} \right|.$$

In particular, setting $\delta = |x-y|$, for $|a| > \delta$, we can deduce that

$$\left| \frac{x-b}{a} \right| \leq |\xi| + \frac{\delta}{|a|} \leq |\xi| + 1.$$

Therefore, we have

$$1 + \left| \frac{x-b}{a} \right|^\sigma \leq 1 + (|\xi| + 1)^\sigma \leq (1 + 2^\sigma)(|\xi|^\sigma + 1).$$

Subsequently, we arrive at

$$\frac{1}{1+|\xi|^\sigma} \leq (1 + 2^\sigma) \frac{1}{1 + \left| \frac{x-b}{a} \right|^\sigma},$$

and hence, for $\delta = |x-y|$,

$$C_\psi |f(x) - f(y)|$$

$$\begin{aligned}
&\leq C \int_{|a|\leq \delta} |a|^{\alpha-1} da \|\psi\|_1 + C|x-y| \\
&\quad \left(\int_{|a|>\delta} \int_{\mathbb{R}} |a|^{\alpha-3} (1+2^\sigma) \frac{1}{1+|\frac{x-b}{a}|^\sigma} dadb \right) \\
&= C \frac{1}{\alpha} \delta^\alpha + C|x-y| \int_{|a|>\delta} |a|^{\alpha-2} da \int_{\mathbb{R}} \frac{db}{1+|b|^\sigma} \\
&\leq C \frac{1}{\alpha} \delta^\alpha + C|x-y| \frac{1}{1-\alpha} \delta^{\alpha-1} \\
&\leq C \left(\frac{1}{\alpha} + \frac{1}{1-\alpha} \right) |x-y|^\alpha.
\end{aligned}$$

That means $f \in C^\alpha(\mathbb{R})$. The proof is complete. \blacksquare

Note: In the proof, we have used the following obvious results, namely, if ξ, a and b are three real numbers, ξ is between a and b , then $|a| \leq |\xi| + |a-b|$. In fact, (1) If a, b have the different sign, then $|a| \leq |a-b| \leq |\xi| + |a-b|$; (2) If a, b have same sign, then $|a| \leq |\xi| \leq |b|$ or $|b| \leq |\xi| \leq |a|$. In the first case, $|a| \leq |\xi| \leq |\xi| + |a-b|$ already holds; In the second case, we have $|a| \leq |b| + |a-b| \leq |\xi| + |a-b|$. These are the results we wanted.

The above theorem can be extended to the case of $\alpha > 1$ as follows. We omit the proof here.

Theorem 2.12 *Let n be a no-negative integer and $\alpha \in \mathbb{R}$ satisfy $n < \alpha < n+1$. Suppose $\psi \in L^2(\mathbb{R})$ has vanishing moments of order n and satisfies:*

$$(1+|x|)^\alpha \psi(x), \psi, \psi', \dots, \psi^{(n)} \in L^1(\mathbb{R}),$$

$$|\psi^{(n+1)}(x)| = O\left(\frac{1}{1+|x|^\sigma}\right), \quad (\exists \sigma > 1).$$

Then, $\forall f \in L^1(\mathbb{R})$, the following holds

$$f \in C^\alpha(\mathbb{R}) \iff \exists C > 0 : |W_\psi f(a, b)| \leq C|a|^{\alpha+\frac{1}{2}}, \quad (\forall a, b \in \mathbb{R}).$$

2.4 Some Examples of Basic Wavelets

There are a great number of basic wavelets. Generally speaking, the derivative of a compact support function, which is continuous differentiable, is a basic wavelet. Several examples of basic wavelets and their basic properties are given in this section.

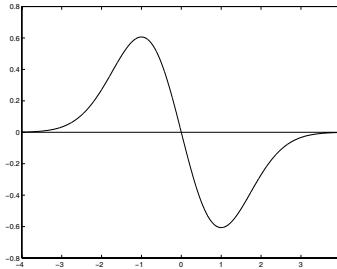


Fig. 2.3 Gaussian wavelet $f(x) = xe^{-\frac{x^2}{2}}$ and its Fourier transform.

- **Gaussian Wavelets:**

$$\psi_\alpha(x) = -\frac{x}{4\alpha\sqrt{\pi\alpha}}e^{-\frac{x^2}{4\alpha}}.$$

It is the derivative of Gaussian function

$$g_\alpha(x) = \frac{1}{2\sqrt{\pi\alpha}}e^{-\frac{x^2}{4\alpha}},$$

i.e., $\psi_\alpha(x) = g'_\alpha(x)$. Its Fourier transform is

$$\hat{\psi}_\alpha(\xi) = i\xi e^{-\alpha\xi^2}.$$

The time-frequency window of ψ_α is

$$[-\sqrt{3\alpha}, \sqrt{3\alpha}] \times \left[-\sqrt{\frac{3}{4\alpha}}, \sqrt{\frac{3}{4\alpha}}\right].$$

Fig. 7.9 graphically shows Gaussian wavelet with $\alpha = \frac{1}{2}$, and its Fourier transform. Both of them have the same shape. The difference between $\psi_\alpha(x)$ and its Fourier transform is only a constant.

- **Mexico hat-like wavelet:** When $\alpha = \frac{1}{2}$, the second derivative of Gaussian function $g_{\frac{1}{2}}$ is a wavelet, which is referred to as Mexico hat-like wavelet:

$$\psi_M(x) := -g''_{\frac{1}{2}}(x) = \frac{1}{\sqrt{2\pi}}(1-x^2)e^{-\frac{1}{2}x^2}.$$

Its Fourier transform is

$$\hat{\tilde{\psi}}_M(\xi) = \xi^2 e^{-\frac{1}{2}\xi^2}.$$

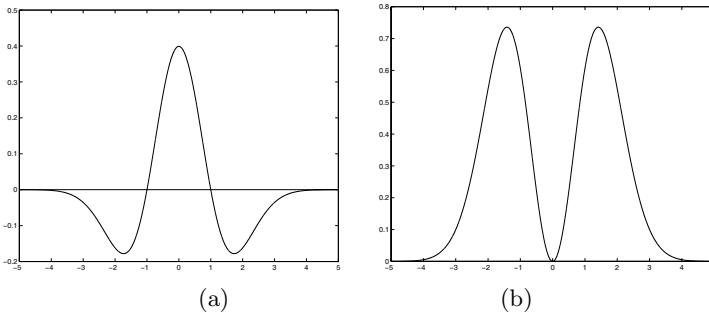


Fig. 2.4 (a): Mexico-hat wavelet ψ_M , and (b): its Fourier transform $\hat{\psi}_M$.

The time-frequency window is

$$\left[-\frac{\sqrt{42}}{6}, \frac{\sqrt{42}}{6} \right] \times \left[-\sqrt{\frac{8}{3\sqrt{\pi}}}, \sqrt{\frac{8}{3\sqrt{\pi}}} \right].$$

Mexico hat-like wavelet and its Fourier transform are shown in Fig. 2.4.

- **Spline wavelet:** For $m \geq 1$, the B-spline of order m is defined by

$$N_m(x) = (\underbrace{N_1 * \dots * N_1}_m)(x) = (N_{m-1} * N_1)(x) = \int_0^1 N_{m-1}(x-t) dt, \quad (2.15)$$

where $N_1(x)$ is a characteristic function of $[0, 1]$, and can be written as

$$N_1(x) := \begin{cases} 1 & x \in [0, 1); \\ 0 & \text{otherwise.} \end{cases}$$

Its Fourier transform is

$$\hat{N}_1(\xi) = \frac{2}{\xi} e^{-i\xi/2} \sin \frac{\xi}{2}.$$

We can easily prove that

$$N_m(x) = \frac{1}{(m-1)!} \sum_{k=0}^m (-1)^k \binom{m}{k} (x-k)_+^{m-1}.$$

$N_m(x)$ has compact support $[0, m]$. For $m \geq 2$, the derivative of $N_m(x)$ is a basic wavelet, which is called the spline wavelet of order

$m - 1$ and can be represented as

$$\psi_{m-1}(x) := N'_m(x) = \frac{1}{(m-1)!} \sum_{k=0}^m (-1)^k \binom{m}{k} (x-k)_+^{m-2} .$$

Its Fourier transform is

$$\begin{aligned} \hat{\psi}_{m-1}(\xi) &= (N'_m)^\wedge(\xi) = i\xi \hat{N}_m(\xi) = i\xi (\hat{N}_1(\xi))^m \\ &= i\xi \left(\frac{2}{\xi} e^{-i\xi/2} \sin \frac{\xi}{2} \right)^m . \end{aligned}$$

In particular, we consider the following two cases, namely, $m = 2$ and $m = 4$:

(1) When $m = 2$, the first spline wavelet is

$$\psi_1(x) := \begin{cases} 1 & x \in (0, 1) \\ -1 & x \in (1, 2) \\ 0 & x \in \mathbb{R} \setminus [0, 1] \end{cases} .$$

By taking dilation to the 1-D spline wavelet, Haar wavelet $h(x) := \psi_1(2x)$ can be achieved:

$$h(x) := \begin{cases} 1 & x \in (0, \frac{1}{2}) \\ -1 & x \in (\frac{1}{2}, 1) \\ 0 & x \in \mathbb{R} \setminus [0, 1] \end{cases} .$$

Its Fourier transform is

$$\hat{h}(\xi) = \frac{1}{2} \hat{\psi}_1\left(\frac{\xi}{2}\right) = \frac{1}{4} i\xi e^{-i\xi/2} \left(\frac{4}{\xi} \sin \frac{\xi}{4}\right)^2 .$$

The time-frequency window is

$$\left[\frac{1}{2} - \frac{1}{2\sqrt{3}}, \frac{1}{2} + \frac{1}{2\sqrt{3}} \right] \times (-\infty, \infty).$$

It is clearly that the width of its frequency window is infinite, Haar wavelet is good for the localized analysis in the time domain but in the frequency domain. Haar wavelet and its Fourier transform are graphically displayed in Fig. 2.5.

(2) When $m = 4$, the quadratic spline wavelet is

$$\psi_3(x) := \frac{1}{6} \sum_{k=0}^4 (-1)^k \binom{4}{k} (x-k)_+^2 .$$

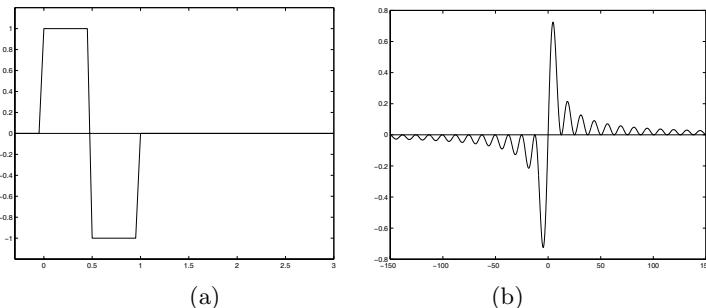


Fig. 2.5 (a): Haar wavelet $h(x)$ and (b): its Fourier transform $-ie^{i\xi/2}\hat{h}(\xi) = \frac{4}{\xi}(\sin \frac{\xi}{4})^2$.

$\tilde{\psi}_3(x) := 4\psi_3(2x + 2)$, the quadratic spline wavelet, is applied widely in signal processing. It is an odd function and has compact support $[-1, 1]$. In $[0, \infty)$, it can be represented as

$$\tilde{\psi}_3(x) := \begin{cases} 8(3x^2 - 2x) & x \in [0, \frac{1}{2}) \\ -8(x - 1)^2 & x \in [\frac{1}{2}, 1) \\ 0 & x \geq 1 \end{cases} .$$

Its Fourier transform can be written as

$$\tilde{\psi}_3(\xi) = i\xi \left(\frac{4}{\xi} \sin \frac{\xi}{4} \right)^4 .$$

The time-frequency window is

$$[-\frac{1}{\sqrt{7}}, \frac{1}{\sqrt{7}}] \times [-4, 4].$$

Quadratic spline wavelet and its Fourier transform are graphically displayed in Fig. 2.6. Note: The width 4 of the frequency window was approximately computed with Matlab system, however, a little error maybe exists.

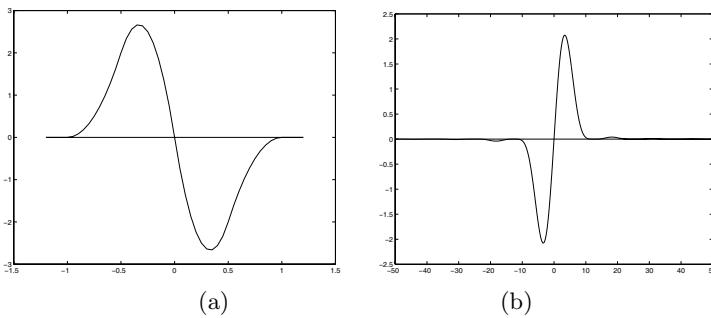


Fig. 2.6 (a): Quadratic Spline wavelet $\tilde{\psi}_3$ and (b): its Fourier transform $-i\hat{\tilde{\psi}}_3$.

This page intentionally left blank

Chapter 3

Multiresolution Analysis and Wavelet Bases

3.1 Multiresolution Analysis

3.1.1 Basic Concept of Multiresolution Analysis (MRA)

It was very difficult to construct a wavelet basis of $L^2(\mathbb{R})$ early on in the history of the wavelet development. From the point of view of function analysis, it is also a non-trivial task to find such a function ψ with good regularity and localization to ensure that $\{\psi_{j,k}(x) := 2^{j/2}\psi(2^j x - k)\}_{j \in \mathbb{Z}, k \in \mathbb{Z}}$ to be an orthonormal basis of $L^2(\mathbb{R})$. For a long time, people doubted the existence of this function. Fortunately, in 1980s, such functions were found, and a standard scheme to construct wavelet bases was set up as well. It has been proved theoretically that almost all useful wavelet bases can be constructed using the standard scheme, namely, Multiresolution Analysis (MRA). MRA was first published in 1989 by Mallat and Meyer [Mallat, 1989a; Meyer, 1990]. Since that time, the multiresolution analysis has become a very important tool in signal processing, image processing, pattern recognition, and other related fields. In this section, we will introduce the intuitive meaning of wavelet and multiresolution analysis. Our purpose is to find a function ψ to ensure that $\{\psi_{j,k}(x) := 2^{j/2}\psi(2^j x - k)\}_{j \in \mathbb{Z}, k \in \mathbb{Z}}$ is an orthonormal bases of $L^2(\mathbb{R})$. Two variables are embedded in them, namely the translation factor $k \in \mathbb{Z}$, and the dilation factor $j \in \mathbb{Z}$. For $\{\psi_{j,k} | k \in \mathbb{Z}\}$, if j is fixed, there will be a fixed bandwidth:

$$|(\psi_{j,k})^\wedge(\xi)| = 2^{-j/2} |\hat{\psi}\left(\frac{\xi}{2^j}\right)|.$$

Thus, the decomposition of $L^2(\mathbb{R})$ in frequency is as follows

$$W_j := \overline{\text{span}\{\psi_{j,k} \mid k \in \mathbb{Z}\}},$$

where the $\{\psi_{j,k}\}_{k \in \mathbb{Z}}$ denotes the orthonormal bases of W_j . We can write

$$L^2(\mathbb{R}) = \cdots \oplus W_{j-1} \oplus W_j \oplus W_{j+1} \oplus \cdots$$

in which, from left to right, the frequency of W_j changes from low to high. We denote that

$$V_j = \cdots \oplus W_{j-1} \oplus W_j, \quad (j \in \mathbb{Z}),$$

where V_j refers to the function set with lower frequency, and satisfies

- $\cdots \subset V_{-1} \subset V_0 \subset V_1 \subset \cdots$
- $\bigcap_{j \in \mathbb{Z}} V_j = \{0\}, \quad \overline{\bigcup_{j \in \mathbb{Z}} V_j} = L^2(\mathbb{R})$
- $f(x) \in V_j \iff f(2x) \in V_{j+1}$

We can also write $W_j = V_{j+1} \ominus V_j$. Therefore $\{W_j\}_{j \in \mathbb{Z}}$ can also be represented by $\{V_j\}_{j \in \mathbb{Z}}$. According to the above definitions, $L^2(\mathbb{R})$ can be divided into the conjoint ‘‘concentric annulus’’ or the expanding ‘‘concentric balls’’ $\{W_j\}_{j \in \mathbb{Z}}$ by $\{V_j\}_{j \in \mathbb{Z}}$. Now, we consider such a question, namely, is there a function ϕ with low frequency such that $\{\phi(\cdot - k)\}_{k \in \mathbb{Z}}$ is an orthonormal or a Riesz basis of V_0 ? The answer is affirmative. In this way, we can find ψ in from

$$\psi \in V_1 \ominus V_0.$$

By $\psi \in V_0 \subset V_1$, we obtain the two-scale equation as follows:

$$\phi(x) = 2 \sum_{k \in \mathbb{Z}} h_k \phi(2x - k), \quad \exists \{h_k\} \in l^2(\mathbb{Z}).$$

This is a famous framework referred to as the multiresolution analysis(MRA), by which wavelet can be constructed. In the following, we will show how to construct wavelet bases and the biorthonormal bases by using MRA. We will give the formal mathematical definitions of MRA.

Definition 3.1 Let H be a Hilbert space. Then a sequence in H , $\{e_j\}_{j=1}^\infty$ is said to be a Riesz basis of H , if the following conditions are satisfied:

- (1) $\overline{\text{span}}\{e_j\}_{j=1}^\infty = H$, i. e., $\forall x \in H$ and $\forall \varepsilon > 0$, there exists $\sum_{j=1}^n c_j e_j$, such that $\|x - \sum_{j=1}^n c_j e_j\| < \varepsilon$;

(2) Constants A and B exist, such that

$$A \sum_{j=1}^{\infty} |c_j|^2 \leq \left\| \sum_{j=1}^{\infty} c_j e_j \right\|^2 \leq B \sum_{j=1}^{\infty} |c_j|^2, \quad \forall \{c_j\}_{j=1}^{\infty} \in l^2.$$

A and B are called the lower and upper bounds of the Riesz basis, respectively.

Particularly, if $A = B = 1$, $\{e_j\}_{j=1}^{\infty}$ is said to be an orthonormal basis.

Definition 3.2 An closed subspace $\{V_j\}_{-\infty}^{\infty}$ of $L^2(\mathbb{R})$ is said to be a multiresolution analysis (MRA), if

- (1). $\cdots \subset V_{-1} \subset V_0 \subset V_1 \subset \cdots$
- (2). $\bigcap_{j \in \mathbb{Z}} V_j = \{0\}$, $\overline{\bigcup_{j \in \mathbb{Z}} V_j} = L^2(\mathbb{R})$
- (3). $f(x) \in V_j \iff f(2x) \in V_{j+1}$

(4). $\exists \phi \in V_0$, such that $\{\phi(\cdot - k)\}_{k \in \mathbb{Z}}$ is a Riesz basis of space V_0 . ϕ is called the scaling function of the MRA.

Especially, if the basis of (4) is orthonormal, the MRA is then called orthonormal multiresolution analysis. Furthermore, if $\{\phi(\cdot - k)\}_{k \in \mathbb{Z}}$ is the Riesz basis of V_0 with both upper bound B and low bound A , we call that ϕ generates a MRA with upper bound B and low bound A .

Note that, certainly, V_j is translation-invariant with $2^{-j}\mathbb{Z}$. Meanwhile, $\{\phi_{j,k}\}_{k \in \mathbb{Z}}$ is a Riesz basis of V_j (the upper and lower bounds are constants for any $j \in \mathbb{Z}$), here $\phi_{j,k}(x) := 2^{j/2}\phi(2^j x - k)$.

In the multiresolution analysis, the key point is how the scale function ϕ can be constructed. Because $\phi \in V_0 \subset V_1$, and $\{\phi(2 \cdot - k) | k \in \mathbb{Z}\}$ is the Riesz basis of V_1 , hence, $\{h_k\} \in l^2(\mathbb{Z})$ exists, such that

$$\phi(x) = 2 \sum_{k \in \mathbb{Z}} h_k \phi(2x - k). \quad (3.1)$$

Applying Fourier transform to both sides of the above equation, we can find that it is equivalent to the following fact: There exists $m_0(\xi) = \sum_{k \in \mathbb{Z}} h_k e^{-ik\xi} \in L^2(\mathbb{T})$, such that

$$\hat{\phi}(\xi) = m_0\left(\frac{\xi}{2}\right) \hat{\phi}\left(\frac{\xi}{2}\right). \quad (3.2)$$

Therefore, we can obtain

$$\hat{\phi}(\xi) = \left(\prod_{j=1}^n m_0\left(\frac{\xi}{2^j}\right) \right) \hat{\phi}\left(\frac{\xi}{2^n}\right), \quad (\forall n \in \mathbb{N}).$$

It means that if $\hat{\phi}(\xi)$ is continuous at $\xi = 0$ and $\phi(0) \neq 0$, we can infer $\prod_{j=1}^{\infty} m_0(2^{-j}\xi)$ is convergent, and

$$\hat{\phi}(\xi) = \prod_{j=1}^{\infty} m_0\left(\frac{\xi}{2^j}\right). \quad (3.3)$$

The construction of ϕ is then equivalent to looking for the function m_0 with $2\pi\mathbb{Z}$ -period. The $\{h_k\} \in l^2(\mathbb{Z})$ can be regarded as the mask of the two-scale equation, and the $m_0(\xi) = \sum_{k \in \mathbb{Z}} h_k e^{-ik\xi}$ can be considered to be the filter function.

As in the above discussion, most of wavelet bases can be constructed by multiresolution analysis, and the key point of the MRA is how to construct ϕ . Moreover, ϕ can be constructed by solving the two-scale equation (3.1) or (3.2). Therefore, one of the most important problems in MRA is:

- The existence of the solution for the two-scale equation (3.1) or (3.2).

From (3.2), we can obtain $\hat{\phi}$ which is represented by (3.3). However, it is only the formal solution, in fact, it is very difficult to find the analytical expression of ϕ for the following reasons: First, we do not know whether (3.3) is convergent. Second, we do not know whether $\hat{\phi}$ belongs to $L^2(\mathbb{R})$. Actually, the analytical expressions do not exist for most of the scaling functions, which are solutions of (3.1). A treatment to handle these problems is to study the convergence of the partial product $\prod_{j=1}^n m_0(2^{-j}\xi)$, and thereafter, to check if it converges in $L^2(\mathbb{R})$.

Although the two-scale equation has a solution ϕ in $L^2(\mathbb{R})$, it cannot guarantee the generation of MRA from ϕ . The main reason is that $\{\phi(\cdot - k)|k \in \mathbb{Z}\}$ must be a Riesz basis of V_0 . Consequently, the second important problem is:

- Whether the solutions of two-scale equation (3.1) can generate a MRA.

Proving that $\{\phi(\cdot - k)|k \in \mathbb{Z}\}$ is a Riesz basis of V_0 is equivalent to the proof of the $[\hat{\phi}, \hat{\phi}]$ has positive upper and lower bounds.

The orthonormality and biorthonormality are very important in multiresolution analysis and wavelet theory. Thus, the third important problem is:

- The orthonormality and biorthonormality of the solutions of two-scale equation.

The orthonormality and biorthonormality can strictly be defined below:

Definition 3.3 Suppose that $\tilde{\phi}, \phi \in L^2(\mathbb{R})$.

(1). If

$$\langle \tilde{\phi}(\cdot - k), \phi(\cdot) \rangle = \delta_{0,k} \quad (\forall k \in \mathbb{Z}),$$

then $\{\tilde{\phi}, \phi\}$ is said to be biorthonormal; furthermore, if $\{\phi, \phi\}$ is biorthonormal, we can say ϕ is orthonormal.

(2). If $\{\tilde{\phi}, \phi\}$ is biorthonormal, and

$$\tilde{V}_j := \{\tilde{\phi}_{j,k} | k \in \mathbb{Z}\}, \quad V_j := \{\phi_{j,k} | k \in \mathbb{Z}\}, \quad (j \in \mathbb{Z}),$$

form MRAs in $L^2(\mathbb{R})$ with scale functions $\tilde{\phi}$ and ϕ respectively, we say that $\{\tilde{\phi}, \phi\}$ generate a pair of biorthonormal MRAs; Particularly, if $\{\phi, \phi\}$ forms a pair of orthonormal MRAs, ϕ is said to generate an orthonormal MRA.

3.1.2 The Solution of Two-Scale Equation

• The General Solution of Two-Scale Equation

As mentioned above, the formal solution of the two-scale equation is

$$\prod_{j=1}^{\infty} m_0(\xi/2^j).$$

The following is a sufficient condition for its convergence:

Theorem 3.1 Let $m_0(\xi) \in C(\mathbb{T})$, and satisfy

$$\exists \varepsilon > 0, \text{ when } |\xi| \text{ is small enough : } m_0(\xi) = 1 + O(|\xi|^\varepsilon).$$

then, $\prod_{j=1}^{\infty} m_0(\xi/2^j)$ is uniformly convergent on any compact subset of \mathbb{R} , consequently, a continuous function can be defined in $C(\mathbb{R})$. Furthermore, there are positive constants C , τ and δ such that the following holds:

$$\left| \prod_{j=1}^n m_0(2^{-j}\xi) \right| \leq \begin{cases} \exp(\tau|\xi|^\varepsilon), & \text{when } |\xi| \leq \delta; \\ C|\xi|^\sigma, & \text{when } |\xi| > \delta, \end{cases} \quad (\forall n \in \mathbb{N} \cup \{+\infty\}),$$

where $\sigma := \log_2 \|m_0\|_{C(\mathbb{T})}$.

Proof. According to the condition, there must exist positive constants M and δ , such that

$$|m_0(\xi) - 1| \leq M|\xi|^\varepsilon < 1, \quad \forall |\xi| \leq \delta.$$

For each compact subset K of \mathbb{R} , we choose $J \in \mathbb{N}$ satisfying

$$|2^{-J}\xi| < \delta, \quad \forall \xi \in K.$$

Then, $\forall J_2 \geq J$, we have

$$\begin{aligned} \ln \left(\prod_{j=J}^{J_2} m_0(2^{-j}\xi) \right) &= \sum_{j=J}^{J_2} \ln m_0(2^{-j}\xi) \\ &= \sum_{j=J}^{J_2} \ln [1 + (m_0(2^{-j}\xi) - 1)] \\ &\leq \sum_{j=J}^{J_2} [m_0(2^{-j}\xi) - 1] \\ &\leq M \sum_{j=J}^{J_2} |2^{-j}\xi|^\varepsilon \\ &\leq M \frac{2^{-\varepsilon(J-1)}}{2^\varepsilon - 1} |\xi|^\varepsilon. \end{aligned} \quad (*)$$

Therefore, when J is large enough, we can write

$$\ln \left(\prod_{j=J}^{J_2} m_0(2^{-j}\xi) \right) = O(2^{-J\varepsilon}) \rightarrow 0, \quad (\forall \xi \in K, J \rightarrow \infty).$$

Consequently, $\prod_{j=1}^{\infty} m_0(\xi/2^j)$ is uniformly convergent on any compact subset of \mathbb{R} . Since $m_0 \in C(\mathbb{T})$, it defines a continuous function.

For the inequality, we only focus on $n \in \mathbb{N}$. It is obvious, for $n = \infty$.

$\forall \xi \in \mathbb{R}$ with $\xi \neq 0$, let $J_1 \in \mathbb{Z}$ satisfy

$$\frac{1}{2}\delta < |2^{-J_1}\xi| \leq \delta.$$

Now, we denote $J := \max(1, J_1)$, then $\forall J_2 \geq J$, according to the formula (*), we have

$$\prod_{j=J}^{J_2} m_0(2^{-j}\xi) \leq \exp\left(M \frac{2^{-\varepsilon(J-1)}}{2^\varepsilon - 1} |\xi|^\varepsilon\right).$$

If $0 < |\xi| \leq \delta$, then $J_1 \leq 0$, $J = 1$, and $\forall n \in \mathbb{N}$, we denote $J_2 = n$, and obtain

$$\left| \prod_{j=1}^m m_0(2^{-j}\xi) \right| = \prod_{j=J}^m m_0(2^{-j}\xi) \leq \exp\left(\frac{M}{2^\varepsilon - 1} |\xi|^\varepsilon\right) = \exp(\tau |\xi|^\varepsilon),$$

where $\tau := \frac{M}{2^\varepsilon - 1} > 0$.

When $\xi = 0$, because $m_0(0) = 1$, it satisfies the inequality.

If $|\xi| > \delta$, then $J_1 \geq 1$, $J = J_1$, therefore, when $n \geq J_1$, we can deduce that

$$\begin{aligned} \left| \prod_{j=1}^m m_0(2^{-j}\xi) \right| &= \left| \prod_{j=1}^{J_1-1} m_0(2^{-j}\xi) \right| \prod_{j=J}^m m_0(2^{-j}\xi) \\ &\leq (\|m_0\|_{C(\mathbb{T})})^{J_1-1} \exp\left(M \frac{2^{-\varepsilon(J-1)}}{2^\varepsilon - 1} |\xi|^\varepsilon\right) \\ &= 2^{(J_1-1) \log_2 \|m_0\|_{C(\mathbb{T})}} \exp\left(M \frac{2^\varepsilon}{2^\varepsilon - 1} |2^{-J_1}\xi|^\varepsilon\right) \\ &\leq \left(\frac{|\xi|}{\delta}\right)^\sigma \exp\left(M \frac{2^\varepsilon}{2^\varepsilon - 1} \delta^\varepsilon\right) \\ &= C|\xi|^\sigma, \end{aligned}$$

where,

$$\sigma := \log_2 \|m_0\|_{C(\mathbb{T})}, \quad C := \delta^{-\sigma} \exp\left(M \frac{2^\varepsilon}{2^\varepsilon - 1} \delta^\varepsilon\right).$$

For $n < J_1$, by $\|m_0\|_{C(\mathbb{T})} \geq |m_0(0)| = 1$, we have

$$\left| \prod_{j=1}^m m_0(2^{-j}\xi) \right| \leq (\|m_0\|_{C(\mathbb{T})})^m \leq (\|m_0\|_{C(\mathbb{T})})^{J_1-1} \leq \left(\frac{|\xi|}{\delta}\right)^\sigma \leq C|\xi|^\sigma.$$

This concludes our proof. ■

Note: Especially, if m_0 is a trigonometric polynomial and $m_0(0) = 1$, then, the conditions of this theorem are satisfied.

We hope that $\prod_{j=1}^{\infty} m_0(\xi/2^j) \in L^2(\mathbb{R})$, so that the inverse Fourier transform of ϕ can be also in $L^2(\mathbb{R})$. This means that ϕ is a solution of equation (3.1). Unfortunately, under the conditions of (3.1), $\prod_{j=1}^{\infty} m_0(\xi/2^j)$ may not be in $L^2(\mathbb{R})$, which can be shown by a simple example, namely, let $m_0(\xi) \equiv 1$, then $\prod_{j=1}^{\infty} m_0(\xi/2^j) \equiv 1 \notin L^2(\mathbb{R})$. We can handle this problem in two ways: First, we can enhance the conditions for $m_0(\xi)$, so as to ensure $\prod_{j=1}^{\infty} m_0(\xi/2^j) \in L^2(\mathbb{R})$. Second, we can discuss the solution of the equation(3.1) in a larger space over $L^2(\mathbb{R})$. We will now focus on the second way. In the conditions of (3.1), $\prod_{j=1}^{\infty} m_0(\xi/2^j)$ is continuous and increase at most polynomially fast at infinite. It belongs to φ' , the space of all the slowly increasing generalized functions, where the operation of Fourier transform is very convenient. We will discuss the solution of two-scale equation in φ' .

As we know, the space φ' consists of all the continuous linear functionals on φ . φ is the space of all the fast decreasing C^∞ functions, whose definition is as below:

$$\varphi := \{\varphi \in C^\infty(\mathbb{R}) \mid \sup_{x \in \mathbb{R}} (1+|x|^2)^{\frac{k}{2}} |\partial^\alpha \varphi(x)| \leq M_{k,\alpha} < \infty \text{ } (k, |\alpha| = 0, 1, \dots)\}.$$

By equipping countable semi-norms as follows:

$$\|\varphi\|_m := \sup_{\substack{|\alpha| \leq m \\ x \in \mathbb{R}}} (1+|x|^2)^{\frac{m}{2}} |\partial^\alpha \varphi(x)|, \quad (m = 0, 1, \dots),$$

φ becomes a countable-norm space, i.e., a B_0^* space (see [Zhang, 1986]).

For $1 \leq p \leq \infty$, we define

$$PL^p := PL^p(\mathbb{R}) := \{f \mid \exists m \in \mathbb{N} : (1+|x|)^{-m} f(x) \in L^p(\mathbb{R})\}. \quad (3.4)$$

Let $f \in PL^1$, and we define

$$\langle f, g \rangle := \int_{\mathbb{R}} f(x)g(x)dx \quad (\forall g \in \varphi).$$

It is obvious that

$$PL^\infty \subset PL^p \subset PL^1 \subset \varphi' \quad (1 \leq p \leq \infty).$$

Now, we will prove that the following two equations are equivalent to each other:

$$\phi = 2 \sum_{k \in \mathbb{Z}} h_k \phi(2 \cdot -k) \quad (\text{In } \varphi', \text{ the right term is convergent}) \quad (3.5)$$

$$\hat{\phi} = m_0\left(\frac{\cdot}{2}\right)\hat{\phi}\left(\frac{\cdot}{2}\right) \quad (\text{In } \varphi') \quad (3.6)$$

Theorem 3.2 Let $m_0(\xi) = \sum_{k \in \mathbb{Z}} h_k e^{-ik\xi} \in C(\mathbb{T})$, and $\phi \in \varphi'$. If one of the following two conditions is satisfied:

(1). $\hat{\phi} \in PL^1$ and $\sum_{k \in \mathbb{Z}} |h_k| < \infty$;

(2). $\sum_{k \in \mathbb{Z}} |h_k| |k|^n < \infty$ ($\forall n \in \mathbb{N}$),

then, ϕ satisfies (3.5) if and only if ϕ satisfies (3.6).

Proof. $\forall \phi \in \varphi'$, $g \in \varphi$, we denote

$$I_n := \lim_{n \rightarrow \infty} \left\langle \left(\sum_{|k| \leq n} h_k e^{-ik\frac{\xi}{2}} \right) \hat{\phi}\left(\frac{\cdot}{2}\right) - m_0\left(\frac{\cdot}{2}\right) \hat{\phi}\left(\frac{\cdot}{2}\right), g \right\rangle.$$

The satisfaction of the condition (1) produces

$$\begin{aligned} |I_n| &= \left| \int_{\mathbb{R}} \hat{\phi}\left(\frac{\xi}{2}\right) \left(\sum_{|k| > n} h_k e^{-ik\frac{\xi}{2}} \right) g(\xi) d\xi \right| \\ &\leq \left(\sum_{|k| > n} |h_k| \right) \int_{\mathbb{R}} \left| \hat{\phi}\left(\frac{\xi}{2}\right) g(\xi) \right| d\xi \\ &\rightarrow 0 \quad (n \rightarrow \infty). \end{aligned}$$

The satisfaction of the condition (2) yields

$$g_n(\xi) := \left(\sum_{|k| > n} h_k e^{-ik\frac{\xi}{2}} \right) g(\xi) \in \varphi.$$

Consequently, we have, $\forall m \in \mathbb{Z}^+$,

$$\begin{aligned} &\|g_n\|_m \\ &= \sup_{\xi \in \mathbb{R}, |\alpha| \leq m} \left| (1 + |\xi|^2)^{\frac{m}{2}} D^\alpha g_n(\xi) \right| \\ &= \sup_{\xi \in \mathbb{R}, |\alpha| \leq m} \left| \left(1 + |\xi|^2\right)^{\frac{m}{2}} \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} \left(\sum_{|k| > n} h_k (-i\frac{k}{2})^\beta e^{-ik\frac{\xi}{2}} \right) D^{\alpha-\beta} g(\xi) \right| \end{aligned}$$

$$\begin{aligned}
&\leq \sup_{\xi \in \mathbb{R}, |\alpha| \leq m} |(1 + |\xi|^2)^{\frac{m}{2}} \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} (\sum_{|k| > n} |h_k| |k|^{\beta}) |D^{\alpha - \beta} g(\xi)| \\
&\leq C (\sum_{|k| > n} |h_k| |k|^m) \sup_{\xi \in \mathbb{R}, |\alpha| \leq m} |(1 + |\xi|^2)^{\frac{m}{2}} |D^\alpha g(\xi)| \\
&= C (\sum_{|k| > n} |h_k| |k|^m) \|g\|_m \\
&\rightarrow 0 \quad (n \rightarrow \infty),
\end{aligned}$$

where, C stands for a constant which depends only on m . Thus, we can obtain $g_n \rightarrow 0$ (in φ as $n \rightarrow 0$). Therefore, we have

$$|I_n| = \left| \langle \hat{\phi}(\frac{\cdot}{2}), \left(\sum_{|k| > n} h_k e^{-ik\frac{\cdot}{2}} \right) g \rangle \right| = \left| \langle \hat{\phi}(\frac{\cdot}{2}), g_n \rangle \right| \rightarrow 0 \quad (n \rightarrow 0).$$

It implies that

$$\begin{aligned}
\phi \text{ satisfies (3.5)} &\iff \hat{\phi} = 2 \sum_{k \in \mathbb{Z}} h_k (\phi(2 \cdot -k))^\circ \\
&\iff \hat{\phi} = \lim_{n \rightarrow \infty} \sum_{|k| \leq n} h_k e^{-ik\frac{\cdot}{2}} \hat{\phi}(\frac{\cdot}{2}) \\
&\iff \langle \hat{\phi}, g \rangle = \lim_{n \rightarrow \infty} \langle \left(\sum_{|k| \leq n} h_k e^{-ik\frac{\cdot}{2}} \right) \hat{\phi}(\frac{\cdot}{2}), g \rangle \quad (\forall g \in \varphi) \\
&\iff \langle \hat{\phi} - m_0(\frac{\cdot}{2}) \hat{\phi}(\frac{\cdot}{2}), g \rangle \\
&= \lim_{n \rightarrow \infty} \langle \left[\sum_{|k| \leq n} h_k e^{-ik\frac{\cdot}{2}} - m_0(\frac{\cdot}{2}) \right] \hat{\phi}(\frac{\cdot}{2}), g \rangle \quad (\forall g \in \varphi) \\
&\iff \langle \hat{\phi} - m_0(\frac{\cdot}{2}) \hat{\phi}(\frac{\cdot}{2}), g \rangle = \lim_{n \rightarrow \infty} I_n = 0 \quad (\forall g \in \varphi) \\
&\iff \phi \text{ satisfy(3.6)}.
\end{aligned}$$

This establishes our proof. ■

As for the solution of two-scale equation in φ' , we have the following results.

Theorem 3.3 *Let $m_0(\xi) = \sum_{k \in \mathbb{Z}} h_k e^{-ik\xi} \in C(\mathbb{T})$, such that $\prod_{j=1}^{\infty} m_0(\frac{\xi}{2^j})$ converges pointwise to $M(\xi) \in PL^1$. Then, $\phi := \check{M} \in \varphi'$ is a solution of (3.6). Furthermore, if $M(\xi)$ is continuous at $\xi = 0$ and $M(0) = 1$, then ϕ*

is the unique solution of (3.6) in φ'_0 ($\varphi'_0 \subset \varphi$) which is defined by:

$$\varphi'_0 := \{\phi \in \varphi' \mid \hat{\phi} \in PL^1, \hat{\phi}(\xi) \text{ is continuous at } \xi = 0, \hat{\phi}(0) = 1\} \quad (3.7)$$

Proof. It is clear that

$$M(\xi) = m_0\left(\frac{\xi}{2}\right)M\left(\frac{\xi}{2}\right), \quad a.e. \xi \in \mathbb{R}.$$

Therefore, as a generalized function in φ' , we have $M(\cdot) = m_0\left(\frac{\cdot}{2}\right)M\left(\frac{\cdot}{2}\right)$, that is, $\phi := \check{M} \in \varphi'$ satisfies

$$\hat{\phi}(\cdot) = m_0\left(\frac{\cdot}{2}\right)\hat{\phi}\left(\frac{\cdot}{2}\right) \quad \text{in } \varphi'.$$

Consequently, ϕ is the solution of (3.6).

If $M(\xi)$ is continuous at $\xi = 0$, and $M(0) = 1$, it is obvious that, $\phi \in \varphi'_0$, i.e. $\phi := \check{M}$ is the solution of (3.6) in φ'_0 .

Now, we will prove the unicity of the solution. Suppose $\phi \in \varphi'_0$ is a solution of (3.6), hence, as a generalized function of φ' , it satisfies

$$\hat{\phi} = m_0\left(\frac{\cdot}{2}\right)\hat{\phi}\left(\frac{\cdot}{2}\right) \quad (\text{in } \varphi').$$

As $\hat{\phi} \in PL^1$, the above result is equivalent to that of ordinary function, thus, we obtain

$$\hat{\phi}(\xi) = m_0\left(\frac{\xi}{2}\right)\hat{\phi}\left(\frac{\xi}{2}\right) \quad a.e. \xi \in \mathbb{R}.$$

Therefore,

$$\hat{\phi}(\xi) = m_0\left(\frac{\xi}{2}\right)\hat{\phi}\left(\frac{\xi}{2}\right) = \cdots = \left(\prod_{j=1}^n m_0\left(\frac{\xi}{2^j}\right)\right)\hat{\phi}\left(\frac{\xi}{2^n}\right) \quad a.e. \xi \in \mathbb{R}.$$

Let $n \rightarrow \infty$, we have

$$\hat{\phi}(\xi) = \prod_{j=1}^{\infty} m_0\left(\frac{\xi}{2^j}\right) = M(\xi) \quad a.e. \xi \in \mathbb{R}.$$

As a conclusion, $\phi = \check{M}$ holds. The proof is complete. ■

Corollary 3.1 Let $m_0(\xi) = \sum_{k \in \mathbb{Z}} h_k e^{-ik\xi} \in C(\mathbb{T})$, such that

$$\sum_{k \in \mathbb{Z}} |h_k| < \infty,$$

and

$$\exists \varepsilon > 0, \text{ when } |\xi| \text{ is small enough, } m_0(\xi) = 1 + O(|\xi|^\varepsilon).$$

Then, (3.5) and (3.6) have the same solution in φ'_0 , which is just the inverse Fourier transform of $\prod_{j=1}^{\infty} m_0(\frac{\cdot}{2^j})$.

Proof. According to theorem 3.2, it is known that the two-scale equations (3.5) and (3.6) have the same solution in φ' . By Theorems 3.1 and 3.3, a unique solution exists in φ'_0 , which is just the inverse Fourier transform of $\prod_{j=1}^{\infty} m_0(\frac{\cdot}{2^j})$. ■

This result is enough for wavelet analysis in $L^2(\mathbb{R})$, although the solution of two-scale equation is in φ'_0 . To solve the unique solution of two-scale equation in $L^2(\mathbb{R})$ is much more difficult than that in φ' . Thus, we discussed the case of φ' first. Meanwhile, from the above analysis, we found that the unicity of the solution in φ'_0 implicates the unicity in $L^2(\mathbb{R})$. Consequently, our task will be changed to investigate the conditions which can guarantee that the solution in φ'_0 belongs to $L^2(\mathbb{R})$.

The Cascade Algorithm to Solve Two-Scale Equation

It is assumed that $m_0(\xi) = \sum_{k \in \mathbb{Z}} h_k e^{-ik\xi} \in C(\mathbb{R})$, we take $\eta_0 \in L^\infty(\mathbb{R}) \cap L^1(\mathbb{R})$, such that $\hat{\eta}_0(0) = 1$. Hence, the general scheme of the Cascade algorithm is

$$\eta_n(x) := 2 \sum_{k \in \mathbb{Z}} h_k \eta_{n-1}(2x - k), \quad (n = 1, 2, \dots). \quad (3.8)$$

In order to guarantee the convergence of the series in the Cascade algorithm, we suppose that $\sum_{j=1}^{\infty} |h_k| < \infty$. It is easy to see that the right hand side of the Cascade algorithm is absolutely convergent, $\eta_n \in L^\infty(\mathbb{R}) \cap L^1(\mathbb{R})$ ($n = 1, 2, \dots$), and

$$\|\eta_n\|_\infty \leq (2 \sum_{k \in \mathbb{Z}} |h_k|) \|\eta_{n-1}\|_\infty, \quad \|\eta_n\|_1 \leq (\sum_{k \in \mathbb{Z}} |h_k|) \|\eta_{n-1}\|_1,$$

for $n = 1, 2, \dots$. Since

$$\hat{\eta}_n(\xi) = \left(\prod_{j=1}^n m_0\left(\frac{\xi}{2^j}\right) \right) \hat{\eta}_0\left(\frac{\xi}{2^n}\right),$$

we have, under the condition of Theorem 3.1, that

$$\hat{\eta}_n(\xi) \rightarrow \prod_{j=1}^{\infty} m_0\left(\frac{\xi}{2^j}\right) \quad a.e. \quad \xi \in \mathbb{R},$$

thus, positive constants C and σ exist, such that

$$|\hat{\eta}_n(\xi)| \leq C(1 + |\xi|)^{\sigma}, \quad (\forall \xi \in \mathbb{R}).$$

Therefore, $\forall g \in \varphi$, using Lebesgue's dominated convergence theorem, we conclude that

$$\langle \hat{\eta}_n, g \rangle = \int_{\mathbb{R}} \hat{\eta}_n(\xi) g(\xi) d\xi \rightarrow 0, \quad (n \rightarrow \infty),$$

that is

$$\hat{\eta}_n \rightarrow \left(\prod_{j=1}^{\infty} m_0\left(\frac{\xi}{2^j}\right) \right) \quad \text{in } \varphi'.$$

Therefore,

$$\eta_n \rightarrow \left(\prod_{j=1}^{\infty} m_0\left(\frac{\xi}{2^j}\right) \right)^{\sim} \quad \text{in } \varphi'.$$

Consequently, the limit of $\{\eta_n\}$ in φ'_0 is just the unique solution of (3.5) in φ' . In summary, we have the following theorem.

Theorem 3.4 *Let $m_0(\xi) = \sum_{k \in \mathbb{Z}} h_k e^{-ik\xi} \in C(\mathbb{T})$ satisfy $\sum_{k \in \mathbb{Z}} |h_k| < \infty$, and*

$$\exists \varepsilon > 0, \text{ when } |\xi| \text{ is small enough, } m_0(\xi) = 1 + O(|\xi|^{\varepsilon}).$$

Let $\eta_0 \in L^\infty(\mathbb{R})$ be a function with compact support $[A, B]$, and satisfy $\hat{\eta}_0(0) = 1$. Then the function sequence $\{\eta_n\}$ defined by Cascade algorithm (3.8) is convergent in φ' . The limit is the unique solution of (3.5) in φ'_0 .

If $\{h_k\}_{k \in \mathbb{Z}}$ has finite length, that is, the nonzero items in $\{h_k\}_{k \in \mathbb{Z}}$ is finit, it can be proved that the solution of (3.5) has a compact support in φ'_0 . The support of $\phi \in \varphi'$ is defined as

$$\text{supp } \phi := \{x \in \mathbb{R} \mid \forall \text{ open set } O_x \ni x, \exists g \in C_c^\infty(O_x) : \langle \phi, g \rangle \neq 0\}, \quad (3.9)$$

where $C_c^\infty(O_x)$ denotes the space of all the C^∞ functions which is compactly supported in O_x .

Obviously, the support defined in the above is a closed set in \mathbb{R} . When ϕ degenerates to an ordinary locally integrable function, the definition of its support is the same as that of the ordinary meaning. The definition of the support of an ordinary function is defined by

$$\text{supp}\phi := \{x \in \mathbb{R}^d \mid \forall O_x \ni x, \text{ always has : } |O_x \cap \{\phi(x) \neq 0\}| > 0\}. \quad (3.10)$$

where, O_x denotes an open set containing x , and $|O_x \cap \{\phi(x) \neq 0\}|$ denotes the Lebesgue measure of the set $O_x \cap \{\phi(x) \neq 0\}$. In order to prove that the unique solution of the equation (3.5) in φ'_0 is compactly supported when $\{h_k\}_{k \in \mathbb{Z}}$ has finite length, we first consider the following lemma:

Lemma 3.1 *Let $\{h_k\}_{k \in \mathbb{Z}}$ be a complex sequence supported in $[N, M]$ in the meaning that $h_k = 0$ if $k \notin [N, M]$. Let $\eta_0 \in L^\infty(\mathbb{R})$ is compactly supported on $[A, B]$. Then, for the function sequence $\{\eta_n\}$ which is given by Cascade algorithm (3.8), we have*

$$\text{supp}\eta_n \subset [A^{(n)}, B^{(n)}], \quad (n = 0, 1, \dots),$$

where $A^{(n)}$, $B^{(n)}$ are defined by

$$\begin{cases} A^{(0)} = A \\ B^{(0)} = B \end{cases} \quad \begin{cases} A^{(n)} = \frac{1}{2}(N + A^{(n-1)}) \\ B_i^{(n)} = \frac{1}{2}(M + B^{(n-1)}) \end{cases}, \quad (n = 1, 2, \dots),$$

and satisfy

$\{A^{(n)}\}_{n=1}^\infty$ is a monotonous sequence between N and A , and $A^{(n)} \rightarrow N$ ($n \rightarrow \infty$);

$\{B^{(n)}\}_{n=1}^\infty$ is a monotonous sequence between M and B , and $B^{(n)} \rightarrow M$ ($n \rightarrow \infty$).

Proof. We use induction for n to prove $\text{supp}\eta_n \subset [A^{(n)}, B^{(n)}]$. Obviously, for $n = 0$, our conclusion is established. Suppose it is right for $n - 1$. As for $\text{supp}\eta_n$, let $\eta_n(x) \neq 0$, then, $\exists k \in [N, M]$, such that $2x - k \in \text{supp}\eta_{n-1} \subset [A^{(n-1)}, B^{(n-1)}]$. This means $A^{(n-1)} \leq 2x - k \leq B^{(n-1)}$ or $\frac{1}{2}(A^{(n-1)} + k) \leq x \leq \frac{1}{2}(B^{(n-1)} + k)$. Therefore, $\frac{1}{2}(A^{(n-1)} + N) \leq x \leq \frac{1}{2}(B^{(n-1)} + M)$. So far, we proved that

$$\text{supp}\eta_n \subset \left[\frac{(A^{(n-1)} + N)}{2}, \frac{B^{(n-1)} + M}{2} \right] = [A^{(n)}, B^{(n)}].$$

By induction, the result holds for $n \in \mathbb{Z}_+$.

We now discuss the monotonicity of $\{A^{(n)}\}_{n=0}^\infty$, in two cases:

If $A^{(0)} = A \geq N$, then, by $A^{(n)} = \frac{1}{2}(N + A^{(n-1)})$, it is easy to prove that $A^{(n)} \geq N$, so $A^{(n)} \leq A^{(n-1)} \leq \dots \leq A^{(0)}$. i. e. $\{A^{(n)}\}_{n=1}^{\infty}$ is a monotonous sequence valued between N and A .

If $A^{(0)} = A \leq N$, the same conclusion can be proved in the same way.

Leting $n \rightarrow \infty$ on the both sides of $A^{(n)} = \frac{1}{2}(N + A^{(n-1)})$, we obtain that $A^{(n)} \rightarrow N$.

We can give the proof in the same way for $B^{(n)}$.

The proof of the Lemma completes. ■

Theorem 3.5 *Let $\{h_k\}_{k \in \mathbb{Z}}$ be a complex sequence which is supported in $[N, M]$, $\sum_{k \in \mathbb{Z}} h_k = 1$. Then, the unique solution ϕ of equation (3.5) in φ'_0 is compactly supported on $[N, M]$.*

Proof. Let $\eta_0 \in L^\infty(\mathbb{R})$ satisfy $\text{supp}\eta_0 \subset [A, B]$, and $\eta_0(0) = 1$. Then, for the function sequence $\{\eta_n\}$ derived by the cascade algorithm, we have $\eta_n \rightarrow \phi$ (in φ'), where ϕ is the unique solution of (3.5)in φ'_0 , and

$$\text{supp}\eta_n \subset [A^{(n)}, B^{(n)}] \quad (n = 0, 1, \dots),$$

where $A^{(n)}$, $B^{(n)}$ satisfy

$$A^{(n)} \rightarrow N, \quad B^{(n)} \rightarrow M, \quad (n \rightarrow \infty).$$

$\forall x \notin [N, M]$, there exists a neighborhood O_x of x , such that the distance between $[N, M]$ and O_x is positive. Then, when n is large enough, we have

$$([A^{(n)}, B^{(n)}]) \bigcap O_x = \text{empty set}.$$

For any $g \in C_c^\infty(O_x)$, we have

$$\langle \phi, g \rangle = \lim_{n \rightarrow \infty} \langle \eta_n, g \rangle = 0.$$

Therefore, $x \notin \text{supp}\phi$. Now that x is an arbitrary point of $\mathbb{R} \setminus [N, M]$, we have

$$\mathbb{R} \setminus [N, M] \subset \mathbb{R} \setminus \text{supp}\phi,$$

i.e., $\text{supp}\phi \subset [N, M]$.

This ends the proof. ■

3.2 The Construction of MRAs

To facilitate the discussion, we assume the $2\pi\mathbb{Z}$ -periodic function, $m_0(\xi)$, is continuous on \mathbb{R} . We will study how MRAs can be constructed from such $m_0(\xi)$ that satisfies some basic conditions as follows.

Definition 3.4 Suppose $m_0(\xi)$ is $2\pi\mathbb{Z}$ -periodic measurable function on \mathbb{R} .

- m_0 is said to satisfy the basic condition I , if $m_0(\xi) \in C(\mathbb{T})$, $m_0(0) = 1$ and $m_0(\pi) = 0$;
- m_0 is said to satisfy the basic condition II , if $\prod_{j=1}^{\infty} m_0(\xi/2^j)$ a.e. converges to a non-zero function, which is denoted by $\hat{\phi}$ and is called the corresponding limit function, (where $\hat{\phi}$ denotes a function instead of the Fourier transform of a function temporarily.)

To study the sufficient conditions such that the solution of the two-scale equation belongs to $L^2(\mathbb{R})$, the transition operators will be introduced first.

Let \tilde{m}_0 , m_0 be $2\pi\mathbb{Z}$ -periodic functions on \mathbb{R} , we define:

$$Tf(\xi) := \left| m_0\left(\frac{\xi}{2}\right) \right|^2 f\left(\frac{\xi}{2}\right) + \left| m_0\left(\frac{\xi}{2} + \pi\right) \right|^2 f\left(\frac{\xi}{2} + \pi\right) \quad (3.11)$$

$$\tilde{T}f(\xi) := \left| \tilde{m}_0\left(\frac{\xi}{2}\right) \right|^2 f\left(\frac{\xi}{2}\right) + \left| \tilde{m}_0\left(\frac{\xi}{2} + \pi\right) \right|^2 f\left(\frac{\xi}{2} + \pi\right) \quad (3.12)$$

$$Sf(\xi) := \tilde{m}_0\left(\frac{\xi}{2}\right)\bar{m}_0\left(\frac{\xi}{2}\right)f\left(\frac{\xi}{2}\right) + \tilde{m}_0\left(\frac{\xi}{2} + \pi\right)\bar{m}_0\left(\frac{\xi}{2} + \pi\right)f\left(\frac{\xi}{2} + \pi\right) \quad (3.13)$$

$$|S|f(\xi) := \left| \tilde{m}_0\left(\frac{\xi}{2}\right)\bar{m}_0\left(\frac{\xi}{2}\right) \right| f\left(\frac{\xi}{2}\right) + \left| \tilde{m}_0\left(\frac{\xi}{2} + \pi\right)\bar{m}_0\left(\frac{\xi}{2} + \pi\right) \right| f\left(\frac{\xi}{2} + \pi\right) \quad (3.14)$$

The following lemma is important.

Lemma 3.2 Let $m_0 \in C(\mathbb{T})$ satisfy the basic condition II and the corresponding limit function $\hat{\phi} \in C(\mathbb{R})$. Let X be a closed subspace of $C(\mathbb{T})$ and satisfy both of the following conditions:

- (1). $T(X) \subset X$ and the spectral radius $r_{\sigma}(T|_X)$ of $T|_X$, which is the restriction of the transition operator T on X , satisfies $r_{\sigma}(T|_X) < 1$;
- (2). There exists a non-negative function $f \in X$ which has at most one zero point $\xi = 0$ in $[-\pi, \pi]$.

Then $\hat{\phi} \in L^2(\mathbb{R})$ and $\|\hat{\phi}_n - \hat{\phi}\|_2 \rightarrow 0$ ($n \rightarrow \infty$), where $\hat{\phi}_n$ is defined by

$$\hat{\phi}_n(\xi) := \prod_{j=1}^n m_0(2^{-j}\xi) \chi_{2^n[-\pi, \pi]}(\xi). \quad (3.15)$$

Proof. It can easily be deduced that $\hat{\phi}(0) = 1$. Let $\delta : 0 < \delta < \pi$ satisfy

$$\frac{1}{2} \leq |\hat{\phi}(\xi)| \leq 2, \quad \forall \xi \in [-\delta, \delta].$$

Since f is continuous and positive on $[-\pi, \pi] \setminus (-\frac{1}{2}\delta, \frac{1}{2}\delta)$, a constant $C > 0$ must exist, such that

$$C^{-1} \leq f(\xi) \leq C, \quad \forall \xi \in [-\pi, \pi] \setminus (-\frac{1}{2}\delta, \frac{1}{2}\delta).$$

We denote

$$E_n := [-2^n\delta, 2^n\delta], \quad (n = 0, 1, \dots).$$

Then $2^{-n}\xi \in E_0 = [-\delta, \delta]$ for any $\xi \in E_n$. Therefore,

$$|\hat{\phi}(\xi)| = |\hat{\phi}_n(\xi)\hat{\phi}(2^n\xi)| \leq |\hat{\phi}_n(\xi)| \sup_{\xi \in E_0} |\hat{\phi}(\xi)| \leq 2|\hat{\phi}_n(\xi)|.$$

Hence we have

$$\begin{aligned} \int_{E_n \setminus E_{n-1}} |\hat{\phi}(\xi)|^2 d\xi &\leq 4 \int_{E_n \setminus E_{n-1}} |\hat{\phi}_n(\xi)|^2 d\xi \\ &\leq 4 \int_{[-2^n\pi, 2^n\pi] \setminus E_{n-1}} |\hat{\phi}_n(\xi)|^2 d\xi \\ &\leq 4C \int_{[-2^n\pi, 2^n\pi] \setminus E_{n-1}} |\hat{\phi}_n(\xi)|^2 f(2^{-n}\xi) d\xi \\ &\leq 4C \int_{\mathbb{R}} |\hat{\phi}_n(\xi)|^2 f(2^{-n}\xi) d\xi \\ &= 4C \int_{\mathbb{T}} (T^n f)(\xi) d\xi \\ &= 4C \int_{\mathbb{T}} ((T|_X)^n f)(\xi) d\xi \\ &\leq 4C(2\pi) \|f\|_{C(\mathbb{T})} \|(T|_X)^n\|. \end{aligned}$$

Let ρ satisfy $r_\sigma(T|_X) < \rho < 1$. According to the spectral radius formula, we deduce that $\lim_{n \rightarrow \infty} \|(T|_X)^n\|^{1/n} = r_\sigma(T|_X)$. Thus there exists $N > 0$

such that $\|(T|x)^n\| \leq \rho^n$ for $n > N$, which concludes that

$$\int_{E_n \setminus E_{n-1}} |\hat{\phi}(\xi)|^2 d\xi \leq 4 \int_{[-2^n \pi, 2^n \pi] \setminus E_{n-1}} |\hat{\phi}_n(\xi)|^2 d\xi \leq 4C(2\pi) \|f\|_{C(\mathbb{T})} \rho^n.$$

Consequently,

$$\begin{aligned} \int_{\mathbb{R}} |\hat{\phi}(\xi)|^2 d\xi &= \int_{E_N} |\hat{\phi}(\xi)|^2 d\xi + \sum_{n=N+1}^{\infty} \int_{E_n \setminus E_{n-1}} |\hat{\phi}(\xi)|^2 d\xi \\ &\leq (2^{N+1} \delta) \|\hat{\phi}\|_{C(E_N)} + 4C(2\pi) \|f\|_{C(\mathbb{T})} \sum_{n=N+1}^{\infty} \rho^n \\ &< \infty. \end{aligned}$$

$\hat{\phi} \in L^2(\mathbb{R})$ is proved.

To prove $\|\hat{\phi}_n - \hat{\phi}\|_2 \rightarrow 0$ ($n \rightarrow \infty$). We deduce that

$$\begin{aligned} \|\hat{\phi}_n\|_2^2 &= \int_{\mathbb{R}} |\hat{\phi}_n(\xi)|^2 d\xi = \int_{[-2^n \pi, 2^n \pi]} |\hat{\phi}_n(\xi)|^2 d\xi \\ &= \int_{[-2^n \pi, 2^n \pi] \setminus E_{n-1}} |\hat{\phi}_n(\xi)|^2 d\xi + \int_{E_{n-1}} |\hat{\phi}_n(\xi)|^2 d\xi. \end{aligned}$$

Since $2^{-n} \xi \in [-\frac{1}{2} \delta, \frac{1}{2} \delta] \subset [-\delta, \delta]$ for any $\xi \in E_{n-1}$, we have

$$|\hat{\phi}(\xi)| = |\hat{\phi}_n(\xi) \hat{\phi}(2^{-n} \xi)| \geq \frac{1}{2} |\hat{\phi}_n(\xi)|.$$

Hence

$$|\hat{\phi}_n(\xi)|^2 \chi_{E_{n-1}}(\xi) \leq 4 |\hat{\phi}(\xi)|^2, \quad \forall \xi \in \mathbb{R}.$$

Now that $\hat{\phi}_n(\xi) \chi_{E_{n-1}}(\xi) \rightarrow \hat{\phi}(\xi)$ a.e. $\xi \in \mathbb{R}$, the Lebesgue dominated convergence theorem deduces that

$$\int_{E_{n-1}} |\hat{\phi}_n(\xi)|^2 d\xi = \int_{\mathbb{R}} |\hat{\phi}_n(\xi)|^2 \chi_{E_{n-1}}(\xi) d\xi \rightarrow \int_{\mathbb{R}} |\hat{\phi}(\xi)|^2 d\xi = \|\hat{\phi}\|_2^2,$$

which together with the following result

$$\int_{[-2^n \pi, 2^n \pi] \setminus E_{n-1}} |\hat{\phi}_n(\xi)|^2 d\xi \leq C(2\pi) \|f\|_{C(\mathbb{T})} \rho^n \rightarrow 0 \quad (n \rightarrow 0)$$

concludes that $\|\hat{\phi}_n\|_2 \rightarrow \|\hat{\phi}\|_2$ ($n \rightarrow \infty$). Note that $\hat{\phi}_n(\xi) \rightarrow \hat{\phi}(\xi)$ a.e. $\xi \in \mathbb{R}$ ($n \rightarrow \infty$), we obtain $\|\hat{\phi}_n - \hat{\phi}\|_2 \rightarrow 0$ ($n \rightarrow \infty$). Thus, the proof of the lemma is complete. \blacksquare

Particularly, if m_0 satisfies the basic conditions I and II, we set

$$X = \dot{C}(\mathbb{T}) := \{f \in C(\mathbb{T}) | f(0) = 0\} \quad \text{or} \quad X = \dot{\mathcal{P}}_N := \{f \in \mathcal{P}_N | f(0) = 0\}$$

corresponding to $m_0 \in C(\mathbb{T})$ or $m_0 \in \mathcal{P}_N^+$ respectively, where

$$\begin{aligned} \mathcal{P}_N &:= \left\{ \sum_{n=-N}^N c_n e^{-in\xi} \mid \forall \text{ sequence } \{c_n\}_{n \in \mathbb{Z}} \right\}, \\ \mathcal{P}_N^+ &:= \left\{ \sum_{n=0}^N c_n e^{-in\xi} \mid \forall \text{ sequence } \{c_n\}_{n \in \mathbb{Z}} \right\}. \end{aligned}$$

Then we have $T(X) \subset X$. According to the above lemma, we obtain the following theorem.

Theorem 3.6 *Suppose m_0 meets the basic conditions I and II and its limit function satisfies $\hat{\phi} \in C(\mathbb{R})$. If the spectral radius $r_\sigma(T|_{\dot{C}(\mathbb{T})})$ of $T|_X$, the restriction of T defined by (3.11), satisfies $r_\sigma(T|_{\dot{C}(\mathbb{T})}) < 1$, then $\hat{\phi} \in L^2(\mathbb{R})$ and $\|\hat{\phi}_n - \hat{\phi}\|_2 \rightarrow 0$ ($n \rightarrow \infty$), where $\hat{\phi}_n$ is defined by (3.15).*

Proof. According to the above lemma, the theorem holds obviously if we set $X = \dot{C}(\mathbb{T})$ and $f(\xi) := (1 - \cos \xi)^2$. ■

Now, we give a theorem for the case that m_0 is a trigonometric polynomial. The readers can refer to [Long, 1995] for the proof.

Theorem 3.7 *Let $m_0 \in \mathcal{P}_N^+$ ($N \neq 0$), $m_0(0) = 1, m_0(\pi) = 0$ and $\hat{\phi}$ be its limit function. We denote*

$$\dot{\mathcal{P}}_N := \{f \in \mathcal{P}_N | f(0) = 0\}.$$

Then the following three conditions are equivalent to each other:

(1). $\lambda = 1$ is the simple eigenvalue of the restriction $T|_{\mathcal{P}_N}$ of T on \mathcal{P}_N , and each of its other eigenvalues λ satisfies $|\lambda| < 1$, where T is defined by (3.11).

(2). each eigenvalue λ of the restriction $T|_{\dot{\mathcal{P}}_N}$ of T on $\dot{\mathcal{P}}_N$ satisfies $|\lambda| < 1$.

(3). $\hat{\phi} \in L^2(\mathbb{R})$ and $\|\hat{\phi}_n - \hat{\phi}\|_2 \rightarrow 0$ ($n \rightarrow \infty$), where $\hat{\phi}_n$ is defined by (3.15)

Furthermore, each of them implies the following results:

(i). $\Phi := [\hat{\phi}, \hat{\phi}] \in \mathcal{P}_N$ and Φ is the eigenvector of $T|_{\mathcal{P}_N}$ corresponding to the simple eigenvalue $\lambda = 1$

(ii). There exist constants $0 < \varepsilon < 2$ and $C > 0$ such that

$$\sum_{k \in \mathbb{Z}} |\hat{\phi}(\xi + 2k\pi)|^{2-\varepsilon} \leq C, \quad |\hat{\phi}(\xi)| \leq C(1 + |\xi|)^{-\varepsilon}, \quad (\forall \xi \in \mathbb{R}),$$

where

$$\Phi(\xi) := [\hat{\phi}, \hat{\phi}] := \sum_{k \in \mathbb{Z}} |\hat{\phi}(\xi + 2\pi k)|^2. \quad (3.16)$$

We consider whether ϕ generates an orthonormal MRA or a biorthogonal MRA. The estimation of the lower and upper bounds of $\Phi = [\hat{\phi}, \hat{\phi}]$ is key to this question. In mathematics, it is difficult since Φ is defined by an infinite sum and not easily to be expressed analytically in general. The above theorem, however, tells us that under certain conditions, Φ is just the eigenvector of $T|_{\mathcal{P}_N}$ corresponding to the simple eigenvalue $\lambda = 1$. The latter can be solved with a computer.

A sufficient condition such that $\hat{\phi} \in L^2(\mathbb{R})$ and $\|\hat{\phi}_n - \hat{\phi}\|_2 \rightarrow 0$ ($n \rightarrow \infty$) is given as follows. Its proof is omitted here and can be found in some references such as [Daubechies, 1992; Long, 1995].

Theorem 3.8 Suppose m_0 can be written as

$$m_0(\xi) = \left(\frac{1 + e^{-i\xi}}{2}\right)^L M_0(\xi)$$

where $L \in \mathbb{Z}_+$ and $M_0 \in C(\mathbb{T})$ satisfies:

- (1). $M_0(\pi) \neq 0$;
- (2). There exists a constant $0 < \delta < 1$ satisfying $M_0(\xi) = 1 + O(|\xi|^\delta)$ in some neighborhood of $\xi = 0$;
- (3). There exists $k \in \mathbb{N}$ such that

$$B_k := \max_{\xi \in \mathbb{R}} |M_0(\xi) \cdots M_0(2^{k-1}\xi)|^{1/k} < 2^{L-\frac{1}{2}}.$$

Then, $\hat{\phi} \in L^2(\mathbb{R})$, $\|\hat{\phi}_n - \hat{\phi}\|_2 \rightarrow 0$ ($n \rightarrow \infty$) and a constant $C > 0$ exists, such that

$$|\hat{\phi}(\xi)| \leq \frac{C}{(1 + |\xi|)^{\frac{1}{2} + \varepsilon}}, \quad (\forall \xi \in \mathbb{R}),$$

where $\varepsilon := L - \frac{1}{2} - \log B_k > 0$ and $\hat{\phi}_n$ is defined by (3.15).

The following is a sufficient and necessary condition, which can guarantee that ϕ generates a MRA.

Theorem 3.9 *Let m_0 satisfy the basic conditions I and II and $\hat{\phi} \in C(\mathbb{R})$ be its limit function. Then ϕ generates MRA if and only if both of the following two conditions hold:*

(1). $\hat{\phi} \in L^2(\mathbb{R})$ and $\|\hat{\phi}_n - \hat{\phi}\|_2 \rightarrow 0$ ($n \rightarrow \infty$), where $\hat{\phi}_n$ is defined by (3.15).

(2). Two positive constants A and B exist, such that $A \leq (T^n 1)(\xi) \leq B$ a.e. $\xi \in \mathbb{T}$ for any $n \in \mathbb{N}$, where T is the transition operator defined by (3.11).

Proof. To prove the necessity we assume that ϕ generates a MRA of $L^2(\mathbb{R})$. Two constants A and B exist, such that Φ defined by (3.16) satisfies

$$A \leq \Phi(\xi) \leq B \quad \text{a.e. } \xi \in \mathbb{T},$$

which concludes (1).

Now, we prove (2). Since T^n is a positive linear operator for any $n \in \mathbb{N}$, we deduce that $T^n f(\xi) \leq T^n g(\xi)$ a.e. ξ for any $f(\xi) \leq g(\xi)$ a.e. ξ . Using $T^n \Phi = \Phi$, we have:

$$\begin{aligned} A \leq \Phi(\xi) &= T^n \Phi(\xi) \leq (T^n B)(\xi) = B(T^n 1)(\xi), \quad \text{a.e. } \xi \in \mathbb{T}; \\ B \geq \Phi(\xi) &= T^n \Phi(\xi) \geq (T^n A)(\xi) = A(T^n 1)(\xi), \quad \text{a.e. } \xi \in \mathbb{T}. \end{aligned}$$

Therefore, $\frac{A}{B} \leq (T^n 1)(\xi) \leq \frac{B}{A}$ a.e. $\xi \in \mathbb{T}$. The proof of (2) is complete.

To prove the sufficiency, we suppose (1) and (2) hold. For any function g which is $2\pi\mathbb{Z}$ -periodic, bounded, non-negative and measurable, the following holds:

$$\int_{\mathbb{R}} |\hat{\phi}_n(\xi)|^2 g(\xi) d\xi = \int_{\mathbb{T}} (T^n 1)(\xi) g(\xi) d\xi.$$

Hence

$$A \int_T g(\xi) d\xi \leq \int_{\mathbb{R}} |\hat{\phi}_n(\xi)|^2 g(\xi) d\xi \leq B \int_T g(\xi) d\xi.$$

Let $n \rightarrow \infty$, we have

$$A \int_T g(\xi) d\xi \leq \int_{\mathbb{R}} |\hat{\phi}(\xi)|^2 g(\xi) d\xi \leq B \int_T g(\xi) d\xi,$$

i.e.

$$A \int_{\mathbb{T}} g(\xi) d\xi \leq \int_{\mathbb{T}} \Phi(\xi) g(\xi) d\xi \leq B \int_{\mathbb{T}} g(\xi) d\xi.$$

Thus, $A \leq \Phi(\xi) \leq B$ a.e. $\xi \in \mathbb{T}$ which implies consequently that $\{\phi(\cdot - k) | k \in \mathbb{Z}\}$ constitute a Riesz basis of $V_0 := \overline{\text{span}}\{\phi(\cdot - k) | k \in \mathbb{Z}\}$. By the fact of $\hat{\phi}(0) = m_0(0) = 1$ we deduce that $\overline{\cup_{j \in \mathbb{Z}} V_j} = L^2(\mathbb{R})$ for $V_j := \{f(2^j \cdot) | f \in V_0\}$ ($j \in \mathbb{Z}$). Therefore, ϕ generates a MRA of $L^2(\mathbb{R})$. This ends the proof of the theorem. ■

For the case that m_0 is trigonometric polynomial, we further have following result.

Corollary 3.2 *Let $m_0 \in \mathcal{P}_N^+$ (where $N \in \mathbb{Z}_+$, $N \neq 0$). Then $\hat{\phi}(\xi) := \prod_{j=1}^{\infty} m_0(\xi/2^j)$ converges a.e. $\xi \in \mathbb{R}$, belongs to $L^2(\mathbb{R})$ and ϕ generates a MRA, if and only if the following three conditions hold:*

(1). $m_0(0) = 1$ and $m_0(\pi) = 0$,

(2). $\lambda = 1$ is the simple eigenvalue of $T|_{\mathcal{P}_N}$ which is the restriction of T defined by (3.11) on \mathcal{P}_N , and each of its other eigenvalues λ satisfies $|\lambda| < 1$.

(3). The eigenvector $g \in \mathcal{P}_N$ of $T|_{\mathcal{P}_N}$ corresponding to the eigenvalue 1 has positive lower bound if $g(0) = 1$, that is, there is a constant $C > 0$ such that $g(\xi) \geq C$ ($\forall \xi \in \mathbb{T}$).

Proof. To prove the necessity, we choose $\xi_0 \in \mathbb{R}$ such that $\prod_{j=1}^{\infty} m_0(\xi_0/2^j)$ converges. Then $\lim_{j \rightarrow \infty} m_0(\xi_0/2^j) = 1$, i.e. $m_0(0) = 1$.

Since $\hat{\phi} \in L^2(\mathbb{R})$, and $m_0 \in \mathcal{P}_N^+$ implies that ϕ is compactly supported, we conclude that $\Phi := [\hat{\phi}, \hat{\phi}]$ is a trigonometric polynomial. According to the fact that ϕ generates a MRA, it can be shown easily that Φ has positive upper and lower bounds on \mathbb{T} . Therefore, (1) holds.

By Theorems 3.7 and 3.9, it can be concluded that (2) holds and $\Phi = [\hat{\phi}, \hat{\phi}]$ is the eigenvector of $T|_{\mathcal{P}_N}$ corresponding to the eigenvalue 1. Since 1 is the simple eigenvalue of $T|_{\mathcal{P}_N}$, we deduce that $g \in \mathcal{P}_N$ is just some constant times of Φ . Now, $g(0) = \Phi(0) = 1$, we further have $g \equiv \Phi$. Hence g has positive lower bound and (3) is proved.

To prove the sufficiency, we deduce that $\hat{\phi} \in L^2(\mathbb{R})$, $\Phi \in \mathcal{P}_N$ and Φ has positive lower bound according to Theorem 3.7 and condition (3). On the other hand, $\Phi \in \mathcal{P}_N$ implies that Φ has positive upper bound. Hence, Φ has positive upper and lower bounds, which together with the equality $\hat{\phi}(0) = 1$ concludes that ϕ generates a MRA. The proof is complete. ■

This corollary is more convenient to be applied in practice. It can be easily implemented with a computer.

3.2.1 The Biorthonormal MRA

The purpose of this section is to construct such $\tilde{\phi}$, $\phi \in L^2(\mathbb{R})$ that generate a pair of biorthonormal MRA. The first task is to construct the biorthonormal functions $\{\tilde{\phi}, \phi\}$, or equivalently, to find out the conditions for masks \tilde{m}_0 and m_0 , such that following two-scale relation:

$$\hat{\phi}(\xi) = m_0\left(\frac{\xi}{2}\right)\hat{\phi}\left(\frac{\xi}{2}\right) \quad (3.17)$$

$$\hat{\tilde{\phi}}(\xi) = \tilde{m}_0\left(\frac{\xi}{2}\right)\hat{\tilde{\phi}}\left(\frac{\xi}{2}\right) \quad (3.18)$$

Generally, it is very difficult to find out the sufficient and necessary conditions. To avoid such difficulty, we will first study some basic necessary conditions for $\{\tilde{m}_0, m_0\}$. Thereafter, by enhancing the conditions until we meet the sufficient ones.

Let $\{\tilde{\phi}, \phi\}$ generate a pair of biorthonormal MRA. There exists a constant $C > 0$ such that

$$C^{-1} \leq \tilde{\Phi}(\xi), \Phi(\xi) \leq C \quad a.e. \xi \in \mathbb{R},$$

where $\Phi(\xi)$ is defined by (3.16), and

$$\tilde{\Phi}(\xi) := [\hat{\tilde{\phi}}, \hat{\tilde{\phi}}] := \sum_{k \in \mathbb{Z}} |\hat{\tilde{\phi}}(\xi + 2k\pi)|^2. \quad (3.19)$$

Enhance the above inequality by letting

$$C^{-1} \leq \tilde{\Phi}(\xi), \Phi(\xi) \leq C \quad (\forall \xi \in \mathbb{R}).$$

Then it can be deduced that

$$m_0(0) = 1, m_0(\pi) = 0.$$

In order to define $\tilde{\phi}$, ϕ from \tilde{m}_0 , m_0 based on the two-scale relation, we assume that $\prod_{j=1}^{\infty} \tilde{m}_0(\xi/2^j)$ and $\prod_{j=1}^{\infty} m_0(\xi/2^j)$ are pointwise convergent. This assumption is necessary if $\hat{\tilde{\phi}}(\xi)$ and $\hat{\phi}(\xi)$ are expected to be continuous at $\xi = 0$ and are not zero functions.

According to the biorthonormal property, we have (see Theorem 3.11):

$$\tilde{m}_0(\xi)\bar{m}_0(\xi) + \tilde{m}_0(\xi+\pi)\bar{m}_0(\xi+\pi) = 1 \quad a.e. \xi \in \mathbb{R}.$$

To facilitate the discussion, we assume that $2\pi\mathbb{Z}$ -period functions $\tilde{m}_0(\xi)$ and $m_0(\xi)$ are continuous and satisfy the following basic conditions. We will study such sufficient conditions $\{\tilde{m}_0(\xi), m_0(\xi)\}$ satisfy, that $\{\hat{\phi}(\xi), \phi(\xi)\}$ generate a pair of biorthonormal.

Definition 3.5 Let $\{\tilde{m}_0(\xi), m_0(\xi)\}$ be a pair of $2\pi\mathbb{Z}$ -period measurable functions on \mathbb{R} . The following three conditions are called the basic conditions:

- Basic condition I: $\tilde{m}_0(\xi), m_0(\xi) \in C(\mathbb{T})$, $\tilde{m}_0(0) = m_0(0) = 1$, $\tilde{m}_0(\pi) = m_0(\pi) = 0$.
- Basic condition II: $\prod_{j=1}^{\infty} \tilde{m}_0(\xi/2^j)$ and $\prod_{j=1}^{\infty} m_0(\xi/2^j)$ are pointwise convergent to non-zero functions $\hat{\phi}$ and $\hat{\phi}$ respectively, which are called the corresponding limit functions. (Here $\hat{\phi}, \hat{\phi}$ denote two functions instead of the Fourier transforms of functions temporarily).
- Basic condition III: $\tilde{m}_0(\xi)\bar{m}_0(\xi) + \tilde{m}_0(\xi+\pi)\bar{m}_0(\xi+\pi) = 1$, a.e. $\xi \in \mathbb{R}$.

$\{\tilde{m}_0(\xi), m_0(\xi)\}$ is called to be basic if it satisfies the above three conditions.

Particularly, if $\tilde{m}_0 = m_0$, the corresponding basic conditions are defined below:

Definition 3.6 Let $m_0(\xi)$ be a $2\pi\mathbb{Z}$ -period measurable functions on \mathbb{R} . The following three conditions are called the basic conditions:

- Basic condition I: $m_0(\xi) \in C(\mathbb{T})$ and $m_0(0) = 1, m_0(\pi) = 0$.
- Basic condition II: $\prod_{j=1}^{\infty} m_0(\xi/2^j)$ is pointwise convergent to a non-zero function $\hat{\phi}$, which is called the corresponding limit functions. (Here $\hat{\phi}$ denotes a function instead of the Fourier transform of a function temporarily).
- Basic condition III: $|m_0(\xi)|^2 + |m_0(\xi+\pi)|^2 = 1$, a.e. $\xi \in \mathbb{R}$.

$m_0(\xi)$ is called to be basic if it satisfies the above three conditions.

Note 1: if $m_0(\xi)$ is basic and its limit function $\hat{\phi} \in L^2(\mathbb{R}) \cap C(\mathbb{R})$, then $\hat{\phi}(2\pi\alpha) = \delta_{0,\alpha}$ ($\forall \alpha \in \mathbb{Z}$).

Note 2: Differing from the case of function pair in Definition 3.5, the conditions $m_0(0) = 1$ and $m_0(\pi) = 0$ in the basic condition I of Definition 3.6 is implied in the basic conditons II and III of Definition 3.6. Furthermore, $\hat{\phi} \in L^2(\mathbb{R})$ can also be concluded by them. The details are included in the following theorem.

Theorem 3.10 *Assume that $\{\tilde{m}_0, m_0\} \subset C(\mathbb{R})$ satisfies the basic condition II, $\hat{\phi}, \bar{\phi}$ are the corresponding limit functions, then*

(1). *if there is a constant $B > 0$, such that operator $|S|$, which is defined by (3.14), satisfies $|S|^n 1(\xi) \leq B$ (a.e. $\xi \in \mathbb{T}$, for all $n \in N$), then, $\hat{\phi}\bar{\phi} \in L^1(\mathbb{R})$.*

(2). *if m_0 satisfies the basic condition III, then $\hat{\phi} \in L^2(\mathbb{R})$, $m_0(0) = 1$, and $m_0(\pi) = 0$.*

Proof. We will prove (1) first. Similar to (3.15), we denote

$$\hat{\phi}_n(\xi) := \prod_{j=1}^n \tilde{m}_0(2^{-j}\xi) \chi_{2^n[-\pi, \pi]}(\xi), \quad (3.20)$$

then it is easy to see that

$$\int_{\mathbb{R}} |\hat{\phi}_n(x)\bar{\phi}_n(x)| dx = \int_{\mathbb{R}} (|S|^n 1)(\xi) d\xi \leq B(2\pi) \quad (\forall n \in \mathbb{N}).$$

Since

$$\hat{\phi}_n(\xi) \rightarrow \hat{\phi}(\xi) \quad \hat{\phi}_n(\xi) \rightarrow \hat{\phi}(\xi) \quad a.e. \xi \in \mathbb{R}.$$

By Fatou lemma, letting $n \rightarrow \infty$, we have

$$\int_{\mathbb{R}} |\hat{\phi}(\xi)\bar{\phi}(\xi)| d\xi \leq B2\pi.$$

This establishes $\hat{\phi}\bar{\phi} \in L^1(\mathbb{R})$.

To prove (2), let $\hat{\phi} = \hat{\phi}$. It is clear that the corresponding operator $|S|$ satisfies $|S|^n 1(\xi) \equiv 1$ (a.e. $\xi \in \mathbb{T}$, $\forall n \in \mathbb{N}$), therefore, $\hat{\phi}\bar{\phi} \in L^1(\mathbb{R})$, or equivalently, $\hat{\phi} \in L^2(\mathbb{R})$. Since

$$\hat{\phi}_{n+1}(\xi) = m_0(\xi/2^{n+1})\hat{\phi}_n(\xi/2), \quad (\forall \xi \in 2^n[-\pi, \pi]),$$

by letting $n \rightarrow \infty$, we have

$$\hat{\phi}(\xi) = m_0(0)\hat{\phi}(\xi) \quad a.e. \xi \in \mathbb{R}.$$

It can be observed easily that $\hat{\phi}(0) = 1$, therefore $m_0(0) = 1$. Furthermore, since m_0 satisfies the basic condition **III**, (2) is proven. ■

Theorem 3.11 *Let $\tilde{\phi}, \phi \in L^2(\mathbb{R})$. Then,*

$$\begin{aligned} \{\tilde{\phi}, \phi\} \text{ is biorthonormal} &\iff F(\xi) = 1 \quad \text{a.e. } \xi \in \mathbb{R} \\ &\implies \exists C > 0 : |F|(\xi) \geq C \quad \text{a.e. } \xi \in \mathbb{R} \\ &\implies \exists C > 0 : \Phi(\xi)\tilde{\Phi}(\xi) \geq C \quad \text{a.e. } \xi \in \mathbb{T}, \end{aligned}$$

where

$$F(\xi) := \sum_{k \in \mathbb{Z}} \hat{\tilde{\phi}}(\xi + 2k\pi) \bar{\hat{\phi}}(\xi + 2k\pi). \quad (3.21)$$

Furthermore,

(1). If there are $2\pi\mathbb{Z}$ -periodic measurable functions $\tilde{m}_0(\xi)$ and $m_0(\xi)$ such that two-scale equations (3.17) and (3.18) hold, then

$$\begin{aligned} \{\tilde{\phi}, \phi\} \text{ is biorthonormal} &\implies \\ \tilde{m}_0(\xi)\bar{m}_0(\xi) + \tilde{m}_0(\xi + \pi)\bar{m}_0(\xi + \pi) &= 1 \quad \text{a.e. } \xi \in \mathbb{R}. \end{aligned}$$

(2). If $\tilde{\Phi}, \Phi \in L^\infty(\mathbb{T})$, then

$\{\tilde{\phi}, \phi\}$ is biorthonormal \implies There is a constant $C > 0$, such that :

$$C^{-1} \leq \tilde{\Phi}(\xi), \Phi(\xi) \leq C \quad \text{a.e. } \xi \in \mathbb{T}.$$

Proof. By

$$\begin{aligned} \int_{\mathbb{R}} \tilde{\phi}(x - k) \bar{\phi}(x) dx &= \left(\frac{1}{2\pi} \right) \int_{\mathbb{R}} \hat{\tilde{\phi}}(\xi) \bar{\hat{\phi}}(\xi) e^{-ik\xi} d\xi \\ &= \left(\frac{1}{2\pi} \right) \int_{\mathbb{T}} F(\xi) e^{-ik\xi} d\xi \quad (\forall k \in \mathbb{Z}), \end{aligned}$$

we have

$$\int_{\mathbb{R}} \tilde{\phi}(x - k) \bar{\phi}(x) dx = \delta_{0,k} \quad (\forall k \in \mathbb{Z}) \iff F(\xi) = 1 \quad \text{a.e. } \xi \in \mathbb{T}.$$

(1). If there are $2\pi\mathbb{Z}$ -periodic measurable functions $\tilde{m}_0(\xi)$ and $m_0(\xi)$ such that two-scale equations (3.17) and (3.18) hold, by

$$F(2\xi) = \tilde{m}_0(\xi)\bar{m}_0(\xi)F(\xi) + \tilde{m}_0(\xi + \pi)\bar{m}_0(\xi + \pi)F(\xi + \pi),$$

we conclude that

$$F(\xi) = 1 \quad a.e. \xi \in \mathbb{R}$$

implies

$$\tilde{m}_0(\xi)\bar{m}_0(\xi) + \tilde{m}_0(\xi + \pi)\bar{m}_0(\xi + \pi) = 1, \quad a.e. \xi \in \mathbb{R}.$$

(2). If $\tilde{\Phi}, \Phi \in L^\infty(\mathbb{T})$, then

$$\begin{aligned} & \{\tilde{\phi}, \phi\} \text{ is biorthonormal} \\ \implies & \exists C > 0 : \Phi(\xi)\tilde{\Phi}(\xi) \geq C \quad a.e. \xi \in \mathbb{T} \\ \implies & \exists C > 0 : \tilde{\Phi}(\xi), \Phi(\xi) \geq C \quad a.e. \xi \in \mathbb{T} \\ \implies & \exists C > 0 : C^{-1} \leq \tilde{\Phi}(\xi), \Phi(\xi) \leq C \quad a.e. \xi \in \mathbb{T}. \end{aligned}$$

This finishes our proof. \blacksquare

Note: According to the above discussion, $\forall \tilde{\phi}, \phi \in L^2(\mathbb{R})$, if $\{\tilde{\phi}, \phi\}$ is biorthonormal and $\tilde{\Phi}, \Phi \in L^\infty(\mathbb{T})$, then $\{\tilde{\phi}(\cdot - k)|k \in \mathbb{Z}\}$ is a Riesz basis of $\tilde{V}_0 := \overline{\text{span}}\{\tilde{\phi}(\cdot - k)|k \in \mathbb{Z}\}$ and $\{\phi(\cdot - k)|k \in \mathbb{Z}\}$ is a Riesz basis of $V_0 := \overline{\text{span}}\{\phi(\cdot - k)|k \in \mathbb{Z}\}$. Furthermore, if $\tilde{\phi}, \phi \in L(\mathbb{R})$, then $\hat{\tilde{\phi}}(0) = \hat{\phi}(0) = 1$. Consequently, biorthonormal MRA $\{\tilde{V}_j\}, \{V_j\}$ can be generated from $\{\tilde{\phi}, \phi\}$, here,

$$\tilde{V}_j := \{f(2^j \cdot) | f \in \tilde{V}_0\}, \quad V_j := \{f(2^j \cdot) | f \in V_0\} (\forall j \in \mathbb{Z}).$$

Theorem 3.12 Let $\{\tilde{m}_0(\xi), m_0(\xi)\}$ be basic, $\{\hat{\tilde{\phi}}, \hat{\phi}\} \subset L^2(\mathbb{R}) \cap C(\mathbb{R})$ be their limit functions. Suppose $\{\hat{\phi}_n\}$ is defined by (3.15) and $\{\hat{\tilde{\phi}}_n\}$ is defined by (3.20). Then,

$$\begin{aligned} \{\tilde{\phi}, \phi\} \text{ is biorthonormal} & \iff F(\xi) = 1 \quad a.e. \xi \in \mathbb{R} \\ & \iff \exists C > 0 : |F|(\xi) \geq C \quad a.e. \xi \in \mathbb{R} \\ & \iff \exists C > 0 : \Phi(\xi)\tilde{\Phi}(\xi) \geq C \quad a.e. \xi \in \mathbb{T} \\ & \iff \|\hat{\tilde{\phi}}_n \overline{\hat{\phi}}_n - \hat{\tilde{\phi}} \overline{\hat{\phi}}\|_1 \rightarrow 0 (n \rightarrow \infty). \end{aligned}$$

Furthermore, if $\tilde{\Phi}, \Phi \in L^\infty(\mathbb{T})$, then

$$\begin{aligned} \{\tilde{\phi}, \phi\} & \text{ can generate a pair of biorthonormal MRAs} \\ \iff & F(\xi) = 1 \quad a.e. \xi \in \mathbb{R} \\ \iff & \exists C > 0 : |F|(\xi) \geq C \quad a.e. \xi \in \mathbb{R} \end{aligned}$$

$$\begin{aligned} &\iff \exists C > 0 : \Phi(\xi)\tilde{\Phi}(\xi) \geq C \quad a.e. \xi \in \mathbb{T} \\ &\iff \|\hat{\tilde{\phi}}_n \bar{\hat{\phi}}_n - \hat{\tilde{\phi}} \bar{\hat{\phi}}\|_1 \rightarrow 0 \quad (n \rightarrow \infty). \end{aligned}$$

Proof. According to Theorem 3.11, we only need to prove that

$$\exists C > 0 : \Phi(\xi)\tilde{\Phi}(\xi) \geq C \quad a.e. \xi \in \mathbb{T} \implies \|\hat{\tilde{\phi}}_n \bar{\hat{\phi}}_n - \hat{\tilde{\phi}} \bar{\hat{\phi}}\|_1 \rightarrow 0 \quad (n \rightarrow \infty); \quad (3.22)$$

and

$$\|\hat{\tilde{\phi}}_n \bar{\hat{\phi}}_n - \hat{\tilde{\phi}} \bar{\hat{\phi}}\|_1 \rightarrow 0 \quad (n \rightarrow \infty) \implies \{\tilde{\phi}, \phi\} \text{ is biorthonormal.} \quad (3.23)$$

It is known that

$$\begin{aligned} \int_{\mathbb{R}} \tilde{\phi}(x) \bar{\phi}(x-k) dx &= (\frac{1}{2\pi}) \int_{\mathbb{R}} \hat{\tilde{\phi}}(\xi) \bar{\hat{\phi}}(\xi) e^{ik\xi} d\xi \\ &= \lim_{n \rightarrow \infty} (\frac{1}{2\pi}) \int_{\mathbb{R}} \hat{\tilde{\phi}}_n(\xi) \bar{\hat{\phi}}_n(\xi) e^{ik\xi} d\xi \\ &= \delta_{0,k}. \end{aligned}$$

Therefore, the second implication is proven. We will ignore the proof of the first implication for short (see [Long, 1995]).

When $\tilde{\Phi}, \Phi \in L^\infty(\mathbb{T})$, the equivalent equations in the theorem can easily be proved based on the note of Theorem 3.11. ■

A key problem of the theory of the wavelet construction is to find out the sufficient conditions for $\{\tilde{m}_0, m_0\}$ to ensure that $\{\tilde{\phi}, \phi\}$ is biorthonormal. Two important results are given below.

Theorem 3.13 Let m_0 satisfy the conditions in Theorem 3.8, and

$$|m_0(\xi)|^2 + |m_0(\xi + \pi)|^2 = 1, \quad \forall \xi \in \mathbb{R}.$$

Then, $\prod_{j=1}^{\infty} m_0(\xi/2^j)$ is pointwise convergent to $\hat{\phi} \in L^2(\mathbb{R}) \cap C(\mathbb{R})$, and ϕ can generate an orthonormal MRA.

Proof. By Theorems 3.8 and 3.12, it is clear that the theorem holds. ■

The famous Daubechies wavelet is based on this construction method. The key problem is to ensure that m_0 satisfies both Theorem 3.8 and the basic condition III. This means to construct the M_0 in Theorem 3.8, and

ensure that m_0 satisfies the basic condition III (see Chapter 4 of this book and [Daubechies, 1992]).

Theorem 3.14 *Let $\{\tilde{m}_0, m_0\} \subset \mathcal{P}_N^+$ ($N \neq 0$) be basic; $\{\hat{\phi}, \tilde{\phi}\}$ be their limit functions. Then, the following two statements are equivalent:*

- (1). $\{\tilde{\phi}, \phi\}$ can generate a pair of biorthonormal MRAs.
- (2). Each eigenvalue λ of $\tilde{T}|_{\dot{\mathcal{P}}_N}$ and $T|_{\dot{\mathcal{P}}_N}$, which are the restrictions of the translation operators \tilde{T} and T in $\dot{\mathcal{P}}_N$, satisfies $|\lambda| < 1$.

Corollary 3.3 *Let $\{\tilde{m}_0, m_0\} \subset \mathcal{P}_N$ ($N \neq 0$) satisfy the basic condition I; $\{\hat{\phi}, \tilde{\phi}\}$ be their limit functions. Then, the following two statements are equivalent:*

- (1). $\hat{\phi}, \tilde{\phi} \in L^2(\mathbb{R})$, and $\{\tilde{\phi}, \phi\}$ can generate a pair of biorthonormal MRAs.
- (2). $\tilde{m}_0(\xi)\bar{m}_0(\xi) + \tilde{m}_0(\xi + \pi)\bar{m}_0(\xi + \pi) = 1$ a.e. $\xi \in \mathbb{R}$, and each eigenvalue λ of $\tilde{T}|_{\dot{\mathcal{P}}_N}$ and $T|_{\dot{\mathcal{P}}_N}$, which are the restrictions of the translation operators \tilde{T} and T in $\dot{\mathcal{P}}_N$, satisfies $|\lambda| < 1$.

Proof. Obviously, $\{\tilde{m}_0, m_0\}$ satisfies the basic condition II.

“(1)⇒(2)” : By (1) in Theorem 3.11, we obtain

$$\tilde{m}_0(\xi)\bar{m}_0(\xi) + \tilde{m}_0(\xi + \pi)\bar{m}_0(\xi + \pi) = 1 \quad \text{a.e. } \xi \in \mathbb{R},$$

i.e., $\{\tilde{m}_0, m_0\}$ satisfies the basic condition III. Thus, $\{\tilde{m}_0, m_0\}$ is basic. By Theorem 3.14, (2) is proven.

“(2)⇒(1)” : Since $\{\tilde{m}_0, m_0\}$ is basic, according to Theorem 3.14, (1) is true. ■

3.2.2 Examples of Constructing MRA

Examples of constructing MRAs based on Theorem 3.3 will be discussed in this subsection.

First of all, the basic condition I can be equivalently illustrated by masks. That is:

Theorem 3.15 *Let $m_0(\xi) := \sum_{k \in \mathbb{Z}} h_k e^{-ik\xi}$, where $\{h_k\} \in l^1(\mathbb{Z})$. Then*

$$m_0(0) = 1 \text{ and } m_0(\pi) = 0 \iff \sum_{k \in \mathbb{Z}} h_{2k} = \sum_{k \in \mathbb{Z}} h_{2k+1} = \frac{1}{2}.$$

Proof. For $\mu = 0$ or 1 , we have

$$\begin{aligned} m_0(\pi\mu) &= \sum_{k \in \mathbb{Z}} h_k e^{-i\pi k \mu} \\ &= (\sum_{k \in \mathbb{Z}} h_{2k}) + (\sum_{k \in \mathbb{Z}} h_{2k+1}) e^{-i\pi \mu}, \end{aligned}$$

or,

$$\begin{cases} m_0(0) = \sum_{k \in \mathbb{Z}} h_{2k} + \sum_{k \in \mathbb{Z}} h_{2k+1} \\ m_0(\pi) = \sum_{k \in \mathbb{Z}} h_{2k} - \sum_{k \in \mathbb{Z}} h_{2k+1} \end{cases}$$

Therefore,

$$\begin{aligned} \begin{cases} m_0(0) = 1 \\ m_0(\pi) = 0 \end{cases} &\iff \begin{cases} \sum_{k \in \mathbb{Z}} h_{2k} + \sum_{k \in \mathbb{Z}} h_{2k+1} = 1 \\ \sum_{k \in \mathbb{Z}} h_{2k} - \sum_{k \in \mathbb{Z}} h_{2k+1} = 0 \end{cases} \\ &\iff \sum_{k \in \mathbb{Z}} h_{2k} = \sum_{k \in \mathbb{Z}} h_{2k+1} = \frac{1}{2}. \end{aligned}$$

This finishes our proof. \blacksquare

For $m_0 \in \mathcal{P}_N^+$ ($N \neq 0$), by corollary 3.2, we see that whether ϕ , which is defined by m_0 , can generates a MRA, depends on the properties of the eigenvalues and eigenvectors of $T|_{\mathcal{P}_N}$. It is easy to see that $T|_{\mathcal{P}_N}$ can be represented by a $2N + 1$ -order matrix. Since the eigenvalues and eigenvectors of a matrix can be calculated easily with a computer, the construction of MRAs in this case is always feasible.

Let $m_0(\xi) = \sum_{k \in \mathbb{Z}} h_k e^{-ik\xi} \in \mathcal{P}_N^+$ ($N \neq 0$). $\{e^{-ik\xi} | k \in \mathbb{Z} \cap [-N, N]\}$ be a set of basis of \mathcal{P}_N . Our first step is to get the matrix of $T|_{\mathcal{P}_N}$ based on this basis. We have that

$$\begin{aligned} (Te^{-ilx})(\xi) &= |m_0(\frac{\xi}{2})|^2 e^{-il\frac{\xi}{2}} + |m_0(\frac{\xi}{2}\pi)|^2 e^{-il\frac{\xi}{2}} e^{-il\pi} \\ &= \left(|m_0(\frac{\xi}{2})|^2 + |m_0(\frac{\xi}{2} + \pi)|^2 e^{-il\pi} \right) e^{-il\frac{\xi}{2}} \\ &= \sum_{n \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} h_n \bar{h}_k e^{-i(n-k)\frac{\xi}{2}} [1 + e^{-i(n-k+l)\pi}] e^{-il\frac{\xi}{2}} \\ &= \sum_{n \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} h_n \bar{h}_k e^{-i(n-k+l)\frac{\xi}{2}} [1 + e^{-i(n-k+l)\pi}] \\ &= \sum_{n \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} h_n \bar{h}_{n+l-k} e^{-ik\frac{\xi}{2}} [1 + e^{-ik\pi}] \end{aligned}$$

$$= \sum_{k \in \mathbb{Z}} (2 \sum_{n \in \mathbb{Z}} h_n \bar{h}_{n+l-2k}) e^{-ik\xi},$$

Let $\{r_1, \dots, r_s\}$ ($s := 2N + 1$) be a permutation of the set $\mathbb{Z} \cap [-N, N]$, and

$$(Te^{-ir_1\xi}, \dots, Te^{-ir_s\xi}) = (e^{-ir_1\xi}, \dots, e^{-ir_s\xi})A,$$

where A , the corresponding matrix of $T|_{\mathcal{P}_N}$ based on the basis $\{e^{-ir_1\xi}, \dots, e^{-ir_s\xi}\}$, is a $2N + 1$ -order matrix:

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,s} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,s} \\ \vdots & \vdots & \vdots & \vdots \\ a_{s,1} & a_{s,2} & \cdots & a_{s,s} \end{pmatrix}.$$

The (k, l) -th element of A is:

$$a_{k,l} = 2 \sum_{n \in \mathbb{Z}} h_n \bar{h}_{n+r_l-2r_k} \quad (k, l = 1, \dots, s).$$

By corollary 3.2, we have

Theorem 3.16 *Let $m_0(\xi) = \sum_{k \in \mathbb{Z}} h_k e^{-ik\xi} \in \mathcal{P}_N^+$, ($N \neq 0$), $\{r_1, \dots, r_s\}$ be a permutation of $\{-N, -N+1, \dots, N\}$ ($s := 2N + 1$), and the (k, l) -th element of A is:*

$$a_{k,l} = 2 \sum_{n=0}^N h_n \bar{h}_{n+r_l-2r_k} \quad (k, l = 1, \dots, s).$$

Then, $\hat{\phi}(\xi) := \prod_{j=1}^{\infty} m_0(\xi/2^j)$ is pointwise convergent on \mathbb{R} and ϕ generates a MRA of $L^2(\mathbb{R})$, if and only if the following three conditions are satisfied:

- (1). $\sum_{k \in \mathbb{Z}} h_{2k} = \sum_{k \in \mathbb{Z}} h_{2k} = \frac{1}{2}$;
- (2). 1 is the simple eigenvalue of matrix A , and each of the other eigenvalues λ satisfies $|\lambda| < 1$;

(3). The eigenpolynomial, $g(\xi) := \sum_{k=1}^s b_k e^{-ir_k \xi}$, has a positive lower-bound, where $(b_1, \dots, b_s)^T$ is the eigenvector of matrix A , which corresponds to the eigenvalue 1 and satisfies $\sum_{k=1}^s b_k = 1$.

Proof. Due to that A is the corresponding matrix of $T|_{\mathcal{P}_N}$ based on the basis $\{e^{-ir_1\xi}, \dots, e^{-ir_s\xi}\}$ of \mathcal{P}_N , it is clear that (1) and (2) in the theorem are equivalent to (1) and (2) in corollary 3.2 respectively. We conclude that $g(\xi) := \sum_{k=1}^s b_k e^{-ir_k \xi}$ is the eigenvector of $T|_{\mathcal{P}_N}$ corresponding to the eigenvalue 1 , and satisfies $g(0) = 1$, if and only if, $(b_1, \dots, b_s)^T$ is the

eigenvector of A corresponding to the eigenvalue 1 and satisfies $\sum_{k=1}^s b_k = 1$. In fact, it is easy to see that

$$\begin{aligned} Tg(\xi) &= g(\xi) \\ \iff (Te^{-ir_1\xi}, \dots, Te^{-ir_s\xi}) \begin{pmatrix} b_1 \\ \vdots \\ b_s \end{pmatrix} &= (e^{-ir_1\xi}, \dots, e^{-ir_s\xi}) \begin{pmatrix} b_1 \\ \vdots \\ b_s \end{pmatrix} \\ \iff (e^{-ir_1\xi}, \dots, e^{-ir_s\xi}) A \begin{pmatrix} b_1 \\ \vdots \\ b_s \end{pmatrix} &= (e^{-ir_1\xi}, \dots, e^{-ir_s\xi}) \begin{pmatrix} b_1 \\ \vdots \\ b_s \end{pmatrix} \\ \iff A \begin{pmatrix} b_1 \\ \vdots \\ b_s \end{pmatrix} &= \begin{pmatrix} b_1 \\ \vdots \\ b_s \end{pmatrix}. \end{aligned}$$

Obviously, $g(0) = 1 \iff \sum_{k=1}^s b_k = 1$. By Corollary 3.2, our proof is complete. ■

For a natural permutation of $\mathbb{Z} \cap [-N, N] = \{-N, -N+1, \dots, N-1, N\}$:

$$r_1 = -N, \dots, r_k = -N - 1, \dots, r_{2N+1} = N.$$

The corresponding matrix A is a $2N + 1$ -order matrix whose (k, l) -th elements is:

$$a_{k,l} = 2 \sum_{n=0}^N h_n \bar{h}_{n+N+1+l-2k}, \quad (k, l = 1, \dots, 2N + 1),$$

where $h_k = 0$ for $k < 0$ and $k > N$.

To construct MRA, a necessary condition that $\{h_0, \dots, h_N\}$ should satisfy is

$$\sum_{k \in \mathbb{Z}, 0 \leq 2k \leq N} h_{2k} = \sum_{k \in \mathbb{Z}, 0 \leq 2k+1 \leq N} h_{2k+1} = \frac{1}{2}.$$

Now, for $N = 1, 2$, we discuss the conditions that $\{h_0, \dots, h_N\}$ should satisfy so that it can generate a MRA. We will focus on the three conditions in Theorem 3.16.

1). For $N = 1$, we have $h_0 = h_1 = \frac{1}{2}$. A is a 3-order matrix whose elements are:

$$a_{k,l} = 2(h_0\bar{h}_{2+l-2k} + h_1\bar{h}_{3+l-2k}) = h_{2+l-2k} + h_{3+l-2k}, \quad (k, l = 1, 2, 3).$$

Thus, A can be represented as:

$$A = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 1 & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} \end{pmatrix}.$$

The eigenvalues of A are $1, \frac{1}{2}$. The eigenvector corresponding to 1 is $(b_1, b_2, b_3)^t = (0, 1, 0)^t$, and the sum of its components equals to 1. Hence, the eigenpolynomial

$$g(\xi) = e^{-ir_2\xi} = e^{-i0\xi} \equiv 1$$

has a positive lower-bound. This tells us that there is only one MRA when $N = 1$. Now, we intend to find out its corresponding scale function. By

$$m_0(\xi) = e^{-i\frac{1}{2}\xi} \cos \frac{\xi}{2},$$

we have

$$\begin{aligned} \hat{\phi}(\xi) &= \lim_{n \rightarrow \infty} \prod_{j=1}^n m_0\left(\frac{\xi}{2^j}\right) \\ &= \lim_{n \rightarrow \infty} \prod_{j=1}^n \left(e^{-i\frac{\xi}{2^{j+1}}} \cos \frac{\xi}{2^{j+1}}\right) \\ &= e^{-i\frac{\xi}{2}} \lim_{n \rightarrow \infty} \prod_{j=1}^n \cos \frac{\xi}{2^{j+1}} \\ &= e^{-i\frac{\xi}{2}} \lim_{n \rightarrow \infty} \frac{\sin \frac{\xi}{2}}{2^n \sin \frac{\xi}{2^{n+1}}} \\ &= e^{-i\frac{\xi}{2}} \frac{2}{\xi} \sin \frac{\xi}{2} \end{aligned}$$

This is just the Fourier transform of the characteristic function $\chi_{[0,1]}(x)$ of interval $[0, 1]$. Therefore,

$$\phi(x) = \chi_{[0,1]}(x) := \begin{cases} 1 & \text{when } x \in [0, 1] \\ 0 & \text{otherwise} \end{cases}$$

As a conclusion, for $N = 1$, there is only one MRA, which is generated by $\chi_{[0,1]}(x)$.
2). For $N = 2$, we have

$$h_0 + h_2 = \frac{1}{2}, \quad h_1 = \frac{1}{2}.$$

Therefore, A is a 5-order matrix whose (k, l) -th elements is:

$$a_{k,l} = 2(h_0\bar{h}_{3+l-2k} + \frac{1}{2}\bar{h}_{4+l-2k} + h_2\bar{h}_{5+l-2k}), \quad (k, l = 1, \dots, 5).$$

Thus,

$$\begin{cases} a_{1,l} = 2h_0\bar{h}_{1+l} \\ a_{2,l} = 2(h_0\bar{h}_{l-1} + \frac{1}{2}\bar{h}_l + h_2\bar{h}_{l+1}) \\ a_{3,l} = 2(h_0\bar{h}_{l-3} + \frac{1}{2}\bar{h}_{l-2} + h_2\bar{h}_{l-1}) \\ a_{4,l} = 2(h_0\bar{h}_{l-5} + \frac{1}{2}\bar{h}_{l-4} + h_2\bar{h}_{l-3}) \\ a_{5,l} = 2h_2\bar{h}_{l-5} \end{cases}$$

Hence, matrix A can be written as:

$$\begin{pmatrix} h_0 - 2|h_0|^2 & 0 & 0 & 0 & 0 \\ 1 - h_0 - \bar{h}_0 + 4|h_0|^2 & \frac{1}{2} + h_0 - \bar{h}_0 & h_0 - 2|h_0|^2 & 0 & 0 \\ \bar{h}_0 - 2|h_0|^2 & \frac{1}{2} + \bar{h}_0 - h_0 & 1 - h_0 - \bar{h}_0 + 4|h_0|^2 & \frac{1}{2} + h_0 - \bar{h}_0 & h_0 - 2|h_0|^2 \\ 0 & 0 & \bar{h}_0 - 2|h_0|^2 & \frac{1}{2} + \bar{h}_0 - h_0 & 1 - h_0 - \bar{h}_0 + 4|h_0|^2 \\ 0 & 0 & 0 & 0 & \bar{h}_0 - 2|h_0|^2 \end{pmatrix}.$$

The eigenpolynomial of A is

$$\begin{aligned} |\lambda I - A| &= (\lambda - h_0 + 2|h_0|^2)(\lambda - \bar{h}_0 + 2|h_0|^2) \\ &\cdot \begin{vmatrix} \lambda - \frac{1}{2} - h_0 + \bar{h}_0 & -h_0 + 2|h_0|^2 & 0 \\ -\frac{1}{2} - \bar{h}_0 + h_0 & \lambda - 1 + h_0 + \bar{h}_0 - 4|h_0|^2 & -\frac{1}{2} - h_0 + \bar{h}_0 \\ 0 & -\bar{h}_0 + 2|h_0|^2 & \lambda - \frac{1}{2} - \bar{h}_0 + h_0 \end{vmatrix} \\ &= (\lambda - h_0 + 2|h_0|^2)(\lambda - \bar{h}_0 + 2|h_0|^2)(\lambda - 1) \\ &\cdot \begin{vmatrix} \lambda - \frac{1}{2} - h_0 + \bar{h}_0 & -h_0 + 2|h_0|^2 & 0 \\ \lambda - 1 & \lambda - 1 & \lambda - 1 \\ 0 & -\bar{h}_0 + 2|h_0|^2 & \lambda - \frac{1}{2} - \bar{h}_0 + h_0 \end{vmatrix} \\ &= (\lambda - h_0 + 2|h_0|^2)(\lambda - \bar{h}_0 + 2|h_0|^2)(\lambda - 1) \end{aligned}$$

$$\begin{aligned}
& \cdot \begin{vmatrix} \lambda - \frac{1}{2} - h_0 + \bar{h}_0 & -h_0 + 2|h_0|^2 & 0 \\ 1 & 1 & 1 \\ 0 & -\bar{h}_0 + 2|h_0|^2 & \lambda - \frac{1}{2} - \bar{h}_0 + h_0 \end{vmatrix} \\
& = (\lambda - h_0 + 2|h_0|^2)(\lambda - \bar{h}_0 + 2|h_0|^2)(\lambda - 1) \\
& \cdot \begin{vmatrix} \lambda - \frac{1}{2} - h_0 + \bar{h}_0 & 1 & 0 \\ -h_0 + 2|h_0|^2 & 1 & -\bar{h}_0 + 2|h_0|^2 \\ 0 & 1 & \lambda - \frac{1}{2} - \bar{h}_0 + h_0 \end{vmatrix} \\
& = -(\lambda - h_0 + 2|h_0|^2)(\lambda - \bar{h}_0 + 2|h_0|^2)(\lambda - 1) \\
& \cdot \begin{vmatrix} 1 & \lambda - \frac{1}{2} - h_0 + \bar{h}_0 & 0 \\ 1 & -h_0 + 2|h_0|^2 & -\bar{h}_0 + 2|h_0|^2 \\ 1 & 0 & \lambda - \frac{1}{2} - \bar{h}_0 + h_0 \end{vmatrix} \\
& = -(\lambda - h_0 + 2|h_0|^2)(\lambda - \bar{h}_0 + 2|h_0|^2)(\lambda - 1) \\
& \cdot \begin{vmatrix} 1 & \lambda - \frac{1}{2} - h_0 + \bar{h}_0 & 0 \\ 0 & -\lambda + \frac{1}{2} + 2|h_0|^2 - \bar{h}_0 & 2|h_0|^2 - \bar{h}_0 \\ 0 & h_0 - 2|h_0|^2 & \lambda - \frac{1}{2} + h_0 - 2|h_0|^2 \end{vmatrix} \\
& = -(\lambda - h_0 + 2|h_0|^2)(\lambda - \bar{h}_0 + 2|h_0|^2)(\lambda - 1) \\
& \cdot \begin{vmatrix} -\lambda + \frac{1}{2} & 2|h_0|^2 - \bar{h}_0 \\ -\lambda + \frac{1}{2} & \lambda - \frac{1}{2} + h_0 - 2|h_0|^2 \end{vmatrix} \\
& = (\lambda - h_0 + 2|h_0|^2)(\lambda - \bar{h}_0 + 2|h_0|^2)(\lambda - 1)(\lambda - \frac{1}{2}) \\
& \cdot \begin{vmatrix} 1 & 2|h_0|^2 - \bar{h}_0 \\ 1 & \lambda - \frac{1}{2} + h_0 - 2|h_0|^2 \end{vmatrix} \\
& = (\lambda - h_0 + 2|h_0|^2)(\lambda - \bar{h}_0 + 2|h_0|^2)(\lambda - 1)(\lambda - \frac{1}{2})
\end{aligned}$$

$$\begin{aligned} & \cdot \begin{vmatrix} 1 & 2|h_0|^2 - \bar{h}_0 \\ 0 & \lambda - \frac{1}{2} + h_0 + \bar{h}_0 - 4|h_0|^2 \end{vmatrix} \\ & = (\lambda - h_0 + 2|h_0|^2)(\lambda - \bar{h}_0 + 2|h_0|^2)(\lambda - 1)(\lambda - \frac{1}{2}) \\ & \quad \cdot (\lambda - \frac{1}{2} + h_0 + \bar{h}_0 - 4|h_0|^2). \end{aligned}$$

Therefore, the eigenvalues of A are:

$$1, \frac{1}{2}, h_0 - 2|h_0|^2, \bar{h}_0 - 2|h_0|^2, \frac{1}{2} + 4|h_0|^2 - h_0 - \bar{h}_0.$$

In order to generate MRA, the moduli of the above eigenvalues, except 1, must be smaller than 1, that is,

$$\begin{cases} |h_0 - 2|h_0|^2| < 1 \\ |\frac{1}{2} + 4|h_0|^2 - h_0 - \bar{h}_0| < 1 \end{cases}.$$

We now simplify these conditions as follows. By

$$\frac{1}{2}|h_0|^2 - h_0 - \bar{h}_0 = \frac{1}{2}\bar{h}_0 - h_0 - \bar{h}_0 = \frac{1}{4}|h_0 - \frac{1}{4}|^2 > 0.$$

we get

$$\left| \frac{1}{2} + 4|h_0|^2 - h_0 - \bar{h}_0 \right| < 1 \iff |h_0 - \frac{1}{4}| < \frac{\sqrt{3}}{4}.$$

For $|h_0 - \frac{1}{4}| < \frac{\sqrt{3}}{4}$, we have:

$$\begin{aligned} |h_0 - 2|h_0|^2| &= \sqrt{|2|h_0|^2 - Re(h_0)|^2 + |Im(h_0)|^2} \\ &= \sqrt{|2|h_0|^2 - \frac{1}{2}(h_0 + \bar{h}_0)|^2 + |Im(h_0)|^2} \\ &= \sqrt{\frac{1}{4}|4|h_0|^2 - h_0 - \bar{h}_0|^2 + |Im(h_0)|^2} \\ &= \sqrt{\frac{1}{4}\left|4|h_0 - \frac{1}{4}|^2 - \frac{1}{4}\right|^2 + |Im(h_0)|^2} \\ &< \sqrt{\frac{1}{4}\left(\frac{1}{2}\right)^2 + |Im(h_0)|^2} \\ &= \frac{1}{2}\sqrt{\frac{1}{4} + 4|Im(h_0)|^2}. \end{aligned}$$

And by

$$\left| \frac{1}{2} - h_0 + \bar{h}_0 \right| = \left| \frac{1}{2} - 2i\text{Im}(h_0) \right| = \sqrt{\frac{1}{4} + 4|\text{Im}(h_0)|^2},$$

we have

$$\begin{aligned} |h_0 - 2|h_0|^2| &< \frac{1}{2} \left| \frac{1}{2} - h_0 + \bar{h}_0 \right| \\ &= \frac{1}{2} \left| \frac{1}{2} - (h_0 - \frac{1}{4}) + (\bar{h}_0 - \frac{1}{2}) \right| \\ &< \frac{1}{2} \left(\frac{1}{2} + \frac{\sqrt{3}}{4} + \frac{\sqrt{3}}{4} \right) \\ &= \frac{1 + \sqrt{3}}{4} < 1, \end{aligned}$$

Hence

$$\begin{cases} |h_0 - 2|h_0|^2| < 1 \\ \left| \frac{1}{2} + 4|h_0|^2 - h_0 - \bar{h}_0 \right| < 1 \end{cases} \iff |h_0 - \frac{1}{4}| < \frac{\sqrt{3}}{4}.$$

Now we suppose $|h_0 - \frac{1}{4}| < \frac{\sqrt{3}}{4}$. To solve the eigenvector corresponding to 1, such that the sum of whose components equals to 1, we apply the elementary row transformation to matrix $I - A$ as follows:

$$\begin{aligned} I - A &\Rightarrow \\ &\left(\begin{array}{ccccc} 1-h_0+2|h_0|^2 & 0 & 0 & 0 & 0 \\ h_0+\bar{h}_0-1-4|h_0|^2 & \frac{1}{2}-h_0+\bar{h}_0 & 2|h_0|^2-h_0 & 0 & 0 \\ 2|h_0|^2-\bar{h}_0 & h_0-\bar{h}_0-\frac{1}{2} & -4|h_0|^2+h_0+\bar{h}_0 & \bar{h}_0-h_0-\frac{1}{2} & 2|h_0|^2-h_0 \\ 0 & 0 & 2|h_0|^2-\bar{h}_0 & \frac{1}{2}+h_0-\bar{h}_0 & h_0+\bar{h}_0-1-4|h_0|^2 \\ 0 & 0 & 0 & 0 & 1+2|h_0|^2-\bar{h}_0 \end{array} \right) \\ &\Rightarrow \left(\begin{array}{ccccc} 1-h_0+2|h_0|^2 & 0 & 0 & 0 & 0 \\ h_0+\bar{h}_0-1-4|h_0|^2 & \frac{1}{2}-h_0+\bar{h}_0 & 2|h_0|^2-h_0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2|h_0|^2-\bar{h}_0 & \frac{1}{2}+h_0-\bar{h}_0 & h_0+\bar{h}_0-1-4|h_0|^2 \\ 0 & 0 & 0 & 0 & 1+2|h_0|^2-\bar{h}_0 \end{array} \right) \\ &\Rightarrow \left(\begin{array}{ccccc} 1 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2}-h_0+\bar{h}_0 & 2|h_0|^2-h_0 & 0 & 0 \\ 0 & 0 & 2|h_0|^2-\bar{h}_0 & \frac{1}{2}+h_0-\bar{h}_0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right). \end{aligned}$$

By solving the following linear system:

$$\begin{cases} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} - h_0 + \bar{h}_0 & 2|h_0|^2 - h_0 & 0 & 0 \\ 0 & 0 & 2|h_0|^2 - \bar{h}_0 & \frac{1}{2} + h_0 - \bar{h}_0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \end{pmatrix} = 0 \\ b_1 + b_2 + b_3 + b_4 + b_5 = 1 \end{cases},$$

we obtain the eigenvector corresponding to the eigenvalue 1, so that the sum of its components equals to 1. It is

$$\begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \end{pmatrix} = \frac{8|\frac{1}{2} - h_0 + \bar{h}_0|^2}{3 - 16|h_0 - \frac{1}{4}|^2} \begin{pmatrix} 0 \\ \frac{h_0 - 2|h_0|^2}{\frac{1}{2} - h_0 + \bar{h}_0} \\ 1 \\ \frac{\bar{h}_0 - 2|h_0|^2}{\frac{1}{2} - h_0 + \bar{h}_0} \\ 0 \end{pmatrix}$$

Thus, the eigenpolynomial is:

$$\begin{aligned} g(\xi) &= b_1 e^{i2\xi} + b_2 e^{i\xi} + b_3 + b_4 e^{-i\xi} + b_5 e^{-i2\xi} \\ &= \frac{8|\frac{1}{2} - h_0 + \bar{h}_0|^2}{3 - 16|h_0 - \frac{1}{4}|^2} \left(\frac{h_0 - 2|h_0|^2}{\frac{1}{2} - h_0 + \bar{h}_0} e^{i\xi} + 1 + \frac{\bar{h}_0 - 2|h_0|^2}{\frac{1}{2} - h_0 + \bar{h}_0} e^{-i\xi} \right) \\ &= \frac{8|\frac{1}{2} - h_0 + \bar{h}_0|^2}{3 - 16|h_0 - \frac{1}{4}|^2} \left(1 + 2\operatorname{Re}\left(\frac{h_0 - 2|h_0|^2}{\frac{1}{2} - h_0 + \bar{h}_0} e^{i\xi}\right) \right) \\ &\geq \frac{8|\frac{1}{2} - h_0 + \bar{h}_0|^2}{3 - 16|h_0 - \frac{1}{4}|^2} \left(1 - 2 \left| \frac{h_0 - 2|h_0|^2}{\frac{1}{2} - h_0 + \bar{h}_0} \right| \right) \\ &> 0, \end{aligned}$$

which shows that $g(\xi)$ has a positive lower-bound.

All the discussion above shows that, for $N = 2$, $m_0(\xi) := h_0 + h_1 e^{-i\xi} + h_2 e^{-i2\xi}$ can generate a MRA based on Theorem 3.16 if and only if:

$$\begin{cases} h_0 \neq 0, h_0 \neq \frac{1}{2} \text{ and } |h_0 - \frac{1}{4}| < \frac{\sqrt{3}}{4} \\ h_1 = \frac{1}{2} \\ h_2 = \frac{1}{2} - h_0 \end{cases} \quad (3.24)$$

There are infinite bank of $\{h_0, h_1, h_2\}$ which satisfy the above conditions. Generally speaking, it is very difficult to find the corresponding scale function ϕ . Now, we consider a particular case.

For $h_0 = \frac{1}{4}$, we have

$$\begin{aligned} m_0(\xi) &= \frac{1}{4} + \frac{1}{2}e^{-i\xi} + \frac{1}{4}e^{-2i\xi} \\ &= \frac{1}{4}(1 + e^{-i\xi})^2 \\ &= \left(e^{-i\xi/2} \cos\left(\frac{1}{2}\xi\right)\right)^2. \end{aligned}$$

Obviously, it is just the square of the filter function m_0 for $N = 1$. Therefore,

$$\hat{\phi}(\xi) = (\hat{\chi}_{[0,1]}(\xi))^2 = (\chi_{[0,1]} * \chi_{[0,1]})(\xi),$$

i.e.,

$$\phi(x) = \chi_{[0,1]} * \chi_{[0,1]}(x) = \begin{cases} x & x \in [0, 1) \\ 2 - x & x \in [1, 2) \\ 0 & \text{otherwise} \end{cases}.$$

This is the well-known first order B-spline function.

3.3 The Construction of Biorthonormal Wavelet Bases

Multiresolution Analysis (MRA), since it was proposed by S. Mallat and Y. Meyer in late 1980's, has become the standard scheme for construction of wavelet bases. It has shown that almost all the wavelet bases with usual properties can be constructed from MRAs. In this section, we will study how to construct a general (orthonormal or not) wavelet basis $\{\psi_{j,k}\}$ from an MRA.

Definition 3.7 $\phi \in L^2(\mathbb{R})$ is said to satisfy the basic smooth condition if there exist positive numbers $\delta < 2$ and C such that

$$\begin{cases} \sum_{k \in \mathbb{Z}} |\hat{\phi}(\xi + 2\pi k)|^{2-\delta} \leq C & \text{a. e. } \xi \in \mathbb{T} \\ \sum_{j=0}^{\infty} |\hat{\phi}(2^j \xi)|^\delta \leq C & \text{a. e. } 1 \leq |\xi| \leq 2. \end{cases}$$

Note 1: The condition in the definition is called the basic smooth condition since the smoothness of ϕ corresponds to the decreasing of its Fourier transform.

Note 2: It can be proved that $\hat{\phi}(\xi) := \prod_{j=1}^{\infty} m_0(\xi/2^j)$ satisfies the basic smooth condition if m_0 satisfies either the following conditions:

- (1) the conditions of Theorem 3.8;
- (2) $m_0 \in \mathcal{P}_N^+$ ($N \neq 0$), $m_0(0) = 1$, $m_0(\pi) = 0$ and any eigenvalue λ of $T_{\mathcal{P}_N}$, the restriction of transition operator T on \mathcal{P}_N , satisfies $|\lambda| < 1$.

Theorem 3.17 Suppose $\tilde{\phi}, \phi \in L^2(\mathbb{R})$, generate biorthonormal MRA $\{\tilde{V}_j\}, \{V_j\}$, satisfy the basic smooth condition and the corresponding filter function $\tilde{m}_0, m_0 \in L^\infty(\mathbb{T})$ satisfy

$$|m_0(\xi)|^2 + |m_0(\xi + \pi)|^2 \neq 0,$$

and

$$\tilde{m}_0(\xi)m_0(\xi) + \tilde{m}_0(\xi + \pi)m_0(\xi + \pi) = 1, \quad a.e. \xi \in \mathbb{R}.$$

Then, for

$$m_1(\xi) = a(2\xi)\bar{\tilde{m}}_0(\xi + \pi)e^{-i\xi}, \quad \tilde{m}_1(\xi) = \overline{a^{-1}(2\xi)}\bar{m}_0(\xi + \pi)e^{-i\xi},$$

where $a(\xi)$ denotes a 2π -periodic function bounded by positive numbers from both the above and below, $\{\tilde{\psi}_{j,k}, \psi_{j,k}\}_{j,k \in \mathbb{Z}}$ which is defined by

$$\hat{\psi}(\xi) := \tilde{m}_1\left(\frac{\xi}{2}\right)\hat{\tilde{\phi}}\left(\frac{\xi}{2}\right), \quad \hat{\psi}(\xi) := m_1\left(\frac{\xi}{2}\right)\hat{\phi}\left(\frac{\xi}{2}\right) \quad (3.25)$$

constitutes a pair of biorthonormal wavelet bases of $L^2(\mathbb{R})$, that is, both $\{\tilde{\psi}_{j,k}\}_{j,k \in \mathbb{Z}}$ and $\{\psi_{j,k}\}_{j,k \in \mathbb{Z}}$ constitute a Riesz basis of $L^2(\mathbb{R})$ and they are biorthonormal each other, i.e.

$$\langle \tilde{\psi}_{j,k}, \psi_{j',k'} \rangle = \delta_{j,j'}\delta_{k,k'} \quad (\forall j, j' \in \mathbb{Z}, k, k' \in \mathbb{Z}).$$

Further, the following results hold:

(1). The operators defined by

$$\tilde{P}_j f := \sum_{k \in \mathbb{Z}} \langle f, \phi_{j,k} \rangle \tilde{\phi}_{j,k} \quad P_j f := \sum_{k \in \mathbb{Z}} \langle f, \tilde{\phi}_{j,k} \rangle \phi_{j,k} \quad (3.26)$$

$$\tilde{Q}_j f := \sum_{k \in \mathbb{Z}} \langle f, \psi_{j,k} \rangle \tilde{\psi}_{j,k} \quad Q_j f := \sum_{k \in \mathbb{Z}} \langle f, \tilde{\psi}_{j,k} \rangle \psi_{j,k} \quad (3.27)$$

satisfy that, $\forall j \in \mathbb{Z}$,

$$\begin{cases} \tilde{P}_{j+1} = \tilde{P}_j + \tilde{Q}_j, & P_{j+1} = P_j + Q_j, \\ \lim_{j \rightarrow -\infty} \|\tilde{P}_j f\|_2 = \lim_{j \rightarrow -\infty} \|P_j f\|_2 = 0 & \forall f \in L^2(\mathbb{R}), \\ \lim_{j \rightarrow +\infty} \|\tilde{P}_j f - f\|_2 = \lim_{j \rightarrow +\infty} \|P_j f - f\|_2 = 0 & \forall f \in L^2(\mathbb{R}). \end{cases}$$

(2). If we denote

$$\tilde{W}_j := \overline{\text{span}}\{\tilde{\psi}_{j,k} | k \in \mathbb{Z}\}, \quad W_j := \overline{\text{span}}\{\psi_{j,k} | k \in \mathbb{Z}\}, \quad (3.28)$$

then, for any $j \in \mathbb{Z}$, operators \tilde{P}_j , P_j , \tilde{Q}_j , Q_j are the projectors from $L^2(\mathbb{R})$ to \tilde{V}_j , V_j , \tilde{W}_j and W_j respectively satisfying:

$$\begin{cases} V_j \cap \tilde{V}_j^\perp = \tilde{V}_j \cap V_j^\perp = \{0\} \\ W_j \cap \tilde{W}_j^\perp = \tilde{W}_j \cap W_j^\perp = \{0\} \end{cases}$$

$$\begin{cases} \tilde{W}_j = \tilde{V}_{j+1} \cap V_j^\perp \\ W_j = V_{j+1} \cap \tilde{V}_j^\perp \end{cases} \quad \begin{cases} \tilde{V}_{j+1} = \tilde{V}_j \dot{+} \tilde{W}_j \\ V_{j+1} = V_j \dot{+} W_j \end{cases},$$

where $\dot{+}$ denotes the direct sum.

We omit the proof of the theorem because it is complicated and refers to some mathematical analysis which is not included in this book. The reader can obtain the details from [Daubechies, 1992; Long, 1995] and other related references.

Sometimes, $V_j \perp W_j$ is useful in practice. We will discuss this question here. A lemma is given first.

Lemma 3.3 Let m_0 be a 2π -periodic measurable function on \mathbb{R} satisfying $|m_0(\xi)|^2 + |m_0(\xi + \pi)|^2 \neq 0$ a.e. $\xi \in \mathbb{R}$. Then, a 2π -periodic measurable function m_1 satisfies

$$m_0(\xi)m_1(\xi) + m_0(\xi + \pi)m_1(\xi + \pi) = 0, \quad \text{a.e. } \xi \in \mathbb{R},$$

if and only if a 2π -periodic measurable function ν exists such that

$$m_1(\xi) = e^{-i\xi} \nu(2\xi) m_0(\xi + \pi), \quad \text{a.e. } \xi \in \mathbb{R}.$$

Proof. The part of the sufficiency holds obviously. We need to prove only the necessity, and we assume without losing generality that, $\forall \xi \in \mathbb{R}$,

$$\begin{cases} m_0(\xi) = m_0(\xi + 2\pi), & m_1(\xi) = m_1(\xi + 2\pi) \\ |m_0(\xi)|^2 + |m_0(\xi + \pi)|^2 \neq 0 \\ m_0(\xi)m_1(\xi) + m_0(\xi + \pi)m_1(\xi + \pi) = 0 \end{cases}.$$

It is easy to see that

$$\lambda(\xi) := \begin{cases} -\frac{m_1(\xi+\pi)}{m_0(\xi)} & \text{for } m_0(\xi) \neq 0 \\ \frac{m_1(\xi)}{m_0(\xi+\pi)} & \text{for } m_0(\xi) = 0. \end{cases}$$

is a 2π -periodic measurable function. We claim that

$$\lambda(\xi) + \lambda(\xi + \pi) = 0, \quad \forall \xi \in \mathbb{T}.$$

In fact,

(1). If $m_0(\xi) = 0$, it is clear that $m_0(\xi + \pi) \neq 0$, therefore,

$$\lambda(\xi) + \lambda(\xi + \pi) = \frac{m_1(\xi)}{m_0(\xi + \pi)} - \frac{m_1(\xi + 2\pi)}{m_0(\xi + \pi)} = 0;$$

(2). If $m_0(\xi + \pi) = 0$, it is still easy to deduce that $m_0(\xi) \neq 0$. By the result of (1), we have $\lambda(\xi + \pi) + \lambda(\xi + 2\pi) = 0$, i.e.

$$\lambda(\xi) + \lambda(\xi + \pi) = 0;$$

(3). At last, if $m_0(\xi) \neq 0$ and $m_0(\xi + \pi) \neq 0$, we can obtain

$$\begin{aligned} \lambda(\xi) + \lambda(\xi + \pi) &= -\frac{m_1(\xi + \pi)}{m_0(\xi)} - \frac{m_1(\xi)}{m_0(\xi + \pi)} \\ &= -\frac{m_1(\xi)m_0(\xi) + m_1(\xi + \pi)m_0(\xi + \pi)}{m_0(\xi)m_0(\xi + \pi)} \\ &= 0. \end{aligned}$$

Our claim is proved.

If $m_0(\xi) = 0$, we have

$$m_1(\xi) = \lambda(\xi)m_0(\xi + \pi).$$

If $m_0(\xi) \neq 0$, it is obvious that

$$m_1(\xi + \pi) = -\lambda(\xi)m_0(\xi),$$

which concludes that

$$\begin{aligned} m_1(\xi)m_0(\xi) &= -m_1(\xi + \pi)m_0(\xi + \pi) \\ &= \lambda(\xi)m_0(\xi)m_0(\xi + \pi). \end{aligned}$$

Hence, we also have

$$m_1(\xi) = \lambda(\xi)m_0(\xi + \pi).$$

Obviously, $\lambda(\xi) + \lambda(\xi + \pi) = 0$ implies that $\lambda(\xi)e^{-i\xi}$ is π -periodic. We denote

$$\nu(\xi) := \lambda\left(\frac{\xi}{2}\right)e^{-i\frac{1}{2}\xi},$$

then $\nu(\xi)$ is a 2π -periodic measurable function and $\lambda(\xi) = \nu(2\xi)e^{i\xi}$. Therefore,

$$m_1(\xi) = e^{i\xi}\nu(2\xi)m_0(\xi + \pi).$$

This ends the proof. \blacksquare

According to the above lemma, the conditions of $V_j \perp W_j$ can be characterized as follows:

Theorem 3.18 *Let m_0 and ϕ be the filter function and scale function of a MRA $\{V_j\}_{j \in \mathbb{Z}}$ respectively, $m_1 \in L^\infty(\mathbb{T})$. We denote*

$$\hat{\psi}(\xi) := m_1\left(\frac{\xi}{2}\right)\hat{\phi}\left(\frac{\xi}{2}\right),$$

$$W_j := \overline{\text{span}}\{\psi(2^j \cdot - k) | k \in \mathbb{Z}\}, \quad (\forall j \in \mathbb{Z}).$$

Then $V_j \perp W_j$ ($\forall j \in \mathbb{Z}$) if and only if there is a 2π -periodic measurable function $\nu(\xi)$ such that

$$m_1(\xi) = e^{-i\xi}\nu(2\xi)\bar{m}_0(\xi + \pi)\Phi(\xi + \pi),$$

where $\Phi := [\hat{\phi}, \hat{\phi}]$.

Proof. It is easy to see that

$$\begin{aligned} V_j \perp W_j \quad (\forall j \in \mathbb{Z}) &\iff V_0 \perp W_0 \\ &\iff \langle \psi(\cdot - k), \phi(\cdot) \rangle = 0 \quad (\forall k \in \mathbb{Z}) \\ &\iff [\hat{\psi}, \hat{\phi}](\xi) = 0 \quad a.e. \xi \in \mathbb{T}. \end{aligned}$$

Since

$$\begin{aligned} [\hat{\psi}, \hat{\phi}](\xi) &= \sum_{k \in \mathbb{Z}} \hat{\psi}(\xi + 2\pi k) \bar{\hat{\phi}}(\xi + 2\pi k) \\ &= \sum_{k \in \mathbb{Z}} m_1\left(\frac{\xi}{2} + \pi k\right) \hat{\phi}\left(\frac{\xi}{2} + \pi k\right) \bar{m}_0\left(\frac{\xi}{2} + \pi k\right) \bar{\hat{\phi}}\left(\frac{\xi}{2} + \pi k\right) \end{aligned}$$

$$\begin{aligned}
&= \sum_{k \in \mathbb{Z}} m_1\left(\frac{\xi}{2} + \pi k\right) \bar{m}_0\left(\frac{\xi}{2} + \pi k\right) |\hat{\phi}\left(\frac{\xi}{2} + \pi k\right)|^2 \\
&= \sum_{\alpha \in \mathbb{Z}} \left[m_1\left(\frac{\xi}{2}\right) \bar{m}_0\left(\frac{\xi}{2}\right) |\hat{\phi}\left(\frac{\xi}{2} + 2\pi\alpha\right)|^2 \right. \\
&\quad \left. + m_1\left(\frac{\xi}{2} + \pi\right) \bar{m}_0\left(\frac{\xi}{2} + \pi\right) |\hat{\phi}\left(\frac{\xi}{2} + 2\pi\alpha + \pi\right)|^2 \right] \\
&= m_1\left(\frac{\xi}{2}\right) \bar{m}_0\left(\frac{\xi}{2}\right) \Phi\left(\frac{\xi}{2}\right) + m_1\left(\frac{\xi}{2} + \pi\right) \bar{m}_0\left(\frac{\xi}{2} + \pi\right) \Phi\left(\frac{\xi}{2} + \pi\right),
\end{aligned}$$

we conclude that

$$\begin{aligned}
V_j \perp W_j \ (\forall j \in \mathbb{Z}) \iff & m_1(\xi) \bar{m}_0(\xi) \Phi(\xi) \\
& + m_1(\xi + \pi) \bar{m}_0(\xi + \pi) \Phi(\xi + \pi) \\
& = 0 \text{ a.e. } \xi \in \mathbb{T}.
\end{aligned}$$

Set m_0 in Lemma 3.3 to be $\bar{m}_0(\xi) \Phi(\xi)$ here and notice that ϕ generates an MRA, it can be concluded easily that there exist positive numbers A and B such that

$$A \leq |\bar{m}_0(\xi) \Phi(\xi)|^2 + |\bar{m}_0(\xi + \pi) \Phi(\xi + \pi)|^2 \leq B, \quad \text{a.e. } \xi \in \mathbb{R}.$$

By Lemma 3.3, our theorem is proved. \blacksquare

3.4 S. Mallat Algorithms

Under the conditions of Theorem 3.17, the well-known S. Mallat algorithm can be deduced easily. This algorithm provides a recursive scheme to calculate the coefficients of the biorthonormal wavelet expansion of a signal from one layer to the next layer. It can be applied to image processing, pattern recognition and other related subjects effectively.

Theorem 3.19 *Suppose the conditions of Theorem 3.17 are satisfied and denote \tilde{m}_0 , m_0 , \tilde{m}_1 , m_1 of Theorem 3.17 as follows:*

$$\begin{aligned}
\tilde{m}_0(\xi) &= \sum_{k \in \mathbb{Z}} \tilde{h}_k e^{-ik\xi}, & m_0(\xi) &= \sum_{k \in \mathbb{Z}} h_k e^{-ik\xi}, \\
\tilde{m}_1(\xi) &= \sum_{k \in \mathbb{Z}} \tilde{g}_k e^{-ik\xi}, & m_1(\xi) &= \sum_{k \in \mathbb{Z}} g_k e^{-ik\xi}.
\end{aligned}$$

Then, $\forall f \in L^2(\mathbb{R})$, the following S.Mallat decomposition algorithm holds:

$$\begin{cases} \langle f, \phi_{j,k} \rangle = \sqrt{2} \sum_{l \in \mathbb{Z}} h_l \langle f, \phi_{j+1,2k+l} \rangle \\ \langle f, \psi_{\mu,j,k} \rangle = \sqrt{2} \sum_{l \in \mathbb{Z}} g_l \langle f, \phi_{j+1,2k+l} \rangle \end{cases} \quad (\forall j \in \mathbb{Z}, k \in \mathbb{Z});$$

$$\begin{cases} \langle f, \tilde{\phi}_{j,k} \rangle = \sqrt{2} \sum_{l \in \mathbb{Z}} \tilde{h}_l \langle f, \tilde{\phi}_{j+1,2k+l} \rangle \\ \langle f, \tilde{\psi}_{j,k} \rangle = \sqrt{2} \sum_{l \in \mathbb{Z}} \tilde{g}_l \langle f, \tilde{\phi}_{j+1,2k+l} \rangle \end{cases} \quad (\forall j \in \mathbb{Z}, k \in \mathbb{Z}).$$

And the corresponding reconstruction algorithm holds as follows:

$$\begin{aligned} \langle f, \phi_{j+1,k} \rangle &= \sqrt{2} \sum_{l \in \mathbb{Z}} \tilde{h}_{k-2l} \langle f, \phi_{j,l} \rangle + \sqrt{2} \sum_{l \in \mathbb{Z}} \tilde{g}_{k-2l} \langle f, \psi_{j,l} \rangle, \\ &\quad (\forall j \in \mathbb{Z}, k \in \mathbb{Z}); \\ \langle f, \tilde{\phi}_{j+1,k} \rangle &= \sqrt{2} \sum_{l \in \mathbb{Z}} \bar{h}_{k-2l} \langle f, \tilde{\phi}_{j,l} \rangle + \sqrt{2} \sum_{l \in \mathbb{Z}} \bar{g}_{k-2l} \langle f, \tilde{\psi}_{j,l} \rangle, \\ &\quad (\forall j \in \mathbb{Z}, k \in \mathbb{Z}). \end{aligned}$$

Proof. Using the two-scale relation and the expression of m_0 , we have

$$\phi(x) = 2 \sum_{l \in \mathbb{Z}} h_l \phi(2x - l),$$

which is equivalent to

$$\phi_{j,k}(x) = \sum_{l \in \mathbb{Z}} h_l \phi_{j+1,2k+l}(x).$$

Hence

$$\langle f, \phi_{j,k} \rangle = \sum_{l \in \mathbb{Z}} h_l \langle f, \phi_{j+1,2k+l} \rangle.$$

The other formulae of the decomposition algorithm can be proved similarly. Now we turn to the proof of the reconstruction algorithm.

By Theorem 3.17, we deduce that, $\forall j \in \mathbb{Z}, l \in \mathbb{Z}$,

$$\tilde{P}_{j+1} \tilde{\phi}_{j+1,l} = \tilde{P}_j \tilde{\phi}_{j+1,l} + \tilde{Q}_j \tilde{\phi}_{j+1,l},$$

i.e.

$$\tilde{\phi}_{j+1,l} = \sum_{k \in \mathbb{Z}} \langle \tilde{\phi}_{j+1,l}, \phi_{j,k} \rangle \tilde{\phi}_{j,k} + \sum_{k \in \mathbb{Z}} \langle \tilde{\phi}_{j+1,l}, \psi_{j,k} \rangle \tilde{\psi}_{j,k}$$

$$\begin{aligned}
&= \sum_{k \in \mathbb{Z}} \langle \tilde{\phi}_{j+1,l}, \sum_{m \in \mathbb{Z}} h_m \phi_{j+1,2k+m} \rangle \tilde{\phi}_{j,k} \\
&\quad + \sum_{k \in \mathbb{Z}} \langle \tilde{\phi}_{j+1,l}, \sum_{m \in \mathbb{Z}} g_{\mu,m} \phi_{j+1,2k+m} \rangle \tilde{\psi}_{j,k} \\
&= \sum_{k \in \mathbb{Z}} \bar{h}_{l-2k} \tilde{\phi}_{j,k} + \sum_{k \in \mathbb{Z}} \bar{g}_{l-2k} \tilde{\psi}_{j,k}.
\end{aligned}$$

Therefore, we have

$$\langle f, \tilde{\phi}_{j+1,l} \rangle = \sum_{k \in \mathbb{Z}} \bar{h}_{l-2k} \langle f, \tilde{\phi}_{j,k} \rangle + \sum_{k \in \mathbb{Z}} \bar{g}_{l-2k} \langle f, \tilde{\psi}_{j,k} \rangle.$$

The another formula of the reconstruction algorithm can be shown similarly.
The proof of this theorem is complete. ■

Note: S.Mallat decomposition algorithm and reconstruction algorithm can be illustrated in Table 3.1.

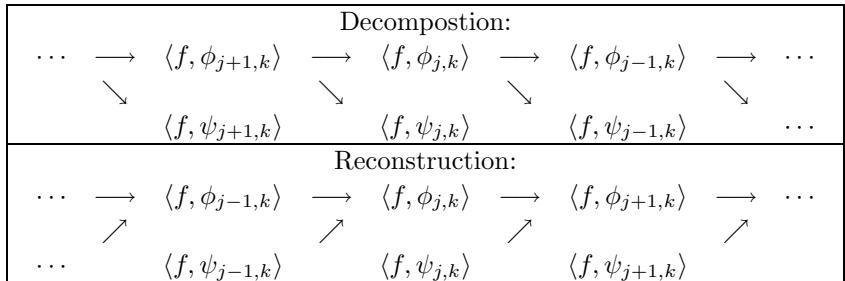


Table 3.1 S. Mallat algorithm

S. Mallat algorithm can be applied effectively to image processing, pattern recognition, fast computation of singular integrals and some other areas. It decomposes a signal from the lower frequency to the higher frequency locally both in time and frequency domains. In practical applications, we usually need the discrete form of S. Mallat algorithm. Let

$$\begin{cases} s_k^j := 2^{j/2} \langle f, \phi_{j,k} \rangle, \\ \tilde{s}_k^j := 2^{j/2} \langle f, \phi_{j,k} \rangle \end{cases}, \quad \begin{cases} t_k^j := 2^{j/2} \langle f, \psi_{j,k} \rangle \\ \tilde{t}_k^j := 2^{j/2} \langle f, \psi_{j,k} \rangle \end{cases}. \quad (3.29)$$

Then, its discrete form can be written as follows: $\forall j \in \mathbb{Z}, k \in \mathbb{Z}$,

Decomposition Algorithm:

$$\begin{cases} s_k^j = \sum_{l \in \mathbb{Z}} h_l s_{2k+l}^{j+1} \\ t_k^j = \sum_{l \in \mathbb{Z}} g_l s_{2k+l}^{j+1} \end{cases} \quad \begin{cases} \tilde{s}_k^j = \sum_{l \in \mathbb{Z}} \tilde{h}_l \tilde{s}_{2k+l}^{j+1} \\ \tilde{t}_k^j = \sum_{l \in \mathbb{Z}} \tilde{g}_l \tilde{s}_{2k+l}^{j+1} \end{cases} \quad (3.30)$$

Reconstruction Algorithm:

$$\begin{cases} s_k^{j+1} = 2^d \sum_{l \in \mathbb{Z}} \bar{h}_{k-2l} s_l^j + 2^d \sum_{l \in \mathbb{Z}} \bar{g}_{k-2l} t_l^j \\ \tilde{s}_k^{j+1} = 2^d \sum_{l \in \mathbb{Z}} \bar{\tilde{h}}_{k-2l} \tilde{s}_l^j + 2^d \sum_{l \in \mathbb{Z}} \bar{\tilde{g}}_{k-2l} \tilde{t}_l^j \end{cases} \quad (3.31)$$

The discrete S.Mallat algorithm is depicted in Table 3.2.

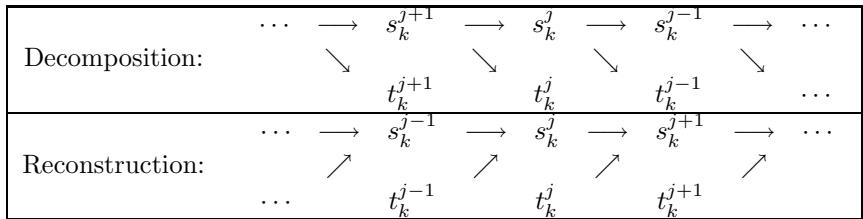


Table 3.2 Discrete form of S. Mallat algorithm

The first formula of the decomposition algorithm (3.30) can be explained as follows: An input signal $\{s_k^{j+1}\}_{k \in \mathbb{Z}}$ is first filtered by a filter $\{h_{-k}\}_{k \in \mathbb{Z}}$, which corresponds to a convolution in mathematics, then the output signal $\{\sum_{l \in \mathbb{Z}} h_l s_{k+l}^{j+1}\}$ is sampled alternately (that is, only the points with even indices are kept), which is called “downsample” and denoted by “ $2 \downarrow$ ”. At last, we get $\{s_k^j\}_{k \in \mathbb{Z}}$. The procedure is represented below:

$$\{s_k^{j+1}\}_{k \in \mathbb{Z}} \longrightarrow \boxed{\{h_{-k}\}_{k \in \mathbb{Z}}} \longrightarrow \boxed{2 \downarrow} \longrightarrow \{s_k^j\}_{k \in \mathbb{Z}}.$$

The other formulae of (3.30) can be illustrated similarly.

The implementation of the first part of the first formula of the reconstruction algorithm (3.31), $x_k^{j+1} := \sum_{l \in \mathbb{Z}} \bar{h}_{k-2l} s_l^j$, is just in inverse order. At first, the input signal is upsampled. More precisely, a zero is placed in between every two consecutive terms of the input sequence $\{s_k^j\}_{k \in \mathbb{Z}}$, which is denoted by “ $2 \uparrow$ ”. Then the upsampled sequence is filtered by $\{\bar{h}_k\}_{k \in \mathbb{Z}}$ and at last the output signal $\{x_k^{j+1}\}_{k \in \mathbb{Z}}$ is obtained. The following is an illustration of the procedure.

$$\{s_k^j\}_{k \in \mathbb{Z}} \longrightarrow \boxed{2 \uparrow} \longrightarrow \boxed{\{\bar{\tilde{h}}_k\}_{k \in \mathbb{Z}}} \longrightarrow \{x_k^{j+1}\}_{k \in \mathbb{Z}}$$

The other parts of (3.31) can be explained similarly.

Chapter 4

Some Typical Wavelet Bases

The purpose of this chapter is to give readers more complete information about the wavelet bases commonly used in the areas of signal processing, image processing and pattern recognition. Some important properties of these wavelet bases are presented. To facilitate the understanding of these properties and further application of these wavelet bases for engineers and scientists, the properties are described in detail while the construction of their mathematical formula is avoided as much as possible. The reason is that in pattern recognition, wavelet transform is only a tool or method which is similar to the Fourier transform in signal analysis.

In practice, we usually apply real wavelet $\psi(t)$, which is a real-valued function. For such a wavelet, it is easy to show that its Fourier transform $\hat{\psi}(\omega)$ satisfies

$$\hat{\psi}(-\omega) = \overline{\psi(\omega)},$$

and $\hat{\psi}(\omega)$ is also a real-valued function if and only if $\psi(t)$ is an even function.

In fact, for real-valued function $\psi(t)$, there holds

$$\begin{aligned}\hat{\psi}(-\omega) &= \int_{\mathbb{R}} \psi(t) e^{it\omega} dt \\ &= \int_{\mathbb{R}} \overline{\psi(t)} e^{-it\omega} dt \\ &= \overline{\int_{\mathbb{R}} \psi(t) e^{-it\omega} dt} \\ &= \overline{\psi(\omega)}.\end{aligned}$$

And since

$$\begin{aligned}\overline{\hat{\psi}(\omega)} &= \overline{\int_{\mathbb{R}} \psi(t) e^{-it\omega} dt} \\ &= \int_{\mathbb{R}} \psi(t) e^{it\omega} dt \\ &= \int_{\mathbb{R}} \psi(-t) e^{-it\omega} dt \\ &= \hat{\psi}(-t)(\omega),\end{aligned}$$

$\hat{\psi}(\omega)$ is real-valued function if and only if

$$\overline{\hat{\psi}(\omega)} = \hat{\psi}(\omega),$$

i.e.,

$$(\psi(-t))(\omega) = (\psi(t))(\omega),$$

i.e.,

$$\psi(-t) = \psi(t),$$

which means $\psi(t)$ is an even function.

In general, wavelet $\psi(t)$ is not an even function, which implies that $\hat{\psi}(\omega)$ is an imaginary function. Therefore we usually indicate the graph of $|\hat{\psi}(\omega)|$ instead. Now that $\hat{\psi}(-\omega) = \overline{\hat{\psi}(\omega)}$, we have that

$$|\hat{\psi}(-\omega)| = |\hat{\psi}(\omega)|,$$

that is, $|\hat{\psi}(\omega)|$ is an even function. Hence we only need to indicate the graph of $|\hat{\psi}(\omega)|$ on positive semi-axis $[0, \infty)$.

Two important construction of orthonormal wavelet bases from MRAs are listed below:

- **Construction of orthonormal wavelet basis from an orthonormal MRA**

Theorem 4.1 *Let ϕ be the scaling function of an orthonormal MRA with filter function m_0 . Then ψ , which is defined by*

$$\hat{\psi}(\xi) := e^{-i\frac{\xi}{2}} \bar{m}_0\left(\frac{\xi}{2} + \pi\right) \hat{\phi}\left(\frac{\xi}{2} + \pi\right),$$

generates an orthonormal wavelet basis of $L^2(\mathbb{R})$.

- Construction of orthonormal wavelet basis from nonorthonormal MRA

Theorem 4.2 Let $m_0 \in C(\mathbb{T})$ satisfy the following conditions:

- (1). $\prod_{j=1}^{\infty} m_0(\xi/2^j)$ converges for almost every $\xi \in \mathbb{R}$;
- (2). $\Phi(\xi) \in C(\mathbb{T})$ and there exist positive constants A and B such that $A \leq \Phi(\xi) \leq B$ ($\forall \xi \in \mathbb{R}$), where

$$\Phi(\xi) := \sum_{k \in \mathbb{Z}} |\hat{\phi}(\xi + 2k\pi)|^2$$

and

$$\hat{\phi}(\xi) := \prod_{j=1}^{\infty} m_0(\xi/2^j).$$

Then,

$$\hat{\phi}^{\sharp}(\xi) := \frac{\hat{\phi}(\xi)}{\sqrt{\Phi(\xi)}}$$

is in $L^2(\mathbb{R})$ and generates an orthonormal MRA of $L^2(\mathbb{R})$. The corresponding scaling function and filter function are respectively $\phi^{\sharp}(\xi)$ and $m_0^{\sharp}(\xi)$ defined by

$$m_0^{\sharp}(\xi) := m_0(\xi) \sqrt{\frac{\Phi(\xi)}{\Phi(2\xi)}}.$$

Proof. It is obvious that $m_0^{\sharp} \in C(\mathbb{T})$ and we can prove that $\Phi(0) = 1$ by Condition (2). It is also easy to see that

$$\prod_{j=1}^{\infty} m_0^{\sharp}(\xi/2^j) = \sqrt{\prod_{j=1}^{\infty} \frac{\Phi(\xi/2^j)}{\Phi(\xi/2^{j-1})}} \prod_{j=1}^{\infty} m_0(\xi/2^j) = \frac{\hat{\phi}(\xi)}{\sqrt{\Phi(\xi)}}.$$

Hence ϕ^{\sharp} , which is defined by

$$\hat{\phi}^{\sharp}(\xi) := \frac{\hat{\phi}(\xi)}{\sqrt{\Phi(\xi)}} = \prod_{j=1}^{\infty} m_0^{\sharp}(\xi/2^j),$$

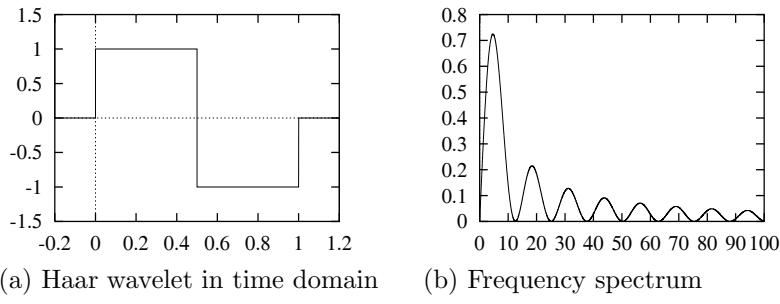


Fig. 4.1 Haar wavelet and its frequency spectrum.

is refineable and the filter function is m_0^\sharp . Since

$$\sum_{k \in \mathbb{Z}} |\hat{\phi}^\sharp(\xi + 2k\pi)|^2 = \frac{\sum_{k \in \mathbb{Z}} |\hat{\phi}(\xi + 2k\pi)|^2}{\Phi(\xi)} = 1, \quad (\forall \xi \in \mathbb{R}),$$

ϕ^\sharp generates an orthonormal MRA of $L^2(\mathbb{R})$. The corresponding scaling function and filter function are $\phi^\sharp(\xi)$ and $m_0^\sharp(\xi)$ respectively. The proof is complete. ■

4.1 Orthonormal Wavelet Bases

4.1.1 Haar Wavelet

The simplest well-known wavelet is the Haar wavelet, which was presented in 1910 by Haar and defined by

$$\psi(t) = \begin{cases} 1 & 0 \leq t < \frac{1}{2}; \\ -1 & \frac{1}{2} \leq t < 1; \\ 0 & \text{otherwise.} \end{cases}$$

Its Fourier transform is

$$\hat{\psi}(\xi) = \frac{1}{4} i \xi e^{-i\xi/2} \left(\frac{4}{\xi} \sin \frac{\xi}{4} \right)^2 \quad (4.1)$$

The Haar wavelet has bad decay in the frequency domain, i.e., its localization in the frequency domain is not good. But in the time domain it is compactly supported on $[0, 1]$. It is also orthonormal and anti-symmetric. Fig. 4.1 shows the Haar wavelet and its frequency spectrum.

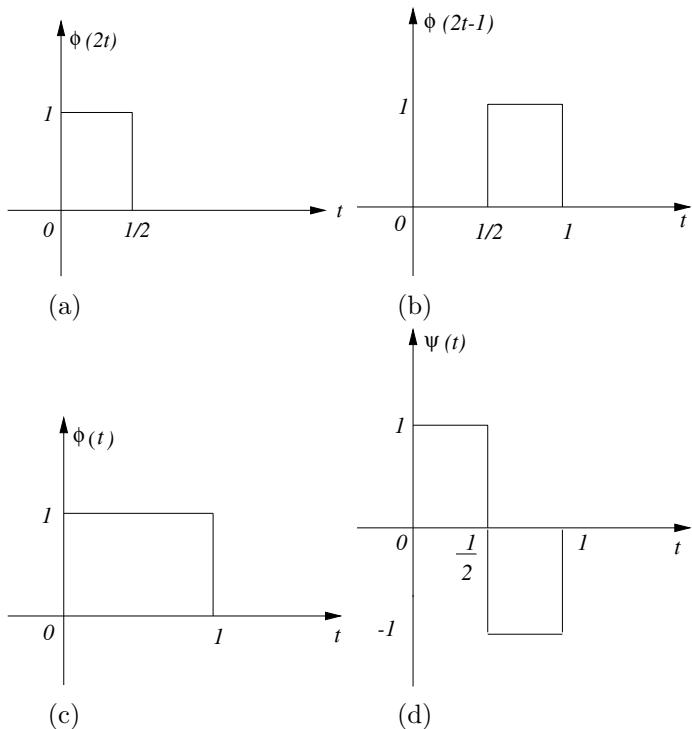


Fig. 4.2 Haar scale function and Haar wavelet: (a) $\phi(2t)$; (b) $\phi(2t - 1)$; (c) $\phi(t)$; (d) $\psi(t)$.

The corresponding scale function of Haar wavelet is as follows:

$$\phi(t) = \begin{cases} 1 & 0 \leq t \leq 1, \\ 0 & \text{otherwise.} \end{cases} \quad (4.2)$$

It is easy to see that $\phi(t)$ satisfies the following two-scale equation:

$$\phi(t) = \phi(2t) + \phi(2t - 1) \quad (4.3)$$

and ψ can be generated by ϕ according to the following equation:

$$\psi(t) = \phi(2t) - \phi(2t - 1). \quad (4.4)$$

Equations (4.2)-(4.4) are graphically illustrated in Fig. 4.2.

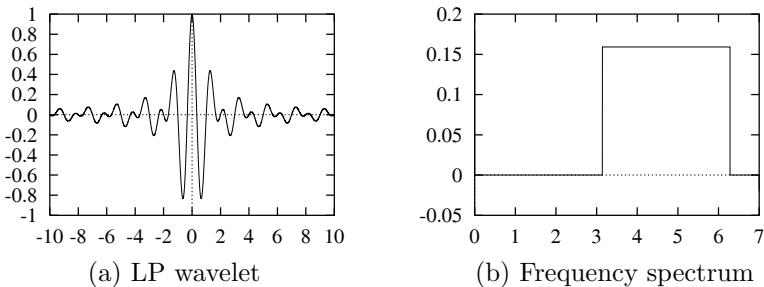


Fig. 4.3 LP wavelet and its frequency spectrum.

4.1.2 Littlewood-Paley (LP) Wavelet

Corresponding to the Haar wavelet, an orthonormal wavelet that is compactly supported in the frequency domain was given by Littlewood and Paley. Its analytical expression in the frequency domain is

$$\hat{\psi}(\xi) = \begin{cases} (2\pi)^{-\frac{1}{2}} & \pi \leq |\xi| \leq 2\pi, \\ 0 & \text{otherwise.} \end{cases} \quad (4.5)$$

Its inverse Fourier transform is

$$\psi(t) = (\pi t)^{-1}(\sin 2\pi t - \sin \pi t) \quad (4.6)$$

Fig. 4.3 shows the LP wavelet and its frequency spectrum. LP wavelet $\psi(t) \in C^\infty$ has order one vanishing moment and is not compactly supported. The slow decay of $\psi(t)$ makes it has not localization in time domain ($\psi(t) \sim |x|^{-1}$ for $t \rightarrow \infty$) but excellent localization in frequency domain since its Fourier transform is compactly supported. Table 4.1 shows the

Wavelet	Time Window		Freq. Window		Time-freq. Area
	Center	Radius	Center	Radius	
LP	0	∞	$\frac{\pi}{2}\sqrt{3}$	12.029	∞

Table 4.1 The centers, radii and area of time window, frequency window and time-frequency window of the LP wavelet

centers and radii of time and frequency windows and the area of the time-frequency window of the LP wavelet.

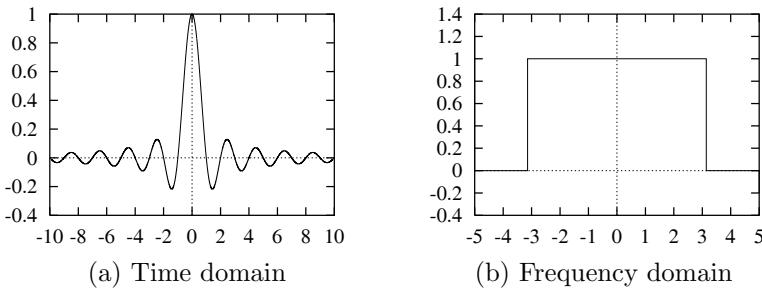


Fig. 4.4 Shannon sampling function and its frequency spectrum.

The corresponding scale function of Littlewood-Paley wavelet is

$$\phi(t) = \frac{\sin \pi t}{\pi t}$$

whose Fourier transform is as follows:

$$\hat{\phi}(\xi) = \begin{cases} 1 & -\pi \leq \xi < \pi, \\ 0 & \text{otherwise.} \end{cases}$$

The two-scale equation $\phi(t)$ satisfies:

$$\hat{\phi}(\xi) = m_0\left(\frac{\xi}{2}\right)\hat{\phi}\left(\frac{\xi}{2}\right),$$

where $m_0(\xi)$ is a 2π -periodic function whose definition on $[-\frac{\pi}{2}, \frac{3\pi}{2})$ is as follows:

$$m_0(\xi) = \begin{cases} 1 & \xi \in [-\frac{\pi}{2}, \frac{\pi}{2}) \\ 0 & \xi \in [\frac{\pi}{2}, \frac{3\pi}{2}) \end{cases}.$$

Since

$$|m_0(\xi)|^2 + |m_0(\xi + \pi)|^2 = 1, \quad \forall \xi \in \mathbb{T},$$

ϕ generates an orthonormal MRA, and consequently ψ , which is defined by

$$\hat{\psi}(\xi) = e^{-i\frac{\xi}{2}} m_0\left(\frac{\xi}{2} + \pi\right) \hat{\phi}\left(\frac{\xi}{2}\right),$$

generates an orthonormal wavelet basis.

It is easy to see that $\phi(t)$ happens to be the Shannon sampling function. Fig. 4.4 shows it and its frequency spectrum.

4.1.3 Meyer Wavelet

Meyer wavelet was constructed by Y. Meyer in 1985. The fourier transform of the corresponding scale function is defined by

$$\hat{\phi}(\xi) := \begin{cases} 1 & |\xi| \leq \frac{2}{3}\pi \\ \cos[\frac{\pi}{2}\nu(\frac{3}{2\pi}|\xi| - 1)] & \frac{2}{3}\pi \leq |\xi| \leq \frac{4}{3}\pi \\ 0 & \text{otherwise,} \end{cases}$$

where $\nu(\xi)$ is an arbitrary C^∞ function satisfying

$$\nu(x) = \begin{cases} 0 & x \leq 0 \\ 1 & x \geq 1 \end{cases} \quad \text{and} \quad \nu(x) + \nu(1-x) = 1 \quad (\forall x \in [0, 1]). \quad (4.7)$$

ϕ generates an orthonormal MRA since we can deduce that

$$\Phi(\xi) = \sum_{k=-\infty}^{\infty} |\hat{\phi}(\xi + 2k\pi)|^2 = 1, \quad (\forall \xi \in \mathbb{T}).$$

It can be proven that the filter function is

$$m_0(\xi) = \sum_{k=-\infty}^{\infty} \hat{\phi}(2(\xi + 2\pi k)),$$

and Meyer wavelet ψ , which is an orthonormal wavelet, is then defined by

$$\hat{\psi}(\xi) = e^{-i\frac{\xi}{2}} \bar{m}_0\left(\frac{\xi}{2} + \pi\right) \hat{\phi}\left(\frac{\xi}{2}\right).$$

It can be shown that

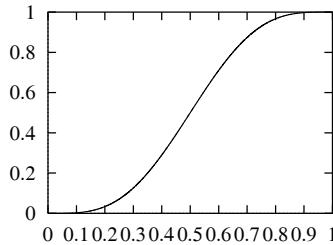
$$\hat{\psi}(\xi) = \begin{cases} e^{-i\frac{\xi}{2}} \sin[\frac{\pi}{2}\nu(\frac{3}{2\pi}|\xi| - 1)] & \frac{2\pi}{3} \leq |\xi| \leq \frac{4\pi}{3} \\ e^{-i\frac{\xi}{2}} \cos[\frac{\pi}{2}\nu(\frac{3}{4\pi}|\xi| - 1)] & \frac{4\pi}{3} \leq |\xi| \leq \frac{8\pi}{3} \\ 0 & \text{otherwise.} \end{cases} \quad (4.8)$$

Obviously, $\hat{\psi}(\xi)$ is compactly supported and its regularity is the same as that of ν .

Particularly, if we take $\nu(t) = 0$ for $t \in (-\infty, 0) \cup (1, \infty)$ and

$$\nu(t) = t^4(35 - 84t + 70t^2 - 20t^3), \quad t \in [0, 1], \quad (4.9)$$

it is easy to see that $\nu(t)$, which is plotted in Fig. 4.5, satisfies (4.7). The corresponding Meyer wavelet, which is the inverse Fourier transform of (4.8), and its frequency property are shown in Fig. 4.6 for ν defined by (4.9). The regularity of $\psi(t)$ is C^∞ . The decay of $\psi(t)$ is faster than any

Fig. 4.5 Function $\nu(t)$ as given by (4.7).

inverse polynomial, i.e., for any natural numbers N , there exists $C_N < \infty$ such that

$$|\psi(t)| \leq C_N(1 + |t|^2)^{-N}.$$

But it cannot decay exponentially fast as $t \rightarrow \infty$. Moreover, Meyer wavelet has linear phase since it is symmetric.

The centers, radii and area of the time-window, frequency-window and time-frequency window of this Meyer wavelet are listed in Table 4.2. The order of its vanishing moment is 0.

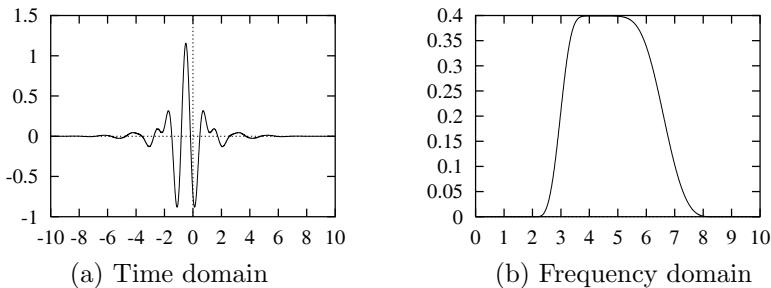


Fig. 4.6 Meyer wavelet and its frequency spectrum.

Wavelet	Time Window		Freq. Window		Time-freq. Area
	Center	Radius	Center	Radius	
Meyer	0.5	0.6774	2.369	1.81349	4.9140

Table 4.2 The centers, radii and area of time window, frequency window and time-frequency window of the Meyer wavelet

4.1.4 Battle-Lemaré-spline Wavelet

The Battle-Lemaré spline wavelet is based on central B-splines, which are polynomial spline functions with equally distance simple knots. For $m \in \mathbb{N}$, the scaling function ϕ is defined by

$$\phi(x) := \begin{cases} \underbrace{\chi_{[-\frac{1}{2}, \frac{1}{2}]} * \cdots * \chi_{[-\frac{1}{2}, \frac{1}{2}]}}_m(x), & \text{if } m \text{ is even;} \\ \underbrace{\chi_{[-\frac{1}{2}, \frac{1}{2}]} * \cdots * \chi_{[-\frac{1}{2}, \frac{1}{2}]}}_m(x - \frac{1}{2}), & \text{if } m \text{ is odd,} \end{cases} \quad (4.10)$$

where $\chi_{[-\frac{1}{2}, \frac{1}{2}]}(x)$ is the characteristic function of $[-\frac{1}{2}, \frac{1}{2}]$ defined by

$$\chi_{[-\frac{1}{2}, \frac{1}{2}]}(x) = \begin{cases} 1 & x \in [-\frac{1}{2}, \frac{1}{2}] \\ 0 & \text{otherwise} \end{cases}$$

and '*' is the convolution operation. It is easy to see that

$$\phi(x) := \begin{cases} N_m(x + \frac{1}{2}m) & \text{if } m \text{ is even;} \\ N_m(x + \frac{1}{2}(m-1)) & \text{if } m \text{ is odd,} \end{cases}$$

where $N_m(x)$ is the spline function of order m defined by (2.15). Thus, $\phi(x)$ is a spline function of order m with integer knots. It is a polynomial on interval $[n, n+1]$ for any $n \in \mathbb{N}$ and $\phi \in C^{m-2}(\mathbb{R})$. The support of ϕ is $[-\frac{m}{2}, \frac{m}{2}]$ if m is even and $[-\frac{m}{2} + \frac{1}{2}, \frac{m}{2} + \frac{1}{2}]$ if m is odd. A simple calculation yields that

$$\hat{\phi}(\xi) := \begin{cases} (\frac{2}{\xi} \sin \frac{\xi}{2})^m & \text{if } m \text{ is even;} \\ e^{-i\frac{\xi}{2}} (\frac{2}{\xi} \sin \frac{\xi}{2})^m & \text{if } m \text{ is odd.} \end{cases}$$

Hence the filter function is

$$m_0(\xi) = \begin{cases} (\cos \frac{\xi}{2})^m & \text{if } m \text{ is even;} \\ e^{-i\frac{\xi}{2}} (\cos \frac{\xi}{2})^m & \text{if } m \text{ is odd.} \end{cases}$$

Since

$$\left(\cos \frac{\xi}{2}\right)^m = \left(\frac{e^{i\frac{\xi}{2}} + e^{-i\frac{\xi}{2}}}{2}\right)^m = e^{i\frac{m}{2}\xi} \left(\frac{1 + e^{-i\xi}}{2}\right)^m,$$

we get

$$\begin{aligned} m_0(\xi) &= \begin{cases} e^{i\frac{m}{2}\xi} \left(\frac{1+e^{-i\xi}}{2} \right)^m & \text{if } m \text{ is even;} \\ e^{i\frac{m-1}{2}\xi} \left(\frac{1+e^{-i\xi}}{2} \right)^m & \text{if } m \text{ is odd.} \end{cases} \\ &= e^{i[\frac{m}{2}]\xi} \left(\frac{1+e^{-i\xi}}{2} \right)^m \end{aligned}$$

where $[\frac{m}{2}]$ is the largest integer not larger than $\frac{m}{2}$.

It is easy to see that

$$|m_0(\xi)|^2 + |m_0(\xi + \pi)|^2 = |\cos \frac{\xi}{2}|^{2m} + |\sin \frac{\xi}{2}|^{2m}.$$

Since

$$|\cos \frac{\xi}{2}|^{2m} + |\sin \frac{\xi}{2}|^{2m} \begin{cases} = |\cos \frac{\xi}{2}|^2 + |\sin \frac{\xi}{2}|^2 = 1 & \text{if } m = 1; \\ < |\cos \frac{\xi}{2}|^2 + |\sin \frac{\xi}{2}|^2 = 1 & \text{if } m > 1, \end{cases}$$

we have that $\forall \xi \in [-\pi, \pi]$,

$$\begin{cases} |m_0(\xi)|^2 + |m_0(\xi + \pi)|^2 = 1 & \text{if } m = 1; \\ 0 < |m_0(\xi)|^2 + |m_0(\xi + \pi)|^2 < 1 & \text{if } m > 1. \end{cases}$$

Therefore, if $m = 1$, ϕ , which is defined by (4.10), i.e.,

$$\begin{aligned} \phi(x) &= \chi_{[-\frac{1}{2}, \frac{1}{2}]}(x - \frac{1}{2}) \\ &= \begin{cases} 1 & x \in [0, 1] \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

generates an orthonormal MRA. Let

$$\psi(t) = \phi(2t) - \phi(2t - 1),$$

we get the well-known Haar wavelet basis $\{\psi_{j,k}(x)\}$, which was discussed in (4.1.1).

And if $m > 1$, ϕ , which is defined by (4.10), can not generate an orthonormal MRA. However, $\forall \xi \in [-\pi, \pi]$, it is easily to see that

$$\begin{aligned} \Phi(\xi) &:= \sum_{k=-\infty}^{\infty} |\hat{\phi}(\xi + 2k\pi)|^2 \\ &= \left| 2 \sin \frac{\xi}{2} \right|^{2m} \sum_{k=-\infty}^{\infty} \frac{1}{|\xi + 2k\pi|^{2m}} \end{aligned}$$

$$= \left| 2 \sin \frac{\xi}{2} \right|^{2m} \left[\frac{1}{|\xi|^{2m}} + \sum_{k=1}^{\infty} \frac{1}{|\xi + 2k\pi|^{2m}} + \sum_{k=1}^{\infty} \frac{1}{|\xi - 2k\pi|^{2m}} \right],$$

and

$$\begin{cases} \sum_{k=1}^{\infty} \frac{1}{|\xi + 2k\pi|^{2m}}, \quad \sum_{k=1}^{\infty} \frac{1}{|\xi - 2k\pi|^{2m}} \leq \left(\frac{1}{\pi}\right)^{2m} \sum_{k=1}^{\infty} \frac{1}{(2k-1)^{2m}}, \\ \left(\frac{2}{\pi}\right)^{2m} \leq \frac{1}{|\xi|^{2m}} \left| 2 \sin \frac{\xi}{2} \right|^{2m} \leq 1, \\ \left| 2 \sin \frac{\xi}{2} \right|^{2m} \leq 2^{2m}, \end{cases}$$

therefore,

$$\left(\frac{2}{\pi}\right)^{2m} \leq \Phi(\xi) \leq 1 + 2^{2m+1} \left(\frac{1}{\pi}\right)^{2m} \sum_{k=1}^{\infty} \frac{1}{(2k-1)^{2m}}.$$

This inequality implies that $\phi(x)$ can generate a nonorthonormal MRA, i.e., $\{\phi(x-k)|k \in \mathbb{Z}\}$ is a Riesz basis of

$$V_0 := \overline{\text{span}\{\phi(x-k)|k \in \mathbb{Z}\}},$$

and $V_j := \{f(2^j x)|f \in V_0\}$ ($j \in \mathbb{Z}$) constitutes a MRA.

To construct an orthonormal MRA, let

$$\hat{\phi}^\sharp(\xi) := \frac{\hat{\phi}(\xi)}{\sqrt{\Phi(\xi)}}$$

and

$$m_0^\sharp(\xi) := m_0(\xi) \sqrt{\frac{\Phi(\xi)}{\Phi(2\xi)}}.$$

Then, by Theorem 4.2, ϕ^\sharp generates an orthonormal MRA of $L^2(\mathbb{R})$. The corresponding scaling function and filter function are $\phi^\sharp(\xi)$ and $m_0^\sharp(\xi)$ respectively.

Let

$$\hat{\psi}^\sharp(\xi) := e^{-i\frac{\xi}{2} \overline{m_0^\sharp}} \left(\frac{\xi}{2} + \pi\right) \hat{\phi}^\sharp\left(\frac{\xi}{2}\right).$$

Then, by Theorem 4.1, ψ^\sharp generates an orthonormal wavelet basis of $L^2(\mathbb{R})$, which is the well-known Battle-Lemaré-spline wavelet basis or B-spline wavelet basis of order $m-1$.

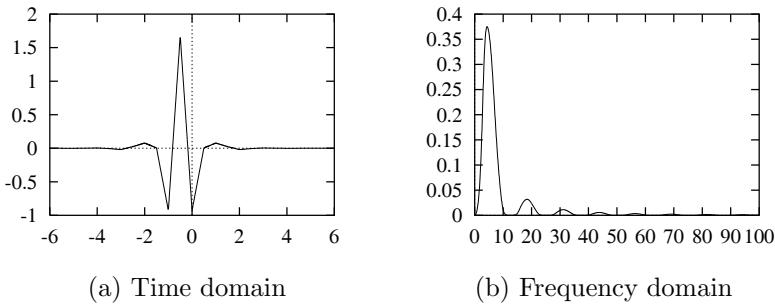


Fig. 4.7 B-spline wavelet of order 1 and its frequency spectrum.

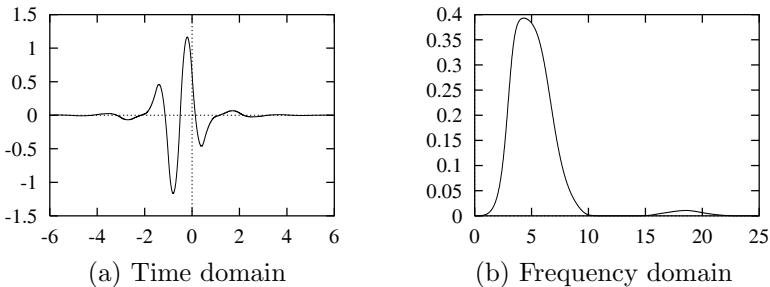


Fig. 4.8 B-spline wavelet of order 2 and its frequency spectrum.

B-spline wavelet has no analytical expression. Fig. 4.7-4.10 show B-spline wavelets of orders $1 \sim 4$ and their frequency properties, respectively.

The centers, radii and areas of the time windows, frequency windows and time-frequency windows of the B-splines wavelets of orders $1 \sim 10$ are shown in Table 4.3.

According to these figures and the table, we easily see that B-spline wavelets are symmetric, with $-\frac{1}{2}$ as the symmetric center. We also observe that the B-spline wavelet of order 2 has the smallest time-frequency localization property. Also, B-spline wavelets are orthonormal and decay exponentially. Table 4.3 shows the relationship between the order number of B-spline wavelets and their time-frequency localization properties. The time-frequency localization property becomes bad with increasing order number except the quadratic spline wavelet.

It can be proven that Battle-Lemaré-spline wavelet decreases exponentially at infinity. Reader are referred to [Daubechies, 1992] for details.

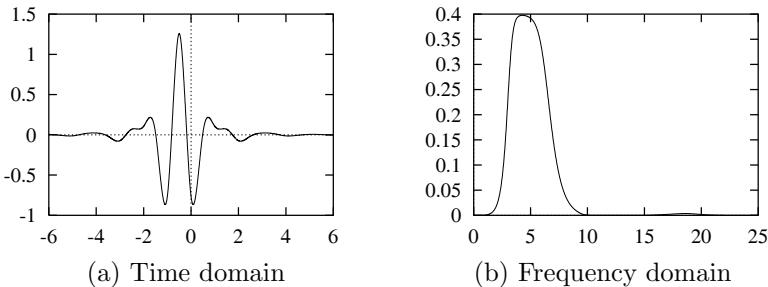


Fig. 4.9 B-spline wavelet of order 3 and its frequency spectrum.

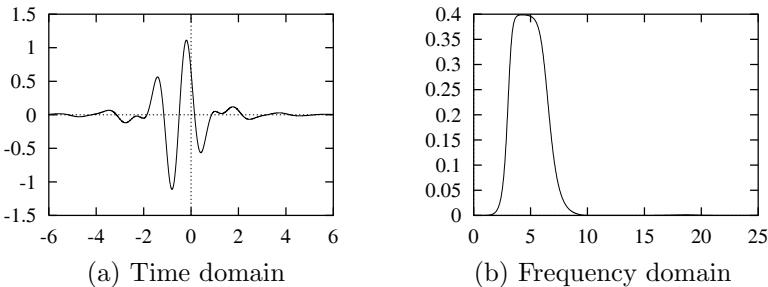


Fig. 4.10 B-spline wavelet of order 4 and its frequency spectrum.

4.1.5 Daubechies' Compactly Supported Orthonormal Wavelets

In the above discussion, except Haar wavelet, all presented orthonormal wavelets, though decay exponentially fast at infinity, do not compactly supported. The compactly supported orthonormal wavelets play an important role in wavelet decomposition of signals and data compression. In practice, the computation with compactly supported orthonormal wavelets need not to be truncated, has higher speed and precision.

Daubechies' compactly supported orthonormal wavelets are constructed by letting the filter function as follows according to Theorem 3.8:

$$m_0(\xi) = \left(\frac{1 + e^{-i\xi}}{2} \right)^L M_0(\xi)$$

To ensure the orthogonality of the wavelet basis, m_0 is constructed so that

$$|m_0(\xi)|^2 + |m_0(\xi + \pi)|^2 = 1, \quad (\forall \xi \in [-\pi, \pi]). \quad (4.11)$$

B-spline	Time Window		Freq. Window		Time-Freq. Area
	Center	Radius	Center	Radius	
1	-0.5	0.3941	2.5177	2.3504	3.70514
2	-0.5	0.4818	2.4104	1.9093	3.67968
3	-0.5	0.5519	2.3857	1.8476	4.07853
4	-0.5	0.5626	2.3772	1.8252	4.10739
5	-0.5	0.6693	2.3896	1.8127	4.85290
6	-0.5	0.7151	2.3607	1.8028	5.15675
7	-0.5	0.7502	2.3524	1.9910	5.39881
8	-0.5	0.7961	2.3377	1.7980	5.72522
9	-0.5	0.8144	2.3087	1.7999	5.86324
10	-0.5	0.8234	2.3940	1.7916	5.9008

Table 4.3 The centers, radii and area of time window, frequency window and time-frequency window of B-spline wavelets of orders 1 ~ 10

It is clear that

$$\begin{aligned}
 |m_0(\xi)|^2 &= \left| \frac{1 + e^{-i\xi}}{2} \right|^{2L} |M_0(\xi)|^2 \\
 &= \left| \cos \frac{\xi}{2} \right|^{2L} |M_0(\xi)|^2 \\
 &= \left(1 - \sin^2 \frac{\xi}{2} \right)^L |M_0(\xi)|^2.
 \end{aligned}$$

On the other hand, to ensure that ϕ and ψ are real-valued functions, we let $M_0(\xi)$ be real-valued functions on $e^{-i\xi}$, then

$$|M_0(\xi)|^2 = M_0(\xi) \overline{M_0(\xi)} = M_0(\xi) M_0(-\xi),$$

which concludes that $|M_0(\xi)|^2$ is an even function on ξ , or equivalently, a function on $\cos \xi$ and therefore on $\sin^2 \frac{\xi}{2}$ since

$$\cos \xi = 1 - 2 \sin^2 \frac{\xi}{2}.$$

Let us denote

$$|M_0(\xi)|^2 = P(\sin^2 \frac{\xi}{2}) = P(y), \quad y = y(\xi) = \sin^2 \frac{\xi}{2}.$$

Then

$$|m_0(\xi)|^2 = \left(1 - \sin^2 \frac{\xi}{2}\right)^L P(y) = (1 - y)^L P(y).$$

Note that

$$y(\xi + \pi) = \sin^2 \frac{\xi + \pi}{2} = \cos^2 \frac{\xi}{2} = 1 - \sin^2 \frac{\xi}{2} = 1 - y(\xi) = 1 - y,$$

by (4.11), $P(y)$ should satisfy

$$(1 - y)^L P(y) + y^L P(1 - y) = 1. \quad (4.12)$$

The general solution of Equation (4.12) is as follows (see [Daubechies, 1992, Chapter 6]):

$$P(y) = P_L(y) + y^L R\left(\frac{1}{2} - y\right),$$

where $R(x)$ is an odd polynomial on x and

$$P_L(y) := \sum_{k=0}^{L-1} \binom{L+k-1}{k} y^k.$$

Particularly, letting $R(x) \equiv 0$ we have $P(y) = P_L(y)$, therefore

$$|M_0(\xi)|^2 = P(\sin^2 \frac{\xi}{2}) = P_L(\sin^2 \frac{\xi}{2}) = \sum_{k=0}^{L-1} \binom{L+k-1}{k} \sin^{2k} \frac{\xi}{2}. \quad (4.13)$$

In general, there can not be more than one real-valued polynomial $M_0(\xi)$ satisfying (4.13). Riesz lemma (see [Daubechies, 1992, Chapter 6]) ensures the existence of such an $M_0(\xi)$: it is a real-valued polynomial of order $L-1$ on $e^{i\xi}$. Since $|M_0(\xi)|^2 = |M_0(-\xi)|^2$, replacing ξ by $-\xi$ we get another solution, which we denote by:

$$M_0(\xi) = \sum_{n=0}^{L-1} b_n e^{-in\xi}, \quad b_n \in \mathbb{R}.$$

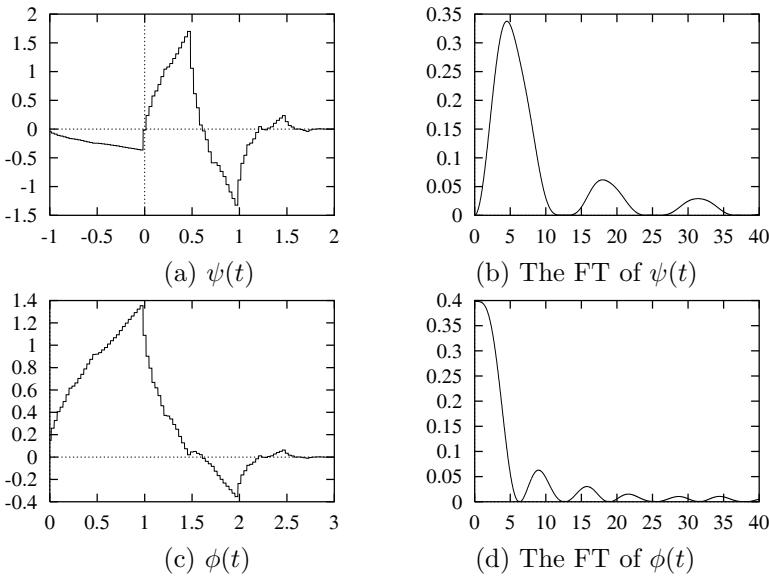


Fig. 4.11 The first type Daubechies orthonormal wavelet and its scaling function ($L=2$).

Therefore,

$$m_0(\xi) = \left(\frac{1 + e^{-i\xi}}{2} \right)^L M_0(\xi) = \left(\frac{1 + e^{-i\xi}}{2} \right)^L \sum_{n=0}^{L-1} b_n e^{-in\xi}$$

satisfies (4.11). It is easy to see that $m_0(\xi)$ can be rewritten as follows:

$$m_0(\xi) = \sum_{n=0}^{2L-1} h_n e^{-in\xi}. \quad (4.14)$$

A family of $m_0(\xi)$ is constructed by I. Daubechies [Daubechies, 1988] by choosing $M_0(\xi)$ satisfying (4.13) ($L \geq 2$) so that the corresponding ψ is a compactly supported wavelet with extremal phase and the highest number of vanishing moments compatible with its support width. The corresponding wavelet bases are called Daubechies' wavelet bases. They are orthonormal, compactly supported and regular wavelet bases. It can be proven that $\psi(t)$ is compactly supported on $[-L+1, L]$ and $\phi(t), \psi(t) \in C^{\mu N}$, where $\mu = \frac{\log 3}{2\log 2} \approx 0.2075$. The filter coefficients $\{h_n\}$ of (4.14) are listed in Table 4.4 for $L = 1 \sim 10$.

n	h_n (L=1)	n	h_n (L=9)
0	0.500000000000	4	-0.159993299446
1	0.500000000000	5	-0.091759032030
n	h_n (L=2)	6	0.068944046487
0	0.341506350946	7	0.019461604854
1	0.591506350946	8	-0.022331874165
2	0.158493649054	9	0.000391625576
3	-0.091506350946	10	0.003378031182
n	h_n (L=3)	11	-0.000761766903
0	0.235233603892	n	h_n (L=7)
1	0.570558457916	0	0.055049715373
2	0.325182500263	1	0.280395641813
3	-0.095467207784	2	0.515574245831
4	-0.060416104155	3	0.332186241105
5	0.024908749866	4	-0.101756911232
n	h_n (L=4)	5	-0.158417505640
0	0.162901714026	6	0.050423232505
1	0.505472857546	7	0.057001722580
2	0.446100069123	8	-0.026891226295
3	-0.019787513118	9	-0.011719970782
4	-0.132253583684	10	0.008874896190
5	0.021808150237	11	0.000303757498
6	0.023251800536	12	-0.001273952359
7	-0.007493494665	13	0.000250113427
n	h_n (L=5)	n	h_n (L=10)
0	0.113209491292	0	0.018858578796
1	0.426971771352	1	0.133061091397
2	0.512163472130	2	0.372787535743
3	0.097883480674	3	0.486814055367
4	-0.171328357691	0	0.038477811054
5	-0.022800565942	1	0.221233623576
6	0.054851329321	2	0.477743075214
7	-0.004413400054	3	0.413908266211
8	-0.008895935051	4	-0.011192867667
9	0.002358713969	5	-0.200829316391
n	h_n (L=6)	6	0.000334097046
0	0.078871216001	7	0.091038178423
1	0.349751907037	8	-0.012281950523
2	0.531131879941	9	-0.031175103325
3	0.222915661465	10	0.009886079648
n	h_n (L=7)	11	0.006184422410
0	0.078871216001	12	-0.003443859628
1	0.349751907037	13	-0.000277002274
2	0.531131879941	14	0.000477614855
3	0.222915661465	15	-0.000083068631
n	h_n (L=8)	16	-0.000484973920
0	0.078871216001	17	-0.000082354503
1	0.349751907037	18	0.000066177183
2	0.531131879941	19	-0.000009379208
n	h_n (L=9)	20	

Table 4.4 The filter coefficients $\{h_n\}$ for compactly supported wavelet with extremal phase and the highest number of vanishing moments compatible with their support width. The $\{h_n\}$ are normalized so that $\sum_n h_n = m_0(0) = 1$

The centers, radii and area of time window, frequency window and time-frequency window of the first type Daubechies wavelets are listed in Table 4.5, whose compactly supported lengths are 3, 5, 7, 9, 11, 13, 15, 17, 19, respectively. The centers, radii and area of time window, frequency window and time-frequency window of the second Daubechies wavelets are shown in Table 4.6, whose compactly supported lengths are 7, 9, 11, 13, 15, 17, 19, respectively.

Daub. I	Time Window		Freq. Window		Time-Freq. Area
	Center	Radius	Center	Radius	
1	0.4353	0.3308	1.3559	3.0720	4.0650
2	0.4793	0.3486	1.3755	3.1644	4.4122
3	0.4748	0.4243	1.2911	3.5206	4.2774
4	0.4682	0.5000	1.2502	2.2042	4.4080
5	0.4597	0.5754	1.2241	2.0580	4.7368
6	0.4585	0.6511	1.2218	2.0000	5.2092
7	0.4512	0.7237	1.2121	1.9440	5.6274
8	0.4446	0.7960	1.2071	1.9330	6.1548
9	0.4378	0.8666	1.2038	1.9212	6.6542

Table 4.5 The centers, radii and area of time window, frequency window and time-frequency window of the first type Daubechies wavelets, whose compactly supported lengths are 3, 5, 7, 9, 11, 13, 15, 17, 19, respectively

It is easy to observe that Daubechies' wavelets ψ are asymmetric. It can be shown that any compactly supported, orthonormal and regular wavelet is always asymmetric. For "least asymmetric", Daubechies chooses another family of $M_0(\xi)$. The corresponding wavelets are called the second type of Daubechies orthonormal wavelets here to differ from those discussed above, the first type. The corresponding filter coefficients $\{h_n\}$ are listed in Table 4.7. For details, readers are referred to [Daubechies, 1992]. Figs. 4.11-4.19 and 4.20-4.26 show the scaling functions and the wavelet functions of these two types of Daubechies wavelets, as well as the Fourier spectra of the scaling functions and the frequency properties of the wavelet functions, respectively.

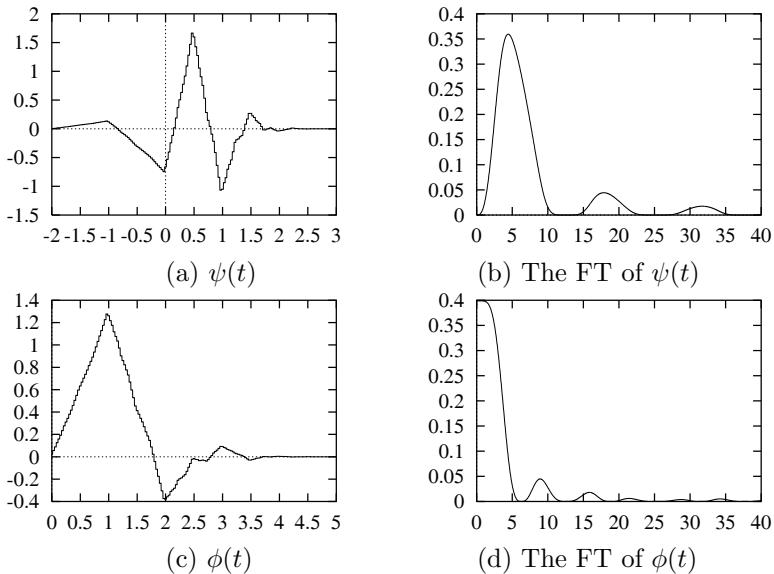


Fig. 4.12 The first type Daubechies orthonormal wavelet and its scaling function ($L=3$).

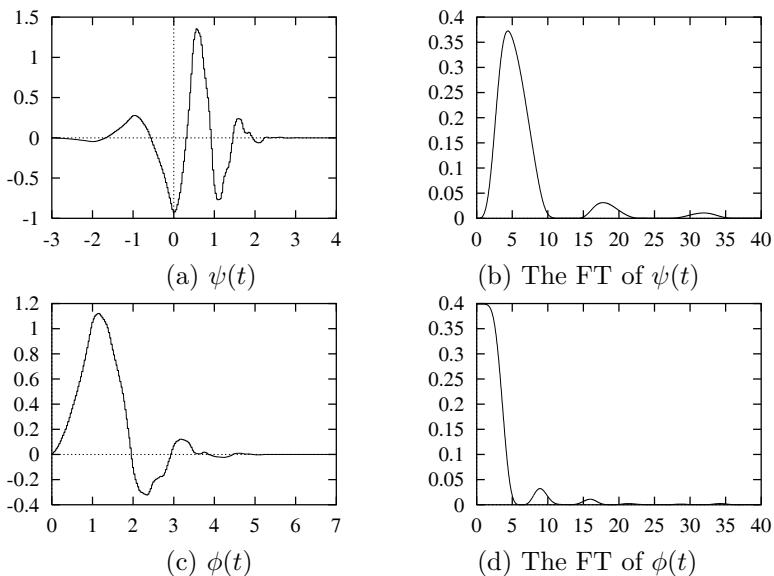


Fig. 4.13 The first type Daubechies orthonormal wavelet and its scaling function ($L=4$).

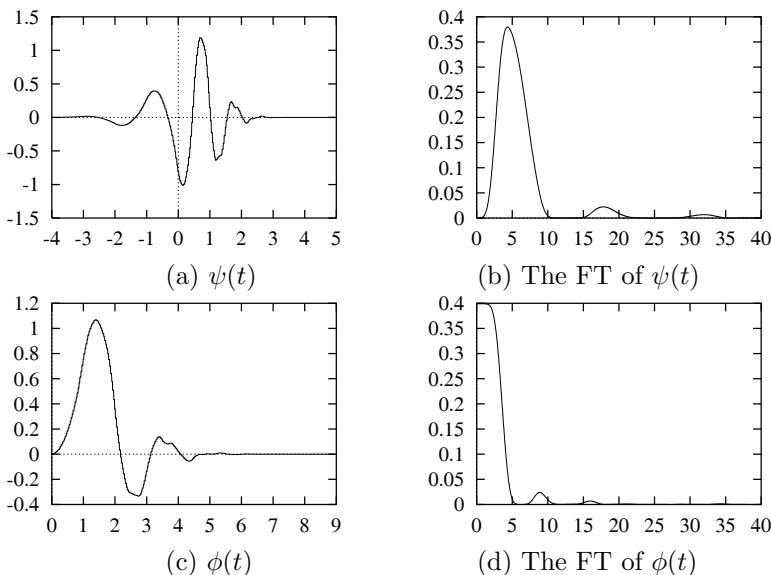


Fig. 4.14 The first type Daubechies orthonormal wavelet and its scaling function ($L=5$).

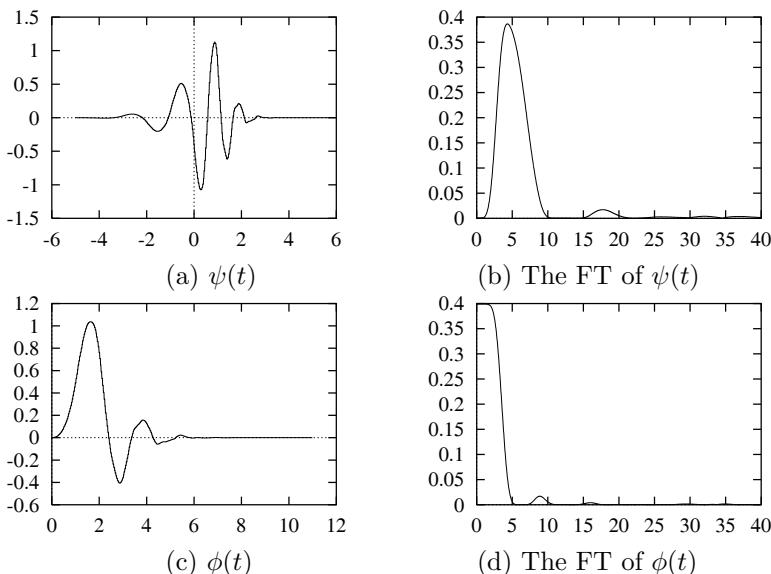


Fig. 4.15 The first type Daubechies orthonormal wavelet and its scaling function ($L=6$).

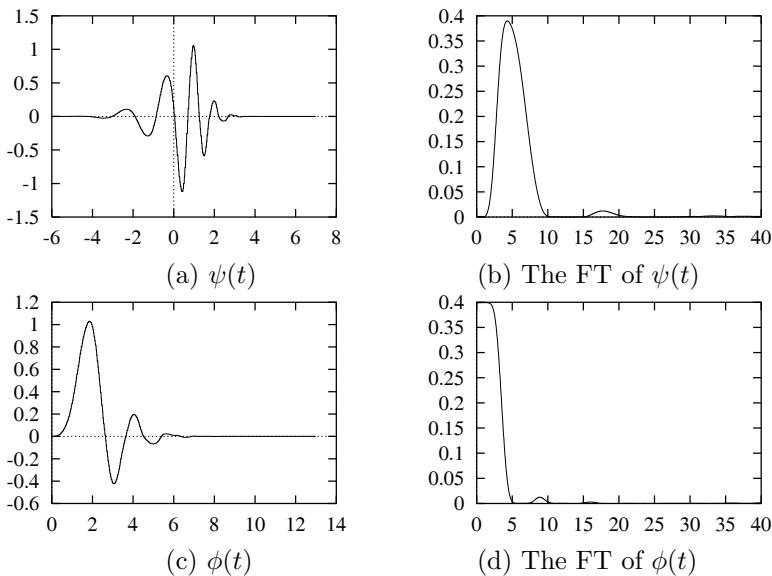


Fig. 4.16 The first type Daubechies orthonormal wavelet and its scaling function (L=7).

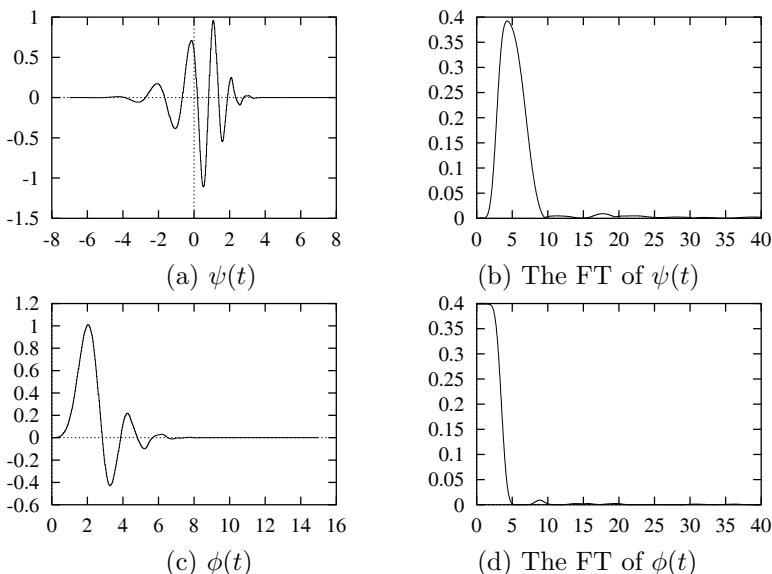


Fig. 4.17 The first type Daubechies orthonormal wavelet and its scaling function (L=8).

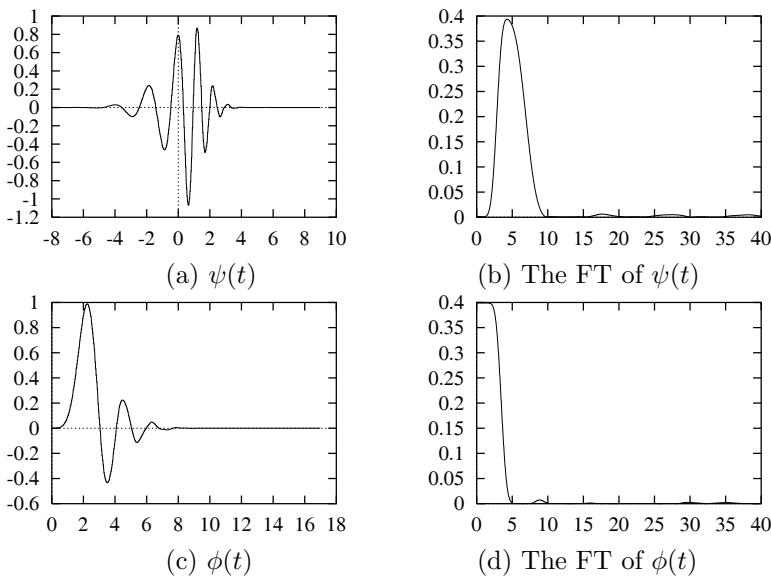


Fig. 4.18 The first type Daubechies orthonormal wavelet and its scaling function ($L=9$).

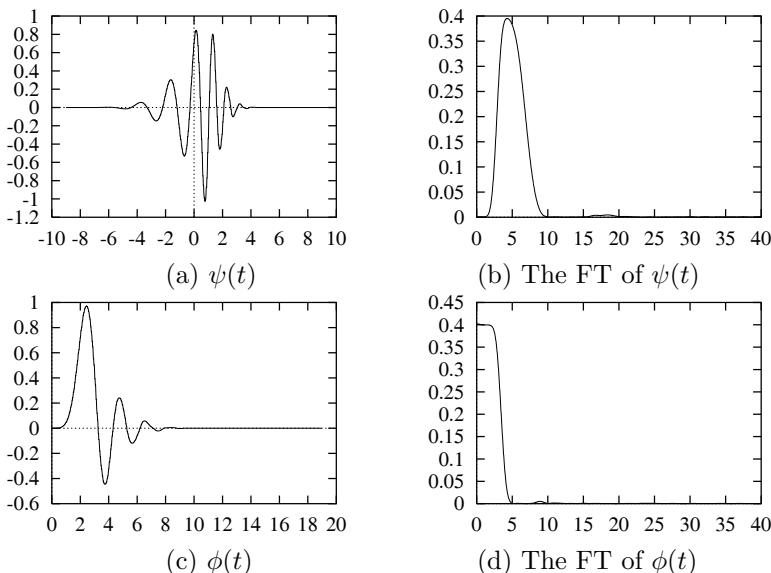


Fig. 4.19 The first type Daubechies orthonormal wavelet and its scaling function ($L=10$).

Daub.II	Time Window		Freq. Window		Time-Freq. Area
	Center	Radius	Center	Radius	
1	0.4077	0.3956	1.2486	2.2000	3.4813
2	0.3809	0.4519	1.1238	2.0537	3.7124
3	0.3417	0.4394	1.2202	2.0038	3.5217
4	0.2727	0.4722	1.2105	1.9462	3.6756
5	0.2542	0.4786	1.2073	1.9335	3.7022
6	0.2226	0.5189	1.2068	1.9333	4.0153
7	0.2110	0.5135	1.1941	1.8898	3.9820

Table 4.6 The centers, radii and area of time window, frequency window and time-frequency window of the second Daubechies wavelets, whose compactly supported lengths are 7, 9, 11, 13, 15, 17, 19, respectively

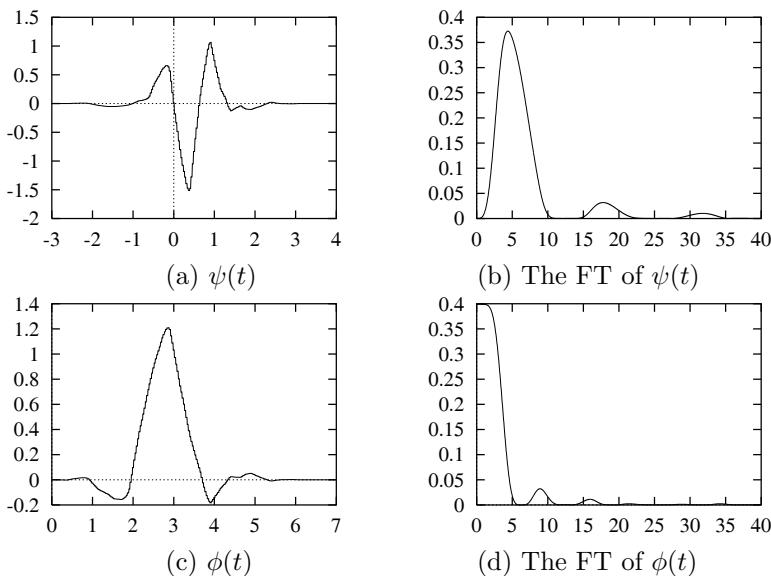


Fig. 4.20 The 2nd type Daubechies orthonormal wavelet and its scaling function ($L=4$).

n	h_n (L=4)	1	-0.000740612958	4	0.043895625777
0	-0.053574450709	2	-0.008935215825	5	-0.012893222965
1	-0.020955482562	3	0.021577726291	6	-0.135446891751
2	0.351869534328	4	0.048007383968	7	0.024941415480
3	0.568329121704	5	-0.035039145611	8	0.436524203675
4	0.210617267102	6	0.012332829745	9	0.507629895416
5	-0.070158812090	7	0.379081300982	10	0.168829461801
6	-0.008912350721	8	0.542891354907	11	-0.038586080548
7	0.022785172948	9	0.204091969863	12	0.000412570465
n	h_n (L=5)	10	-0.099028353403	13	0.021372216801
0	0.019327397978	11	-0.076231935948	14	-0.008151675613
1	0.020873432211	12	0.002835671343	15	-0.009384698418
2	-0.027672093058	13	0.007260697381	16	0.000438251270
n	h_n (L=8)	n	h_n (L=10)		
0	0.001336396696	0	0.000544585223		
1	-0.000214197150	1	0.000067622510		
2	-0.010572843264	2	-0.006110321315		
3	0.002693194377	3	-0.001036181962		
4	0.034745232955	4	0.032475462289		
5	-0.019246760632	5	0.008209434713		
6	-0.036731254380	6	-0.112779486117		
7	0.257699335187	7	-0.050120107516		
8	0.549553315269	8	0.333535669077		
9	0.340372673595	9	0.544125765250		
10	-0.043326807703	10	0.271406505606		
11	-0.101324327643	11	-0.025128270046		
12	0.005379305875	12	-0.022620386109		
13	0.022411811521	13	0.035351783775		
14	-0.000383345448	14	0.004076408400		
15	-0.002391729256	15	-0.014393115963		
n	h_n (L=9)	0	0.000756243654		
0	0.000756243654	1	-0.000334570754		
1	-0.000334570754	2	-0.007257789276		
2	-0.007257789276	3	0.006264448121		
n	h_n (L=7)	3	-0.000324794948		
0	0.001896329267				

Table 4.7 The filter coefficients $\{h_n\}$ for the “least asymmetric” wavelet basis, the second Daubechies’ wavelet basis. The $\{h_n\}$ are normalized so that $\sum_n h_n = m_0(0) = 1$

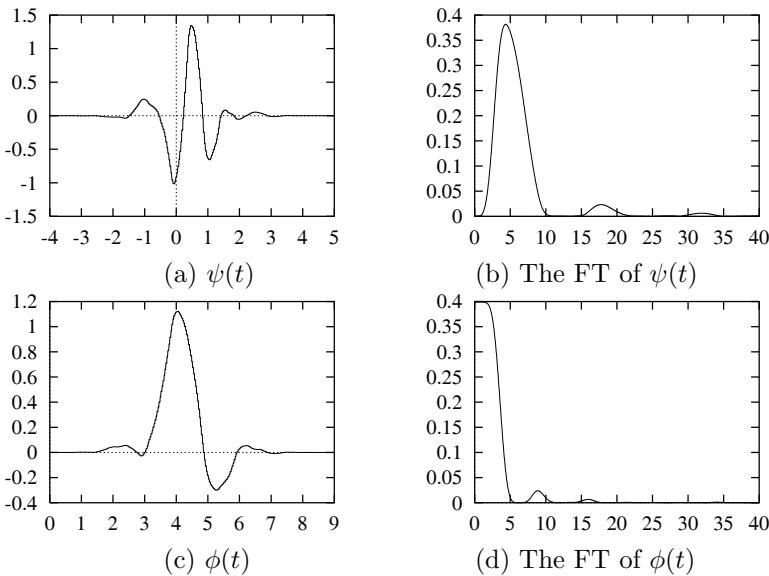


Fig. 4.21 The 2nd type Daubechies orthonormal wavelet and its scaling function ($L=5$).

4.1.6 Coiflet

In the applications of wavelet to the fast computation of singular integrals, A high vanishing moment of the wavelet applied is very important. To quicken the computation, R. Coifman suggested in the spring of 1989 that it might be worthwhile to construct orthonormal wavelet bases with vanishing moments not only for ψ , but also for ϕ . That is, we hope to find ψ and ϕ such that

$$\int_{\mathbb{R}} x^l \psi(x) dx = 0, \quad (l = 0, \dots, L-1), \quad (4.15)$$

and

$$\int_{\mathbb{R}} \phi(x) dx = 0, \quad \int_{\mathbb{R}} x^l \phi(x) dx = 0, \quad (l = 1, \dots, L-1), \quad (4.16)$$

Daubechies studied the construction of such wavelets [Daubechies, 1988; Daubechies, 1992], which are called “coiflet” because they were first requested by Coifman and the order is then called the order of coiflet.

To construct compactly supported orthonormal wavelet bases satisfying (4.15) and (4.16), it can be proven that m_0 has the form (see [Daubechies,

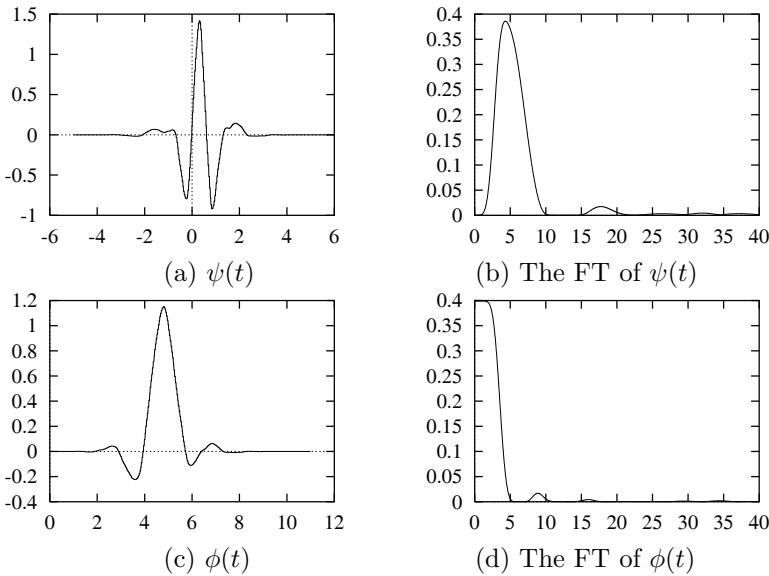


Fig. 4.22 The 2nd type Daubechies orthonormal wavelet and its scaling function ($L=6$).

1992]:

$$\begin{aligned} m_0(\xi) &= \left(\frac{1+e^{-i\xi}}{2}\right)^L M_0(\xi) \\ &= \sum_{n \in \mathbb{Z}} h_n e^{-in\xi}, \end{aligned}$$

where $L = 2K$ is a positive even integer,

$$M_0(\xi) := \sum_{k=0}^{K-1} \binom{K-1+k}{k} \left(\sin^2 \frac{\xi}{2}\right)^k + \left(\sin^2 \frac{\xi}{2}\right)^K f(\xi),$$

and $f(\xi) := \sum_{n=0}^{2K-1} f_n e^{-in\xi}$ is a trigonometric polynomial such that

$$|m_0(\xi)|^2 + |m_0(\xi + \pi)|^2 = 1 \quad (\forall \xi \in \mathbb{R}).$$

Coiflet of order L is a compactly supported orthonormal wavelet basis with support width $3L - 1$. It has vanishing moment of order $L - 1$ and the corresponding scaling function ϕ satisfies (4.16). Table 4.8 shows the localization property of coiflets. Figs. 4.27 ~ 4.31 show the scaling function and its frequency property, as well as the wavelet function and its frequency

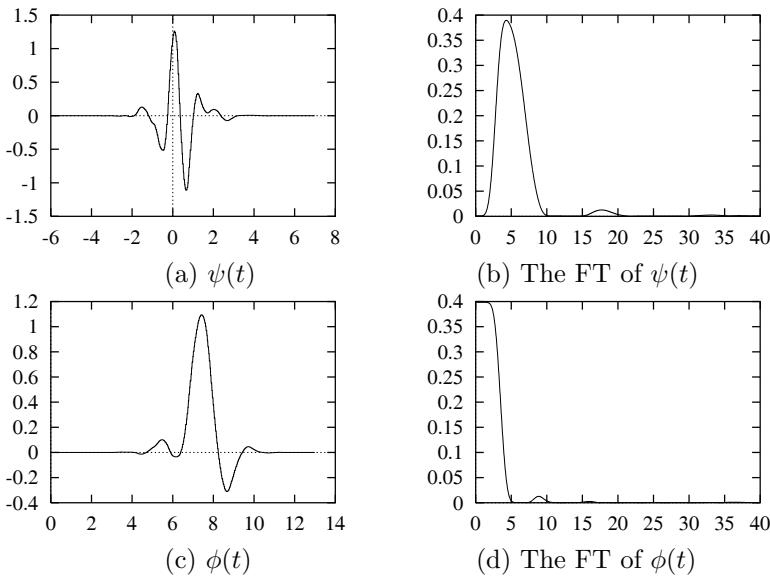


Fig. 4.23 The 2nd type Daubechies orthonormal wavelet and its scaling function ($L=7$).

property. The filter coefficients of the coiflets of orders $L = 2, 4, 6, 8, 10$ are listed in Table 4.9.

Coifman	Time Window		Freq. Window		Time-Freq. Area
	Center	Radius	Center	Radius	
1	0.4050	0.3808	1.3559	3.0720	4.6972
2	0.3756	0.3981	1.2458	2.1827	3.4759
3	0.3135	0.4430	1.2155	1.9945	3.5340
4	0.3761	0.4740	1.2080	1.9682	3.7314
5	0.3767	0.5050	1.2005	1.9419	4.2020

Table 4.8 The centers, radii and area of time window, frequency window and time-frequency window of R. Coifman wavelets, whose compactly supported lengths are 5, 11, 17, 23, 29, respectively

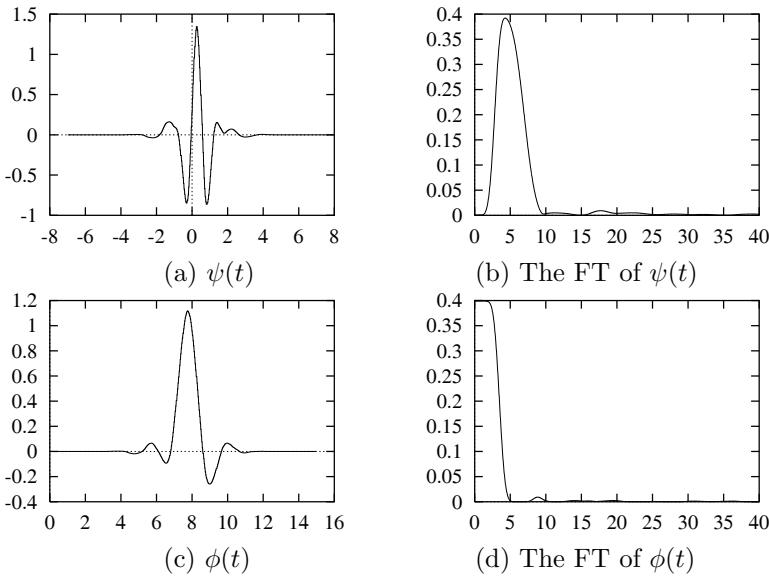


Fig. 4.24 The 2nd type Daubechies orthonormal wavelet and its scaling function ($L=8$).

4.2 Nonorthonormal Wavelet Bases

4.2.1 Cardinal Spline Wavelet

In this sub-section, we aim at constructing interpolatory spline wavelets based on B-splines, that is, fundamental cardinal splines, and discussing their properties.

For $m \geq 1$, the B-spline of order m $N_m(x)$ is defined by

$$N_m(x) = \underbrace{N_1 * \dots * N_1}_{m}(x), \quad (4.17)$$

where

$$N_1(x) = \chi_{[0,1)}(x) = \begin{cases} 1 & x \in [0, 1); \\ 0 & \text{otherwise.} \end{cases}$$

Then it can be deduced that that

$$\hat{N}_m(\xi) = \left(\frac{2}{\xi} e^{-i\xi/2} \sin \frac{\xi}{2} \right)^m.$$

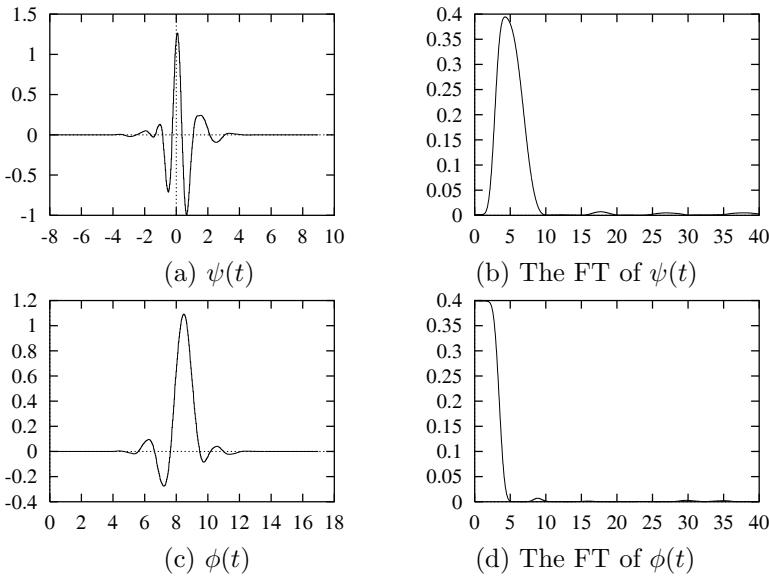


Fig. 4.25 The 2nd type Daubechies orthonormal wavelet and its scaling function ($L=9$).

Therefore, $N_m(x)$ satisfies the following two-scale equation:

$$\hat{N}_m(\xi) = \left(\frac{1 + e^{-i\xi/2}}{2} \right)^m \hat{N}_m(\xi/2).$$

By the expression of $\hat{N}_m(\xi)$, we can conclude that

$$\Phi_m(\xi) := \sum_{k \in \mathbb{Z}} |\hat{N}_m(\xi + 2k\pi)|^2$$

satisfies that

$$\left(\frac{2}{\pi} \right)^{2m} \leq \Phi_m(\xi) \leq 1 + 2^{2m+1} \left(\frac{2}{\pi} \right)^{2m} \sum_{k=1}^{\infty} \frac{1}{(2k-1)^{2m}} \quad (4.18)$$

for $m > 1$, and $\Phi_m(\xi) \equiv 1$ for $m = 1$. Hence, $\{N_m(x-k) | k \in \mathbb{Z}\}$ is a Riesz basis of

$$V_0^m := \overline{\text{span}\{N_m(x-k) | k \in \mathbb{Z}\}} \quad (4.19)$$

and

$$V_j^m := \{f(2^j x) | f(x) \in V_0^m\}, \quad (j \in \mathbb{Z}),$$

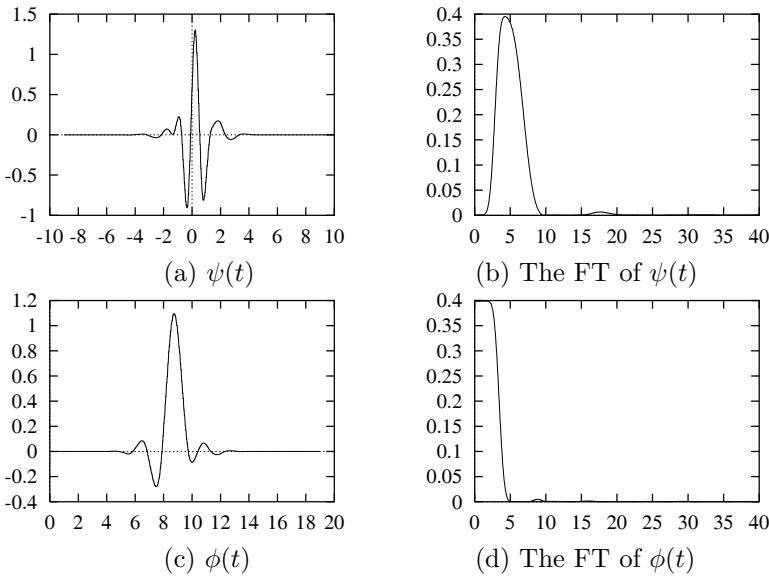


Fig. 4.26 The 2nd type Daubechies orthonormal wavelet and its scaling function ($L=10$).

constitutes a MRA (nonorthonormal except for the case of $m = 1$) of $L^2(\mathbb{R})$.

To estimate the bounds of $\Phi_m(\xi)$ in (4.18) more exactly, we use Poisson's summation formula to deduce that (see [Chui, 1992, p.48, p.89])

$$\begin{aligned}\Phi_m(\xi) &= \sum_{k \in \mathbb{Z}} \left(\int_{\mathbb{R}} N_m(t+k) \overline{N_m(t)} dt \right) e^{-ik\xi} \\ &= \sum_{k=-m+1}^{m-1} N_{2m}(m+k) e^{-ik\xi}.\end{aligned}$$

Note that $\Phi_m(\xi)$ and $N_m(\xi)$ are always nonpositive, we get the upper bound of $\Phi_m(\xi)$, i.e.,

$$\Phi_m(\xi) \leq \sum_{k=-m+1}^{m-1} N_{2m}(m+k) = 1.$$

The estimation of the lower bound of $\Phi_m(\xi)$ is more complicated though.

n	h_n (L=2)	n	h_n (L=10)
-2	-0.051429728471	5	0.011229240962
-1	0.238929728471	6	-0.006369601011
0	0.602859456942	7	-0.001820458916
1	0.272140543058	8	0.000790205101
2	-0.051429972847	9	0.000329665174
3	-0.011070271529	10	-0.000050192775
n	h_n (L=4)	11	-0.000024465734
-4	0.011587596739	n	h_n (L=8)
-3	-0.029320137980	-8	0.000630961046
-2	-0.047639590310	-7	-0.001152224852
-1	0.273021046535	-6	-0.005194524026
0	0.574682393857	-5	0.011362459244
1	0.294867193696	-4	0.018867235378
2	-0.054085607092	-3	-0.057464234429
3	-0.042026480461	-2	-0.039652648517
4	0.016744410163	-1	0.293667390895
5	0.003967883613	0	0.553126452562
6	-0.001289203356	1	0.307157326198
7	-0.000509505399	2	-0.047112738865
n	h_n (L=6)	3	-0.068038127051
-6	-0.002682418671	4	0.027813640153
-5	0.005503126709	5	0.017735837438
-4	0.016583560479	6	-0.010756318517
-3	-0.046507764479	7	-0.004001012886
-2	-0.043220763560	8	0.002652665946
-1	0.286503335274	9	0.000895594529
0	0.561285256870	10	-0.000416500571
1	0.302983571773	11	-0.000183829769
2	-0.050770140755	12	0.000044080354
3	-0.058196250762	13	0.000022082857
4	0.024434094321	14	-0.000002304942
		15	-0.000001262175

Table 4.9 The coefficients $\{h_n\}$ for coiflets of orders $L = 2, 4, 6, 8, 10$. They are normalized so that $\sum_n h_n = 1$

Let

$$E_{2m-1}(z) := (2m-1)! z^{m-1} \sum_{k=-m+1}^{m-1} N_{2m}(m+k) z^k \quad (4.20)$$

be the Euler-Frobenius polynomials of order $2m - 1$, which is in fact a

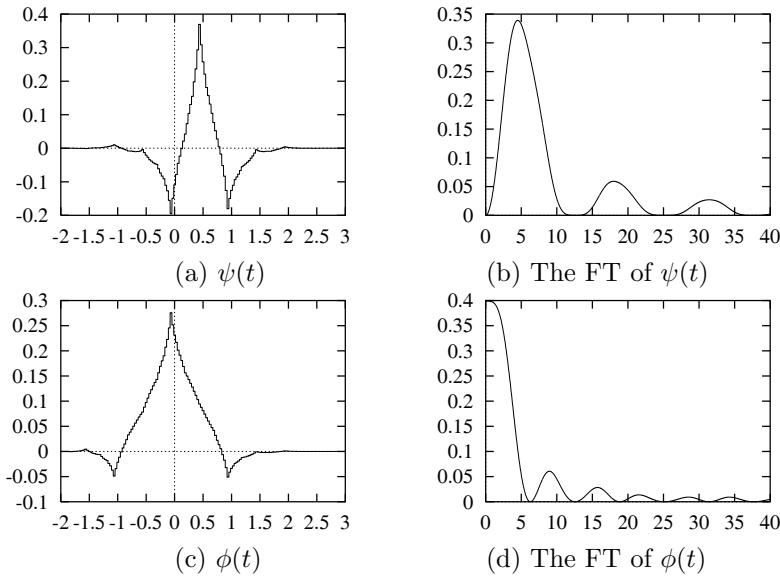


Fig. 4.27 Coiflet of order $L = 2$ and the corresponding scaling function.

polynomial on z of degree $2m - 2$. It can be shown that (see [Chui, 1992, Chapter 6] for details) the $2m - 2$ roots, $\lambda_1, \dots, \lambda_{2m-2}$, of $E_{2m-1}(z)$ are simple, real and negative, and furthermore, when they are arranged in a decreasing order, say,

$$0 > \lambda_1 > \dots > \lambda_{2m-2},$$

they satisfy

$$\lambda_1 \lambda_{2m-2} = \dots = \lambda_{m-1} \lambda_m = 1.$$

Hence we have

$$\begin{aligned} \Phi_m(\xi) &= \sum_{k=-m+1}^{m-1} N_{2m}(m+k) e^{-ik\xi} \\ &= \frac{1}{(2m-1)!} e^{i(m-1)\xi} E_{2m-1}(e^{-i\xi}). \end{aligned}$$

It is easy to see that the coefficient of z^{2m-2} in $E_{2m-1}(z)$ is

$$(2m-1)! N_{2m}(2m-1) = 1.$$

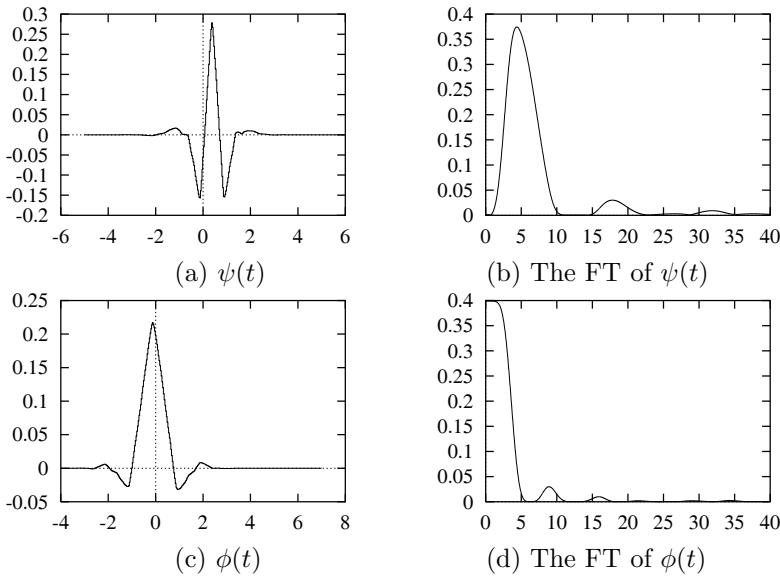


Fig. 4.28 Coiflet of order $L = 4$ and the corresponding scaling function.

Therefore,

$$E_{2m-1}(z) = \prod_{k=1}^{2m-2} (z - \lambda_k),$$

consequently,

$$\begin{aligned} \Phi_m(\xi) &= \frac{1}{(2m-1)!} e^{i(m-1)\xi} \prod_{k=1}^{2m-2} (e^{-i\xi} - \lambda_k) \\ &= \frac{1}{(2m-1)!} \prod_{k=1}^{2m-2} |e^{-i\xi} - \lambda_k| \\ &= \frac{1}{(2m-1)!} \left[\prod_{k=1}^{m-1} |e^{-i\xi} - \lambda_k| \right] \left[\prod_{k=m}^{2m-2} |e^{-i\xi} - \lambda_k| \right] \\ &= \frac{1}{(2m-1)!} \left[\prod_{k=1}^{m-1} |e^{-i\xi} - \lambda_k| \right] \left[\prod_{k=1}^{m-1} |e^{-i\xi} - \lambda_{2m-1-k}| \right] \end{aligned}$$

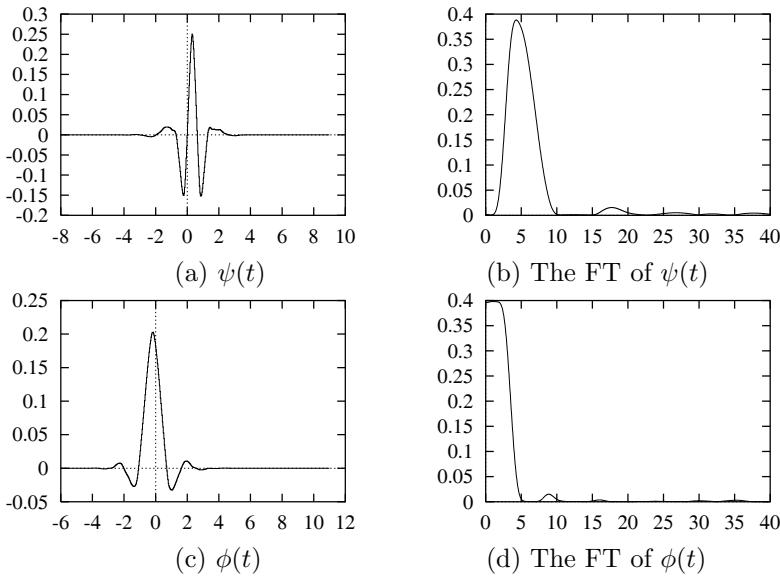


Fig. 4.29 Coiflet of order $L = 6$ and the corresponding scaling function.

$$\begin{aligned}
 &= \frac{1}{(2m-1)!} \prod_{k=1}^{m-1} \left[|e^{-i\xi} - \lambda_k| |e^{-i\xi} - \frac{1}{\lambda_k}| \right] \\
 &= \frac{1}{(2m-1)!} \prod_{k=1}^{m-1} \frac{|(e^{-i\xi} - \lambda_k)(\lambda_k e^{-i\xi} - 1)|}{|\lambda_k|} \\
 &= \frac{1}{(2m-1)!} \prod_{k=1}^{m-1} \frac{1 - 2\lambda_k \cos \xi + \lambda_k^2}{|\lambda_k|} \\
 &\geq \frac{1}{(2m-1)!} \prod_{k=1}^{m-1} \frac{1 + 2\lambda_k + \lambda_k^2}{|\lambda_k|} \\
 &= \Phi_m(\pi).
 \end{aligned}$$

Hence, $\Phi_m(\pi)$ and $\Phi_m(0) = 1$ are the lower and upper bounds of $\Phi_m(\xi)$, respectively, and these bounds are the best possible obviously.

The idea behind the construction of Battle-Lemarie wavelet basis is to

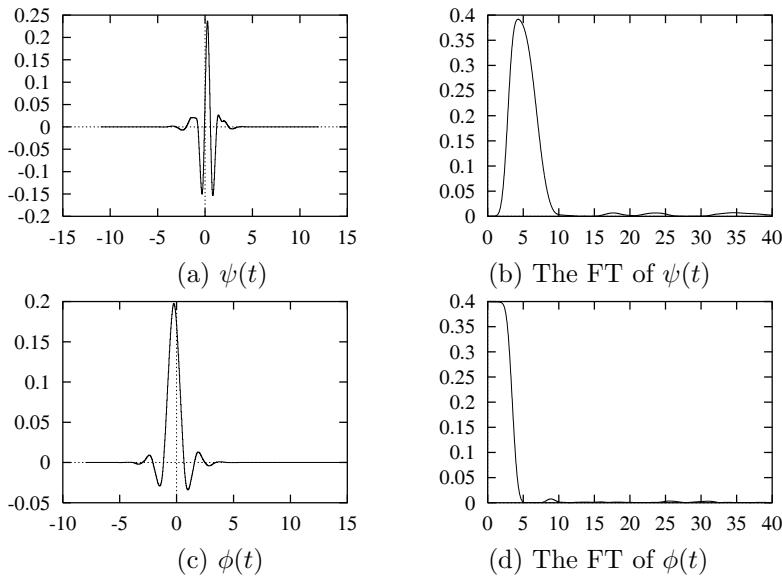


Fig. 4.30 Coiflet of order $L = 8$ and the corresponding scaling function.

orthonormalize the nonorthonormal Riesz basis by letting

$$\hat{\phi}^\sharp(\xi) := \frac{\hat{N}_m(\xi)}{\sqrt{\Phi_m(\xi)}}.$$

A drawback of such construction is that $\phi^\sharp(\xi)$ can neither be expressed explicitly nor be supported compactly as $N_m(\xi)$. To overcome this drawback, we keep V_j^m as above and let W_j^m be the orthogonal complement of V_{j+1}^m and V_j^m , i.e.,

$$V_{j+1}^m = V_j^m \oplus W_j^m. \quad (4.21)$$

Now our task is to construct a wavelet ψ such that

$$\psi_{j,k}(x) := 2^{j/2}\psi(2^j x - k), \quad (\forall j, k \in \mathbb{Z}).$$

constitutes a Riesz basis of W_j for any $j \in \mathbb{Z}$.

Let the m^{th} “fundamental cardinal spline function” $L_m(x)$ be defined

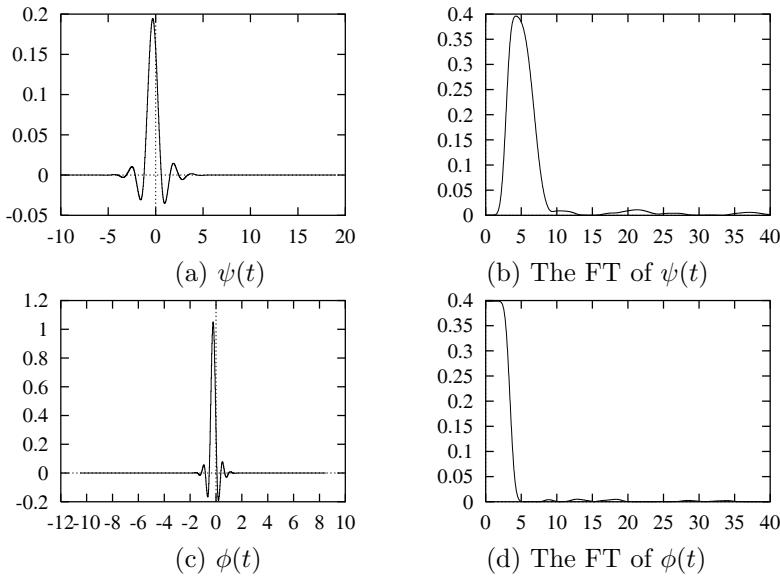


Fig. 4.31 Coiflet of order $L = 10$ and the corresponding scaling function.

by

$$L_m(x) := \sum_{k=-\infty}^{\infty} c_k^{(m)} N_m(x + \frac{m}{2} - k), \quad (4.22)$$

where, $\{c_k^{(m)}\}$ is obtained by solving the following bi-infinite system:

$$\sum_{k=-\infty}^{\infty} c_k^{(m)} N_m(j + \frac{m}{2} - k) = \delta_{j,0}. \quad (4.23)$$

We observe that (4.23) is equivalent to the following equation:

$$\left(\sum_{k=-\infty}^{\infty} c_k^{(m)} z^k \right) \left(\sum_{k=-\infty}^{\infty} N_m(\frac{m}{2} + k) z^k \right) = 1. \quad (4.24)$$

Since

$$\tilde{N}_m(z) := \sum_{k=-\infty}^{\infty} N_m(\frac{m}{2} + k) z^k$$

is a Laurent polynomial, which does not vanish on the unit circle $|z| = 1$ (see [Chui, 1992, p.111] for details), there exist $r, R : r < 1 < R$ such that

$$\sum_{k=-\infty}^{\infty} c_k^{(m)} z^k = \frac{1}{\tilde{N}_m(z)}$$

is analytic on $r < |z| < R$. Choose $z_0 : 1 < |z_0| < R$, then the convergence of $\sum_{k=-\infty}^{\infty} c_k^{(m)} z_0^k$ implies $\lim_{k \rightarrow \infty} c_k^{(m)} z_0^k = 0$. Therefore, $|c_k^{(m)}| = O(|z_0|^{-k})$, $k \rightarrow \infty$, i.e., $c_k^{(m)}$ decays to zero exponentially fast as $k \rightarrow \infty$. Similarly, we have that $c_k^{(m)}$ decays to zero exponentially fast as $k \rightarrow -\infty$. Moreover, it should be pointed out that $\{c_k^{(m)}\}$ is not finit for $m \geq 3$. In fact, it is easy to see that

$$\tilde{N}_m(z) = \frac{1}{z^{\lceil \frac{m-1}{2} \rceil}} \sum_{k=0}^{2[\frac{m-1}{2}]} N_m(\frac{m}{2} - [\frac{m-1}{2}] + k) z^k,$$

where $[x]$ denotes the largest integer not exceeding x . If $\{c_k^{(m)}\}$ is finit, by denoting

$$\sum_{k=-\infty}^{\infty} c_k^{(m)} z_0^k = \frac{1}{z^p} (a_0 + a_1 z + \cdots + a_q z^q),$$

where $a_0 \neq 0$ and p, q are negative integers and using (4.24), we have

$$\frac{1}{z^p} (a_0 + a_1 z + \cdots + a_q z^q) \frac{1}{z^{\lceil \frac{m-1}{2} \rceil}} \sum_{k=0}^{2[\frac{m-1}{2}]} N_m(\frac{m}{2} - [\frac{m-1}{2}] + k) z^k = 1,$$

i.e.,

$$(a_0 + a_1 z + \cdots + a_q z^q) \sum_{k=0}^{2[\frac{m-1}{2}]} N_m(\frac{m}{2} - [\frac{m-1}{2}] + k) z^k = z^{p+[\frac{m-1}{2}]}.$$
 (4.25)

On one hand, the constant item of the left side of (4.25) is $a_0 N_m(\frac{m}{2} - [\frac{m-1}{2}])$, which does not vanish; on the other hand, the constant item of the right side of (4.25) is obviously 0 since $[\frac{m-1}{2}] \geq 1$, provided that $m \geq 3$. The contradiction proves that $\{c_k^{(m)}\}$ must be infinit.

It is now easy to see that L_m satisfies the following interpolation property:

$$L_m(j) = \delta_{j,0}.$$

Therefore, for any series $\{f(k)\}$ with at most polynomial growth, if we define an interpolation spline operator as follows:

$$J_m f(x) := \sum_{k=-\infty}^{\infty} f(k) L_m(x - k), \quad (4.26)$$

then, the right side of (4.26) certainly converges at every $x \in \mathbb{R}$ due to the growth limit of $\{f(k)\}$, and J_m satisfies the following interpolation property:

$$(J_m f)(j) = f(j), \quad (\forall j \in \mathbb{Z}).$$

It can be proven that the coefficient sequence $\{c_k^{(m)}\}$ is not finite for each $m \geq 3$, so that the fundamental cardinal spline $L_m(x)$ does not vanish identically outside a compact set, even though it certainly decays to zero exponentially fast as $x \rightarrow \infty$ (see [Chui, 1992]). The following theorem gives a construction of the so-called fundamental cardinal spline wavelets.

Theorem 4.3 *Let m be any positive integer, and define*

$$\psi_{I,m}(x) := L_{2m}^{(m)}(2x - 1),$$

where $L_{2m}(x)$ is the $(2m)^{\text{th}}$ order fundamental cardinal spline. Then

$$\{\psi_{I,m}(x - k) | k \in \mathbb{Z}\}$$

constitutes a Riesz basis of W_0^m . Consequently,

$$\{2^{j/2} \psi_{I,m}(2^j x - k) | j, k \in \mathbb{Z}\}$$

constitutes a (Riesz) wavelet basis of $L^2(\mathbb{R})$.

For details of the proof, readers are referred to [Chui, 1992, pp. 142–145, 178–181] and are mentioned to notice that $W_i \perp W_j$ ($\forall i \neq j$).

The wavelet $\psi_{I,m}(x)$ constructed by Theorem 4.3 has an explicit expression. It can be deduced that

$$\psi_{I,m}(x) = \sum_{n=-\infty}^{\infty} q_n N_m(2x - n), \quad (4.27)$$

with

$$q_n := \sum_{l=0}^m (-1)^l \binom{m}{l} c_{m+n-1-l}^{(2m)},$$

which decays to zero exponentially fast as $k \rightarrow \pm\infty$ as $c_k^{(m)}$ does.

Denote

$$Q(z) := \sum_{n=-\infty}^{\infty} q_n z^n.$$

It is easy to deduce that

$$Q(z) = z^{1-m} (1-z)^m \sum_{n=-\infty}^{\infty} c_n^{(2m)} z^n.$$

Therefore, to calculate $\{q_n\}$ equals to calculate $\{c_n^{(2m)}\}$. We introduce

$$F_m(z) := \sum_{k=-m+1}^{m-1} N_{2m}(m+k) z^k = \sum_{k=-\infty}^{\infty} N_{2m}(m+k) z^k.$$

Since

$$\begin{aligned} F_m(z) \sum_{n=-\infty}^{\infty} c_n^{(2m)} z^n &= \sum_{k=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} c_n^{(2m)} N_{2m}(m+k) z^{n+k} \\ &= \sum_{n=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} c_n^{(2m)} N_{2m}(m+k-n) z^k \\ &= \sum_{k=-\infty}^{\infty} \left[\sum_{n=-\infty}^{\infty} c_n^{(2m)} N_{2m}(m+k-n) \right] z^k \\ &= \sum_{k=-\infty}^{\infty} L_{2m}(k) z^k \\ &= 1, \end{aligned}$$

we have

$$\sum_{n=-\infty}^{\infty} c_n^{(2m)} z^n = \frac{1}{F_m(z)}.$$

Consequently,

$$Q(z) = z^{1-m} (1-z)^m \frac{1}{F_m(z)} = (2m-1)! \frac{(1-z)^m}{E_{2m-1}(z)},$$

where

$$\begin{aligned} E_{2m-1}(z) &:= (2m-1)!z^{m-1}F_m(z) \\ &= (2m-1)!z^{m-1} \sum_{k=-m+1}^{m-1} N_{2m}(m+k)z^k \end{aligned}$$

is the Euler-Frobenius polynomial of order $2m-1$ defined by (4.20). Generally, the modified Euler-Frobenius polynomial of order n is defined by

$$\tilde{E}_n(z) := n! \sum_{k=0}^{n-1} N_{n+1}(k+1)z^k.$$

In particular, Table 4.10 lists 9 modified Euler - Frobenius polynomials. Although $\psi_{I,m}(x)$ can be expressed analytically, it is however not sup-

n	$\tilde{E}_n(z)$
1	1
2	$1+z$
3	$1+4z+z^2$
4	$1+11z+11z^2+z^3$
5	$1+26z+66z^2+26z^3+z^4$
6	$1+57z+302z^2+302z^3+57z^4+z^5$
7	$1+120z+1191z^2+2416z^3+1191z^4+120z^5+z^6$
8	$1+247z+4293z^2+15619z^3+15619z^4+4293z^5+247z^6+z^7$
9	$1+502z+14608z^2+88234z^3+156190z^4+88234z^5+14608z^6+502z^7+z^8$

Table 4.10 Modified Euler-Frobenius polynomials

ported compactly for $m \geq 2$. The reason is as follows: $\{c_n^{2m}\}$ is not finite for $m \geq 2$, therefore L_{2m} cannot be supported compactly, finally, neither is $\psi_{I,m}(x)$ for $m \geq 2$.

For $m = 1$, by

$$Q(z) = \frac{1-z}{E_1(z)} = 1-z,$$

we get

$$\psi_{I,1}(x) = N_1(2x) - N_1(2x-1),$$

which is the well-known Haar wavelet.

For $m = 2$, by

$$Q(z) = 6 \frac{(1-z)^2}{E_3(z)} = 6 \frac{(1-z)^2}{1+4z+z^2},$$

and

$$\begin{aligned} & \frac{1}{1+4z+z^2} \\ = & -\frac{1}{2\sqrt{3}} \left(\frac{1}{z+2+\sqrt{3}} - \frac{1}{z+2-\sqrt{3}} \right) \\ = & \frac{1}{2\sqrt{3}} \left(\frac{1}{z} \cdot \frac{1}{1+(2-\sqrt{3})z^{-1}} - \frac{1}{2+\sqrt{3}} \cdot \frac{1}{1+z(2+\sqrt{3})^{-1}} \right) \\ = & \frac{1}{2\sqrt{3}} \left[\frac{1}{z} \sum_{k=0}^{\infty} (\sqrt{3}-2)^k z^{-k} + \sum_{k=0}^{\infty} (\sqrt{3}+2)^{-(k+1)} z^k \right] \\ = & \frac{1}{2\sqrt{3}} \left[\sum_{k=0}^{\infty} (\sqrt{3}-2)^k z^{-(k+1)} + \sum_{k=0}^{\infty} (\sqrt{3}-2)^{k+1} z^k \right] \\ = & \frac{1}{2\sqrt{3}} \sum_{k=-\infty}^{\infty} (\sqrt{3}-2)^{|k+1|} z^k \end{aligned}$$

we get

$$\begin{aligned} Q(z) &= 6(1-z)^2 \frac{1}{2\sqrt{3}} \sum_{k=-\infty}^{\infty} (\sqrt{3}-2)^{|k+1|} z^k \\ &= \sqrt{3}(1-2z+z^2) \sum_{k=-\infty}^{\infty} (\sqrt{3}-2)^{|k+1|} z^k \\ &= \sqrt{3} \left[\sum_{k=-\infty}^{\infty} (\sqrt{3}-2)^{|k+1|} z^k - 2 \sum_{k=-\infty}^{\infty} (\sqrt{3}-2)^{|k+1|} z^{k+1} \right. \\ &\quad \left. + \sum_{k=-\infty}^{\infty} (\sqrt{3}-2)^{|k+1|} z^{k+2} \right] \\ &= \sqrt{3} \left[\sum_{k=-\infty}^{\infty} (\sqrt{3}-2)^{|k+1|} z^k \right. \\ &\quad \left. - 2 \sum_{k=-\infty}^{\infty} (\sqrt{3}-2)^{|k|} z^k + \sum_{k=-\infty}^{\infty} (\sqrt{3}-2)^{|k-1|} z^k \right] \end{aligned}$$

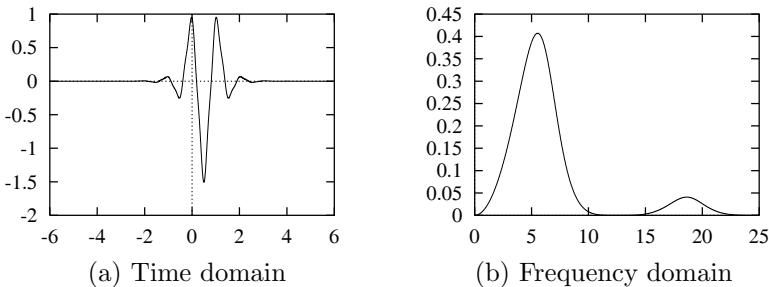


Fig. 4.32 C-spline wavelet $\psi_{I,2}$ of order 2 and its frequency spectrum.

$$= \sqrt{3} \sum_{k=-\infty}^{\infty} [(\sqrt{3}-2)^{|k+1|} - 2(\sqrt{3}-2)^{|k|} + (\sqrt{3}-2)^{|k-1|}] z^k.$$

It can be deduced that

$$q_0 = \sqrt{3} [(\sqrt{3}-2) - 2 + (\sqrt{3}-2)] = 6(1-\sqrt{3});$$

and for $k \neq 0$,

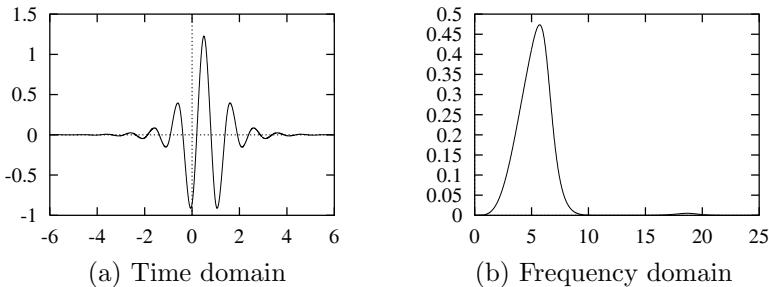
$$\begin{aligned} q_k &= \sqrt{3}[(\sqrt{3}-2)^{|k+1|} - 2(\sqrt{3}-2)^{|k|} + (\sqrt{3}-2)^{|k-1|}] \\ &= \sqrt{3}(\sqrt{3}-2)^{|k|}[(\sqrt{3}-2)^{|k+1|-|k|} - 2 + (\sqrt{3}-2)^{|k-1|-|k|}] \\ &= \sqrt{3}(\sqrt{3}-2)^{|k|}[(\sqrt{3}-2) - 2 + (\sqrt{3}-2)^{-1}] \\ &= \sqrt{3}(\sqrt{3}-2)^{|k|}[(\sqrt{3}-2) - 2 - (\sqrt{3}+2)] \\ &= \sqrt{3}(\sqrt{3}-2)^{|k|}[(\sqrt{3}-2) - 2 - (\sqrt{3}+2)] \\ &= -6\sqrt{3}(\sqrt{3}-2)^{|k|}. \end{aligned}$$

Therefore,

$$\psi_{I,2}(x) = -6\sqrt{3} \sum_{k=-\infty}^{\infty} (\sqrt{3}-2)^{|k|} N_2(2x-k).$$

C-spline wavelets $\psi_{I,2}, \psi_{I,4}$ of orders 2, 4 and their frequency properties, are illustrated in Figs. 4.32 and 4.33, respectively.

The centers, radii and areas of the time windows, frequency windows and time-frequency windows of the orders $2 \sim 10$ C-splines wavelets are given in Table 4.11.

Fig. 4.33 C-spline wavelet $\psi_{I,4}$ of order 4 and its frequency spectrum.

C-spline	Time Window		Freq. Window		Time-Freq. Area
	Center	Radius	Center	Radius	
2	0.5	0.4156	2.7547	2.4243	4.03025
3	0.5	0.4944	2.6757	2.0549	4.06344
4	0.5	0.5675	2.6767	2.0084	4.55896
5	0.5	0.6352	2.6854	1.9960	5.07147
6	0.5	0.6977	2.6930	1.9903	5.55437
7	0.5	0.7539	2.6969	1.9858	5.98852
8	0.5	0.8048	2.6979	1.9814	6.37891
9	0.5	0.8483	2.6946	1.9757	6.70379
10	0.5	0.8892	2.6885	1.9691	7.00385

Table 4.11 The centers, radii and area of time window, frequency window and time-frequency window of C-spline wavelets of orders 2 ~ 10

4.2.2 Compactly Supported Spline Wavelet

As discussed in the above subsection, $\psi_{I,m}$ is not supported compactly. To improve this property, another kind of spline wavelet, which is supported compactly, will be introduced in this sub-section.

Let

$$\psi_m(x) := 2 \sum_{n=-\infty}^{\infty} q_n N_m(2x - n), \quad (4.28)$$

with

$$\begin{aligned} q_n &:= \frac{(-1)^n}{2^m} \sum_{l=0}^m \binom{m}{l} N_{2m}(n+1-l) \\ &= \begin{cases} \frac{(-1)^n}{2^m} \sum_{l=0}^m \binom{m}{l} N_{2m}(n+1-l), & n = 0, 1, \dots, 3m-2, \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

Our main result in this sub-section is as follows:

Theorem 4.4 *Let m be a positive integer and $N_m(x)$ the m^{th} order cardinal B-spline defined by (4.17). Then $\psi_m(x)$, which is defined by (4.28), is a Riesz basis of W_0^m defined by (4.21) and*

$$\text{supp } \psi_m = [0, 2m-1].$$

Consequently,

$$\{2^{j/2} \psi_m(2^j x - k) | j, k \in \mathbb{Z}\}$$

constitutes a (Riesz) wavelet basis of $L^2(\mathbb{R})$.

Furthermore, the dual wavelet $\tilde{\psi}_m$ of ψ_m can be expressed as follows:

$$\tilde{\psi}_m(x) = \frac{(-1)^{m+1}}{2^{m-1}} \sum_{k=-\infty}^{\infty} c_k^{(2m)} \psi_{I,m}(x + m + 1 - k),$$

where $\psi_{I,m}$ is the wavelet deined by (4.27) and $\{c_k^{(2m)}\}$ is determinmed by bi-infinite-system (4.23). On the symmetricity of the wavelets, we can prove the following theorem.

Theorem 4.5 *Wavelets ψ_m , $\tilde{\psi}_m$ and $\psi_{I,m}$ are all symmetric for even m and antisymmetric for odd m . Consequently, they all have generalized linear phases.*

The details of the proof for the theorems are omitted here, readers are refered to [Chui, 1992, pp.183–184].

Figs. 4.34, 4.35 and 4.36 show the compactly supported spline wavelets ψ_m of orders 2, 4, 5 and their frequency properties, respectively. The centers, radii and areas of the time windows, frequency windows and time-frequency windows of ψ_m , the compactly supported splines wavelets of orders 2 ~ 10, are given in Table 4.12.

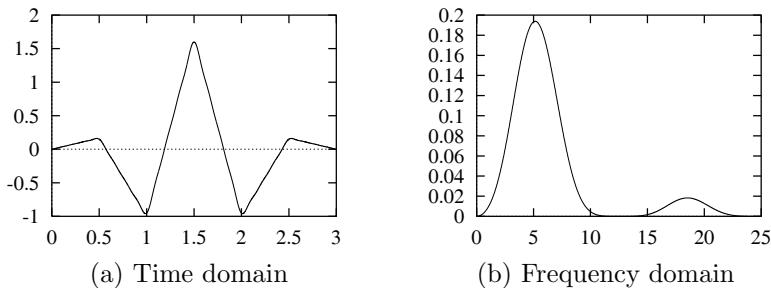


Fig. 4.34 Compactly supported spline wavelet ψ_2 of order 2 and its frequency spectrum.

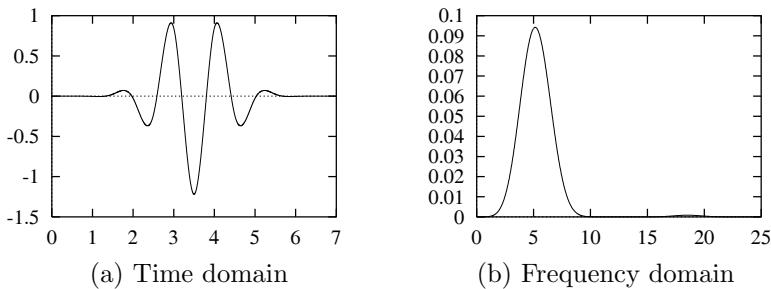


Fig. 4.35 Compactly supported spline wavelet ψ_4 of order 4 and its frequency spectrum.

Figs. 4.37, 4.38 and 4.39 show the dual wavelets $\tilde{\psi}_m$ of ψ_m ($m = 2, 4, 5$) and their frequency property, respectively. The centers, radii and areas of the time windows, frequency windows and time-frequency windows of the dual wavelet $\tilde{\psi}_m$ of orders $2 \sim 10$ are tabulated in Table 4.13.

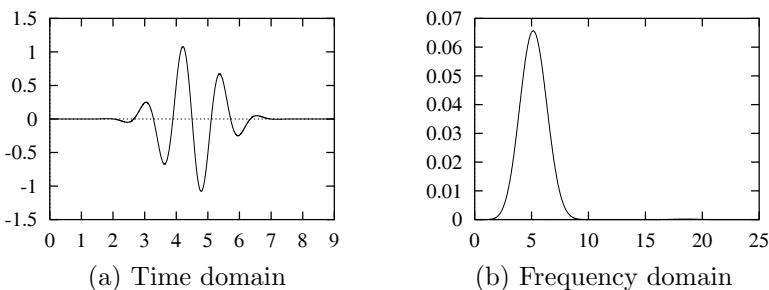


Fig. 4.36 Compactly supported spline wavelet ψ_5 of order 5 and its frequency spectrum.

CS-spline	Time Window		Freq. Window		Time-Freq. Area
	Center	Radius	Center	Radius	
2	1.5	0.3897	2.6788	2.3717	3.6968
3	2.5	0.4728	2.5907	2.0004	3.7835
4	3.5	0.5419	2.5821	1.9410	4.2075
5	4.5	0.6030	2.5975	1.9162	4.6221
6	5.5	0.6585	2.5780	1.9004	5.0059
7	6.5	0.7097	2.5769	1.8891	5.3628
8	7.5	0.7575	2.5762	1.8805	5.6977
9	8.5	0.8024	2.5756	1.8738	6.0141
10	9.5	0.8449	2.5751	1.8684	6.3146

Table 4.12 The centers, radii and area of time window, frequency window and time-frequency window of CS-spline wavelets of orders 2 to 10

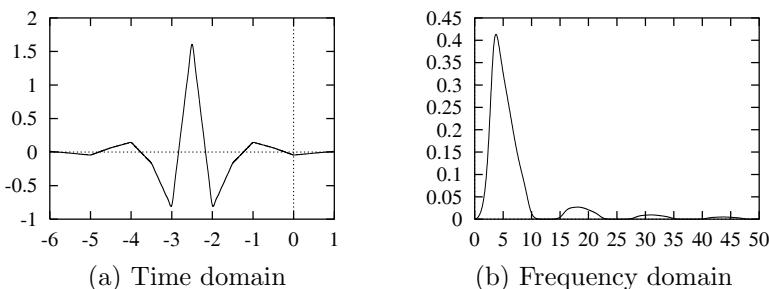


Fig. 4.37 The dual spline wavelet $\tilde{\psi}_2$ of Order 2 and its frequency spectrum.

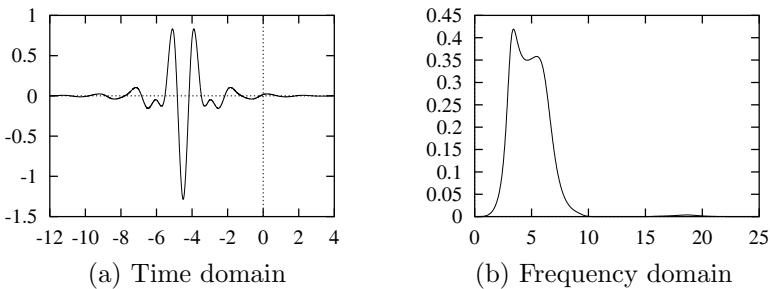


Fig. 4.38 The dual spline wavelet $\tilde{\psi}_4$ of order 4 and its frequency spectrum.

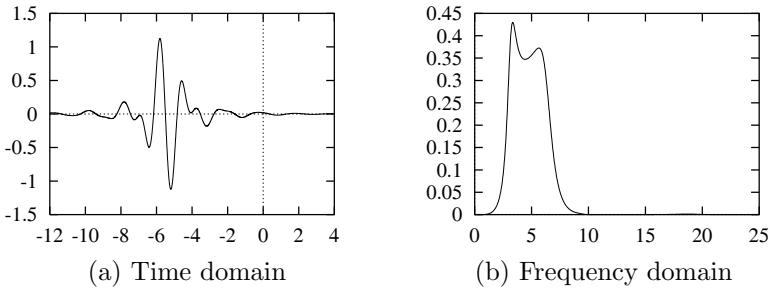


Fig. 4.39 The dual spline wavelet $\tilde{\psi}_5$ of order 5 and its frequency spectrum.

D-spline	Time Window		Freq. Window		Time-Freq. Area
	Center	Radius	Center	Radius	
2	-2.5	0.4554	2.3528	2.1734	3.9590
3	-3.5	0.5258	2.3073	2.1734	3.9202
4	-4.5	0.6164	2.3138	1.8301	4.5124
5	-5.5	0.7088	2.3235	1.8239	5.1709
6	-6.5	0.7981	2.3326	1.8829	6.0112
7	-7.5	0.8816	2.3409	1.8230	6.4281
8	-8.5	0.9583	2.3487	1.8233	6.9890
9	-9.5	1.0220	2.3553	1.8224	7.4498
10	-10.5	1.0900	2.3609	1.8207	7.9382

Table 4.13 The centers, radii and area of time window, frequency window and time-frequency window of D-spline wavelets of orders 2 ~ 10

Chapter 5

Step-Edge Detection by Wavelet Transform

As mentioned in the previous chapters, in pattern recognition, an entire task can be divided into four phases, namely, data acquisition, data pre-processing, feature extraction and classification. In the first phase, the data acquisition, analog data from the physical world are acquired through a transducer, and further digitized to discrete format suitable for computer processing. The physical variables, in the data acquisition phase, are converted into a set of measured data, which are then as the input to the next phase, i.e. the data preprocessing phase. A major function of the data preprocessing phase is to modify the measured data obtained from the data acquisition phase so that those data are more suitable for the further processing in the third phase (feature extraction). Many modifications are made in the data preprocessing phase, For example, some of them are listed below:

- Gray-level histogram modification;
- Smoothing and noise elimination;
- Edge sharpening;
- Boundary detection and contour tracing;
- Thinning;
- Segmentation;
- Morphological processing;
- Texture and object extraction from textural background;
- Approximation of curves and surfaces.

In the above modifications, many items, such as noise, edges, boundaries, surfaces, textures, curves, etc., are of singularities in different patterns.

Wavelet theory will play an important role to analyze and process such singularities. This chapter will give an example to show how the wavelet theory can be used to treat some singularities. Specially, we will study the applications of wavelet transform to edge detection.

5.1 Edge Detection with Local Maximal Modulus of Wavelet Transform

What is the edge of a signal? Intuitively, it is the transient component of the signal. In an one-dimensional signal, for example, the speech signal, the edge corresponds to the abrupt change of the tune which creates a harshness noise to the ears of human. In a two-dimensional signal, for example, an image, the edge refers to the sharp variation in color or gray. Actually, a signal is equivalent to a function, thus, the edge of the signal can be viewed as such a component, whose value varies suddenly. In mathematics, it can be represented by a larger derivative on that point. Consequently, the extraction of the edge eventuates in finding the pixels with the large derivative.

In numerical application, for instance, in pattern recognition, because the signal is in discrete form, the derivative can not be calculated accurately. As an alternation, thus, the difference quotient can be used to approximate it. In this way, many effective and classical edge detection operators were established based on the difference quotient, for example, Robert operator, Sobel operator, Laplace operator and so on. Unfortunately, all of these methods have a poor capability to reduce the noise due to the property of the noise itself, namely, it also is a catastrophe point and has a large value of amplitude in the signals.

As we have known, the noise always distributes randomly. From the point of view of the statistics, the average value of noise is nearly a constant in a certain area. In general, we can suppose the value of this constant is zero, and, take a weighted mean to the signal in this area, This action can be regarded as a low-pass filtering. In this way, the noise will be eliminated considerably. In mathematics, this method corresponds to the smoothing of a function. The scheme of the edge detection is a combination of the filtering and the derivation, in which the signal performs the filtering followed by the derivation. This idea was developed in 70s and 80s. Thereafter, Marr pointed out that the best smoothing operator is the convolution of the

original signal with Gaussian function. Essentially, the edge detection based on the wavelets developed in 1990s is an extension of this method.

If we focus the detection on the location of the edge but the width, then only the local maximal point needs to be found out, without counting other large derivative points in this area. It corresponds to the skeleton of the edge and is effective to the isolated points (in two-dimensional one, they are isolated in one direction). This is the basic idea of the edge detection with local maximal modulus of wavelet transform.

One-dimensional signal

Let $f(x)$ be an original signal, and $\theta(x)$ be a smoothing operator. Denote

$$\theta_s(x) := \frac{1}{s} \theta\left(\frac{x}{s}\right),$$

where $s > 0$ indicates the smoothing scale. To smooth $f(x)$, we take the convolution for $f(x)$:

$$f * \theta_s(x) = \int_{\mathbb{R}} f(x-t) \theta_s(t) dt$$

The purpose of smoothing $f(x)$ is to reduce the noise, but the edge. For this reason, $\theta(x)$ should be localized, such that $f * \theta_s(x) \sim f(x)$ when s is small enough. This implicates that the smoothed signal, which was processed by a little scale, is almost the same as the original one. This conclusion can be easily proved mathematically, when $\theta(x)$ is localized and $\int_{\mathbb{R}} \theta(x) dx = 1$ (assume that the original signal is continuous).

In fact, the edge detection always conflicts with the noise reduction. The more the edge information is extracted, the more the noise will be brought out, and vice versa. In other words, the smoothing should not be too strong, although it is necessary to smooth the original signal. An example of such smoothing function is shown in Fig. 5.1, which meets the above requirement.

Suppose $f(x)$ is smoothed with $\theta(x)$ at first. Its derivative is now computed as follows:

$$\frac{d}{dx}(f * \theta_s)(x) = \int_{\mathbb{R}} f(x-t)(\theta_s)'(t) dt = \int_{\mathbb{R}} f(x-t)(\theta_s)'(t) dt, \quad (5.1)$$

The edge pixel can be obtained by detecting the point, which possesses the local maximum of the absolute derivative.

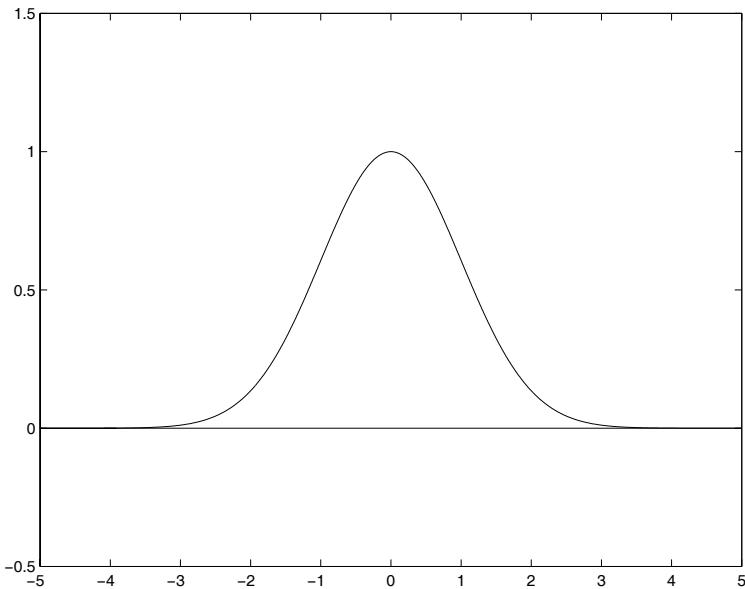


Fig. 5.1 A example of the smoothing function: Gaussian function.

Since

$$\int_{\mathbb{R}} \theta'(x) dx = \theta(+\infty) - \theta(-\infty) = 0,$$

Thus, $\theta'(x)$ is a wavelet function, and the right hand side of Eq. (5.1) is $\frac{1}{s}$ times of its corresponding wavelet transform.

Let $\psi(x) = \theta'(x)$, we have

$$W_s^\psi f(x) = \int_{\mathbb{R}} f(x-t)\psi_s(t)dt = s \frac{d}{dx} (f * \theta_s)(x).$$

This method is referred to as edge detection with local maximal modulus of wavelet transform.

It is known that the senses of audition and vision of human are limited. Hence, it is not necessary to extract all edge pixels, because some are too weak, so that they can not reach the senses of human beings. In practice, a threshold $T > 0$ is used to decide which point will be extract as an edge pixel, that must fulfill the following two conditions:

- (1) $W_s^\psi f(x_0) \geq T$;

(2) $|W_s^\psi f(x)|$ take the local maximum at x_0 .

To extract the edges of a signal $f(x)$, it is required to compute its wavelet transform $\{W_s^\psi f(x)\}$. A discrete input signal $\{f(0), f(1), \dots, f(n)\}$ can be viewed as the result of the A/D transformation from an analog signal. When applying wavelet transform to a digital signal, the discrete signal should be changed to a continuous analog signal by a suitable interpolation. Thereafter, its wavelet transform can be calculated by the following formula

$$W_s^\psi f(x) = \int_{\mathbb{R}} f(x-t) \psi_s(t) dt.$$

After sampling the transform result, we can obtain the discrete data

$$\{W_s^\psi f(0), W_s^\psi f(1), \dots, W_s^\psi f(n)\}.$$

Generally, different results will be achieved by different fitting methods.

It should be mentioned that the well-known zero-crossing technique for edge detection is almost the same as that with local maximal modulus of wavelet transform in principle. The former refers to the calculation of the zero-points of the 2-order derivative of the smoothing function $f * \theta_s(x)$, whereas, the latter focus on the computation of the points, which possess local maxima of 1-order derivative of $f * \theta_s(x)$. In mathematics, a point with local maximum of 1-order derivative must be the zero-point of 2-order derivative. For this reason, these two methods are the same in most of cases. We should note that a zero-point of 2-order derivative might not be the local maximal point of 1-order derivative, and it maybe refers to the minimal point of 1-order derivative of $f * \theta_s(x)$ in some cases. It implies that the pixels extracted by the zero-crossing method are probably of the false-edge. From this point, the edge detection method with local maximal modulus of wavelet transform is better than that of the zero-crossing one. In addition, the 1-order derivative of the smoothing function $f * \theta_s(x)$ can also be obtained by the method with local maximal modulus of wavelet transform. It has been discussed in details by Canny [Canny, 1986].

As the discrete value of wavelet transform, $\{W_s^\psi f(0), W_s^\psi f(1), \dots, W_s^\psi f(n)\}$, has been achieved, the next step is to compute the local maximal modulus of the wavelet transform. In practice, $x = m$ is a point with local maximal modulus, if it meets the following two conditions:

- (1) $|W_s^\psi f(m)| \geq |W_s^\psi f(m-1)|, |W_s^\psi f(m)| \geq |W_s^\psi f(m+1)|;$
- (2) $|W_s^\psi f(m)| > |W_s^\psi f(m-1)|$ or $|W_s^\psi f(m)| > |W_s^\psi f(m+1)|$.

The algorithm of finding the local maximal modulus of the wavelet transform involves the following steps:

Algorithm 5.1 Given an input discrete signal $\{f(0), f(1), \dots, f(n)\}$,

Step 1 To compute its wavelet transform $\{W_s^\psi f(0), W_s^\psi f(1), \dots, W_s^\psi f(n)\}$;

Step 2 To take a threshold $T > 0$, for $m = 0, 1, \dots, n$, if the following conditions satisfy

- (1) $|W_s^\psi f(m)| \geq T$,
- (2) $|W_s^\psi f(m)| \geq |W_s^\psi f(m-1)|$, $|W_s^\psi f(m)| \geq |W_s^\psi f(m+1)|$,
- (3) $|W_s^\psi f(m)| > |W_s^\psi f(m-1)|$ $|W_s^\psi f(m)| > |W_s^\psi f(m+1)|$,

then, $x = m$ is an edge pixel.

An example of the edge detection with the local maximal modulus of wavelet transform is graphically illustrated in Fig. 5.2.

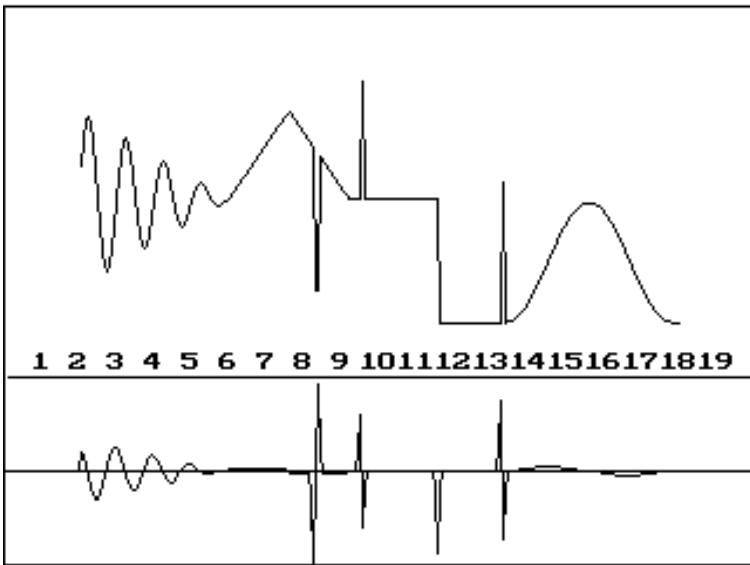


Fig. 5.2 Upper part: the original image; Lower part: the result of wavelet transform.

Two-Dimensional signal

We continue our discussion and extend our effort to 2-D signals. For a 2-D signal, for example, an image, its analysis by the wavelet theory and

the algorithm can be established in a similar way to that of 1-D one. In 2-D signal, the input image $f(x, y)$ and smoothing function $\theta(x, y)$ are 2-variable functions. Meanwhile, $\theta(x, y)$ should have good locality and satisfies

$$\int_{\mathbb{R}} \int_{\mathbb{R}} \theta(x, y) dx dy = 1.$$

Now, function $f(x, y)$ can be smoothed, and we shall describe this by writing $(f * \theta_s)(x, y)$:

$$(f * \theta_s)(x, y) := \int_{\mathbb{R}} \int_{\mathbb{R}} f(x - u, y - v) \theta_s(u, v) du dv,$$

where, $\theta_s(u, v) := \frac{1}{s^2} \theta(\frac{u}{s}, \frac{v}{s})$, and $s > 0$ stands for a smoothing scale.

When we calculate the derivative of a 2-D function, its orientation has to be considered. At each edge pixel, it reaches local maximum along the gradient direction, and can be represented by

$$\text{grad}(f * \theta_s)(x, y) = \mathbf{i} \frac{\partial}{\partial x} (f * \theta_s)(x, y) + \mathbf{j} \frac{\partial}{\partial y} (f * \theta_s)(x, y),$$

where \mathbf{i} and \mathbf{j} correspond to x -axis and y -axis respectively.

Now, we concentrate our discussion on the basic steps, in which, the edge detection with local maximal modulus of wavelet transform can be obtained, namely: first of all, the pixels will be extracted, such that the following modulus will reach its local maximum along the gradient direction

$$|\text{grad}(f * \theta_s)(x, y)| = \sqrt{\left| \frac{\partial}{\partial x} (f * \theta_s)(x, y) \right|^2 + \left| \frac{\partial}{\partial y} (f * \theta_s)(x, y) \right|^2},$$

and these pixels will become the components of the edge.

We can write

$$\psi^1(x, y) := \frac{\partial \theta}{\partial x}(x, y), \quad \psi^2(x, y) := \frac{\partial \theta}{\partial y}(x, y).$$

When $\theta(x, y)$ has good locality, i.e. it necessarily satisfy the conditions

$$\begin{aligned} \int_{\mathbb{R}} \int_{\mathbb{R}} \psi^1(x, y) dx dy &= \int_{\mathbb{R}} [\theta(+\infty, y) - \theta(-\infty, y)] dy = 0, \\ \int_{\mathbb{R}} \int_{\mathbb{R}} \psi^2(x, y) dx dy &= \int_{\mathbb{R}} [\theta(x, +\infty) - \theta(x, -\infty)] dx = 0, \end{aligned}$$

then, $\psi^1(x, y)$ and $\psi^2(x, y)$ become 2-D wavelets.

It is easy to know that

$$\begin{aligned} |grad(f * \theta_s)(x, y)| &= \frac{1}{s} \sqrt{|f * \psi_s^1(x, y)|^2 + |f * \psi_s^2(x, y)|^2} \\ &= \frac{1}{s} \sqrt{|W_s^{\psi^1} f(x, y)|^2 + |W_s^{\psi^2} f(x, y)|^2}, \end{aligned}$$

where

$$W_s^{\psi^i} f(x, y) := (f * \psi^i)(x, y) = \int_{\mathbb{R}} \int_{\mathbb{R}} f(x-u, y-v) \psi_s^i(u, v) du dv \quad (i = 1, 2).$$

stand for the corresponding wavelet transforms of ψ^1 , and ψ^2 respectively.

The modulus of wavelet transform of $f(x, y)$ can be defined by

$$M_s f(x, y) := \sqrt{|W_s^{\psi^1} f(x, y)|^2 + |W_s^{\psi^2} f(x, y)|^2}.$$

It is clear that

$$|grad(f * \theta_s)(x, y)| = \frac{1}{s} M_s f(x, y).$$

Consequently, the calculation of the local maximal modulus of the derivative of a smoothing function along the gradient direction is equivalent to the computation of the local maximal modulus of wavelet transform.

Similar to the case of 1-D, many methods can be used to compute the wavelet transform of a 2-D discrete signal. The next task is to determine the gradient direction of a digital signal as well as its local maximum along this direction.

The gradient direction of a function can be accurately counted in mathematics. However, by contrast, it is difficult to express exactly if it is in discrete form. Fortunately, for the data obtained by equi-spaced sampling, only eight adjacent pixels are around a point. Hence, only the nearest eight points will be taken into account. It is said that the discrete image has eight gradient directions. Therefore, a plane can be divided into eight sectors. A gradient direction is defined by

$$\alpha_s := \arctan \left(\frac{\partial(f * \theta_s)(x, y)}{\partial y} / \frac{\partial(f * \theta_s)(x, y)}{\partial x} \right), \quad s = 1, 2, \dots, 8,$$

When α falls into a sector, it will be quantified to a certain vector, which is represented by a center line of that sector. This can be found in Fig. 5.3, where, the arrow shows a vector, which indicates a direction of the gradient.

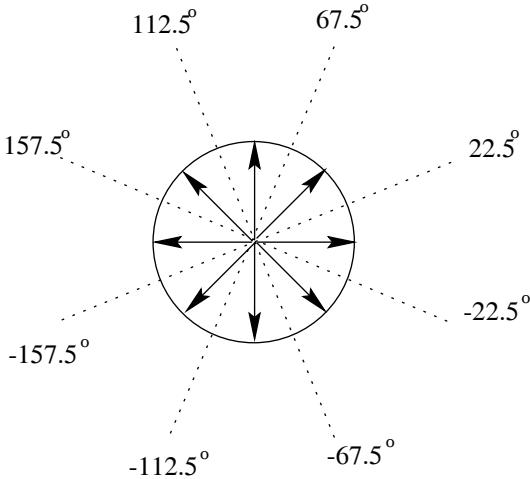


Fig. 5.3 The eight gradient directions (arrows), and the sectors divided by dashed lines.

Along this direction, a local maximal modulus can achieve. The effect of any opposite gradient directions is the same. Thus, only 4 codes are needed to represent the gradient directions. For example, numbers 0, 1, 2, 3 can be used to code these 4 different directions. The tangent of each direction, $\tan \alpha_s$, falls into one of the following intervals:

$$[-1 - \sqrt{2}, 1 - \sqrt{2}), [1 - \sqrt{2}, \sqrt{2} - 1),$$

$$[\sqrt{2} - 1, \sqrt{2} + 1), [\sqrt{2} + 1, +\infty) \cup (-\infty, -1 - \sqrt{2}).$$

We shall describe it by writing

$$A_s f(x, y) := \tan \alpha_s = \frac{\partial(f * \theta_s)(x, y)}{\partial y} \Big/ \frac{\partial(f * \theta_s)(x, y)}{\partial x}.$$

When $A_s f(x, y)$ falls into the above intervals, it will be coded by 0, 1, 2, 3 respectively. They can be viewed as the functions of values 0, 1, 2, 3, and symbolized by $Code A_s f(x, y)$. Similar to the 1-D signals, the algorithm of the edge detection with local maximal modulus of wavelet transform is now established below:

Algorithm 5.2 Given an input digital signal $\{f(k, l) | k = 0, 1, \dots, K; l = 0, 1, \dots, L\}$,

Step 1 To calculate the modulo of its wavelet transform

$$\{M_s f(k, l) | k = 0, 1, \dots, K; l = 0, 1, \dots, L\}$$

as well as the codes

$$\{CodeA_s f(k, l) | k = 0, 1, \dots, K; l = 0, 1, \dots, L\}$$

along the gradient directions;

Step 2 To take a threshold $T > 0$, for $k = 0, 1, \dots, K; l = 0, 1, \dots, L$, if

- (1) $|M_s f(k, l)| \geq T$;
- (2) $|M_s f(k, l)|$ reaches its local maximum along the gradient direction represented by $CodeA_s f(k, l)$,

then, (k, l) is an edge pixel.

An example is given in Fig. 5.4, where a spline function is utilized as the smoothing function:

$$\theta(x, y) := \begin{cases} 8(x^2 + y^2)(\sqrt{x^2 + y^2} - 1) + \frac{4}{3} & 0 \leq x^2 + y^2 \leq \frac{1}{4} \\ -\frac{8}{3}(\sqrt{x^2 + y^2} - 1)^3 & \frac{1}{4} < x^2 + y^2 \leq 1 \\ 0 & x^2 + y^2 > 1, \end{cases}$$

The graphical illustration of this smoothing function $\theta(x, y)$ is shown in Fig. 5.5.

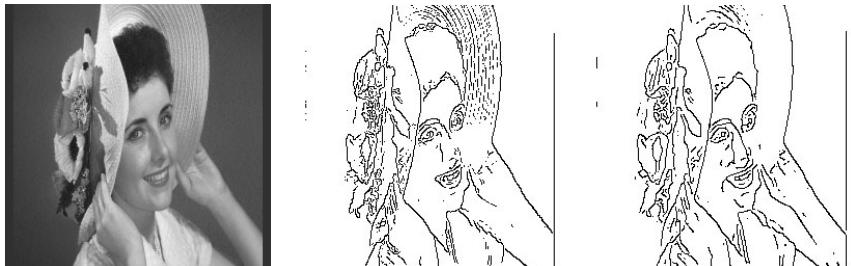


Fig. 5.4 Left: The original image, medial: the edge extracted in $s = 2$, right: the edge extracted in $s = 4$.

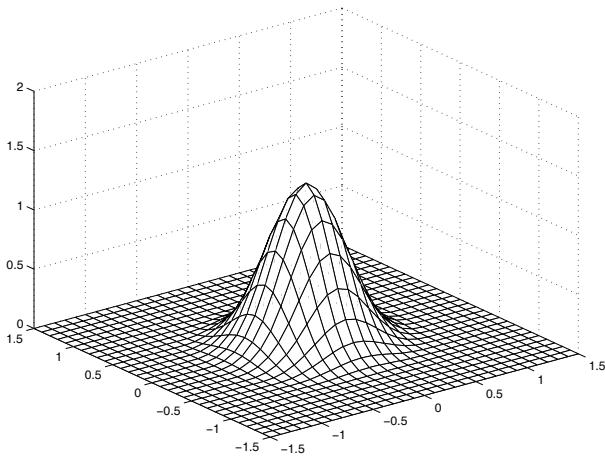


Fig. 5.5 The graphical description of smoothing function $\theta(x, y)$.

5.2 Calculation of $W_s f(x)$ and $W_s f(x, y)$

Calculation of the wavelet transform is a key process in applications. The fast algorithm of computing $W_s f(x)$ for detecting edges and reconstructing the signal can be found in [Mallat and Hwang, 1992] when $\psi(x)$ is a dyadic wavelet defined in that article. However, for some applications such as edge detection, the reconstruction of signals is not required. Therefore, the choice of the wavelet function will not be restricted in the conditions which were presented in [Mallat and Hwang, 1992]. Many wavelets other than dyadic ones can be utilized. In fact, almost all the general integral wavelets satisfy this particular application. It makes much freedom to select the best wavelet ψ for our task. In practice, we should digitize the wavelet transform again by the equi-spaced sampling, and represented by $\{W_s^\psi f(0), W_s^\psi f(1), \dots, W_s^\psi f(n)\}$. In general, different results are produced from the different fitting models. The simplest models are step-function fitting and line-function fitting. They are graphically displayed in Figs. 5.6, 5.7 and 5.8. In the following sub-sections, the wavelet transform will be computed using the step-function fitting.

Of course, the different fitting models and basic wavelets can produce many total different representations of a wavelet transform. The details will be omitted in this book.

Instead of the algorithm stated in [Mallat and Hwang, 1992], we cal-

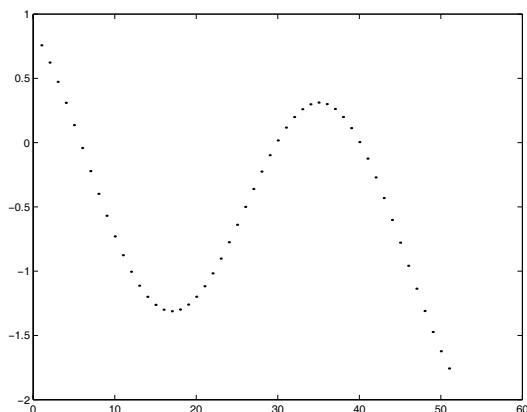


Fig. 5.6 Discrete signal.

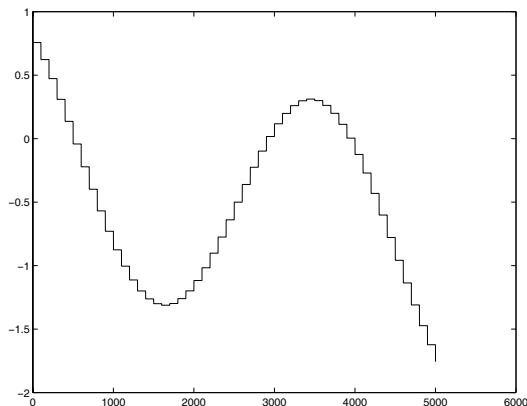


Fig. 5.7 The result fitted by step-function.

culate the corresponding wavelet transforms by the well-known trapezoidal formula of numerical integrals.

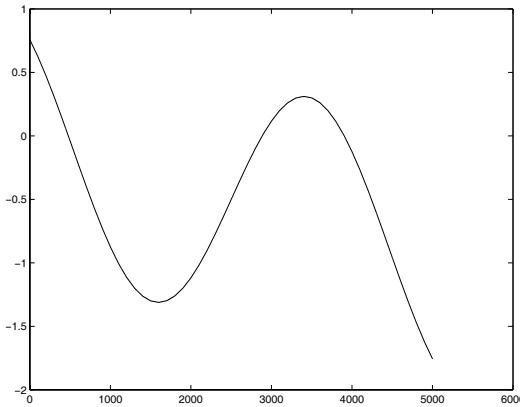


Fig. 5.8 The result from line-function fitting.

5.2.1 Calculation of $W_s f(x)$

The wavelet transform of a one-dimensional signal $f(t)$ can be calculated below:

$$\begin{aligned}
 W_s f(k) &= \int_{-\infty}^{\infty} f(t) \psi_s(k-t) dt \\
 &= \sum_{l \in \mathbb{Z}} \int_{l-1}^l f(l) \psi_s(k-t) dt \\
 &= \sum_{l \in \mathbb{Z}} f(l) \left[\int_{-\infty}^l \psi_s(k-t) dt - \int_{-\infty}^{l-1} \psi_s(k-t) dt \right] \\
 &= \sum_{l \in \mathbb{Z}} f(l) \left[\int_{(k-l)s^{-1}}^{\infty} \psi(t) dt - \int_{(k-l+1)s^{-1}}^{\infty} \psi(t) dt \right] \\
 &= \sum_{l \in \mathbb{Z}} f(l) \int_{(k-l)s^{-1}}^{\infty} \psi(t) dt - \sum_{l \in \mathbb{Z}} f(l+1) \int_{(k-l)s^{-1}}^{\infty} \psi(t) dt \\
 &= \sum_{l \in \mathbb{Z}} [f(l) - f(l+1)] \psi_{k-l}^s,
 \end{aligned} \tag{5.2}$$

where

$$\psi_k^s := \int_{k/s}^{\infty} \psi(t) dt.$$

As an example, let $\psi(x)$ be the quadric spline wavelet which is an odd function. For $x \geq 0$, it can be defined by

$$\psi(x) = \begin{cases} 8(3x^2 - 2x) & x \in [0, 1/2] \\ -8(x-1)^2 & x \in [1/2, 1] \\ 0 & x \geq 1 \end{cases}$$

Since $\psi(x)$ is an odd function, we have

$$\psi_{-k}^s = \int_{-k/s}^{k/s} \psi(t)dt + \int_{k/s}^{\infty} \psi(t)dt = \psi_k^s.$$

Hence, it is enough to calculate ψ_k^s for all non-negative integer k .

- (1) If $k \geq s$, it is obvious that $\psi_k^s = 0$.
- (2) If $\frac{s}{2} \leq k < s$, ψ_k^s becomes

$$\begin{aligned} \psi_k^s &= \int_{k/s}^1 \psi(t)dt \\ &= -8 \int_{k/s}^1 (t-1)^2 dt \\ &= \frac{8}{3} \left(\frac{k}{s} - 1 \right)^3. \end{aligned}$$

- (3) If $0 \leq k < \frac{s}{2}$, the following holds

$$\begin{aligned} \psi_k^s &= \int_{k/s}^1 \psi(t)dt \\ &= 8 \int_{k/s}^{1/2} (3x^2 - 2x)dx - 8 \int_{1/2}^1 (x-1)^2 dx \\ &= 8 \left(\frac{k}{s} \right)^2 \left(1 - \frac{k}{s} \right) - 1 - \frac{8}{3} \left(\frac{1}{2} - 1 \right)^3 \\ &= \frac{1}{3} + 8 \left(\frac{k}{s} \right)^2 \left(1 - \frac{k}{s} \right). \end{aligned}$$

The conclusions can be obtained in accordance with the above discussion, namely

$$\psi_k^s = \begin{cases} \frac{8}{3} \left(\frac{k}{s} - 1\right)^3 & \frac{s}{2} \leq k < s \\ \frac{1}{3} + 8\left(\frac{k}{s}\right)^2 \left(1 - \frac{k}{s}\right) & 0 \leq k < \frac{s}{2} \\ 0 & \text{otherwise} \end{cases}$$

The filtering coefficients $\{\psi_k^s\}$ in Eq. 5.2 can be calculated when ψ is given. If ψ is compactly supported, only finite nonzero items $\{\psi_k^s\}$ will be obtained. Hence, the complexity of this algorithm to compute $\{W_s f(k)\}_{k=0}^N$ is $O(N)$.

In the practical application, it is enough to take $s = 2^j$, $j = 1, 2, 3$. The algorithm presented in this chapter is L^∞ -stable, since

$$\|W_s f\|_{L^\infty(\mathbb{R})} \leq \|f\|_{L^\infty(\mathbb{R})} \|\psi\|_{L^1(\mathbb{R})},$$

where

- $L^1(\mathbb{R})$ denotes the space of all the *Lebesgue* integrable functions on \mathbb{R} and

$$\|\psi\|_{L^1(\mathbb{R})} := \int_{-\infty}^{\infty} |\psi(x)| dx$$

- L^∞ indicates the space of all the essential bounded functions on \mathbb{R} and (see [Rudin, 1973])

$$\|f\|_{L^\infty(\mathbb{R})} := \inf_{|E|=0} \sup_{x \in \mathbb{R} \setminus E} |f(x)|$$

The corresponding filtering coefficients $\{\psi_k^s\}$ in Eq. 5.2 for $j = 1, 2, 3$ are shown in Table 5.1.

5.2.2 Calculation of $W_s f(x, y)$

In two-dimensional signals, the calculation of $W_{2^j}^i f(x, y)$ is more complicated than that of one-dimensional ones. In similar way, we have

$$\begin{aligned} W_s^1 f(n, m) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(u, v) \psi_s^1(n - u, m - v) du dv \\ &= \sum_{k,l} f(k, l) \int \int_{[k, k+1] \times [l, l+1]} \psi_s^1(n - u, m - v) du dv \end{aligned}$$

k	j=1	j=2	j=3
0	-1.3333333731	-1.3333333731	-1.3333333731
1	-0.3333333433	-0.9583333135	-1.2239583731
2		-0.3333333433	-0.9583333135
3		-0.0416666679	-0.6302083135
4			-0.3333333433
5			-0.1406250000
6			-0.0416666679
7			-0.0052083335

Table 5.1 Filtering coefficients $\{\psi_k^{2^j}\}$ and $\{\psi_{k,l}^{2^j,1}\}$ if ψ is a quadric spline wavelet

$$\begin{aligned}
 &= \sum_{k,l} f(k,l) \int \int_{[n-k-1,n-k] \times [m-l-1,m-l]} \psi_s^1(u,v) dudv \\
 &= \sum_{k,l} f(n-1-k, m-1-l) \psi_{k,l}^{s,1}, \tag{5.3}
 \end{aligned}$$

where

$$\begin{aligned}
 \psi_{k,l}^{s,1} &= \int \int_{[k,k+1] \times [l,l+1]} \psi_s^1(u,v) dv \\
 &= \int_{\frac{k}{s}}^{\frac{(k+1)}{s}} du \int_{\frac{l}{s}}^{\frac{(l+1)}{s}} \psi^1(u,v) dv.
 \end{aligned}$$

The problem is, now, to lead to calculation of $\{\psi_{k,l}^{s,1}\}$ in Eq. 5.3. Note that $\psi^1(u,v)$ is odd on u and even on v , we have

$$\begin{aligned}
 \psi_{-k,l}^{s,1} &= \int \int_{[-k,-k+1] \times [l,l+1]} \psi_s^1(u,v) dudv \\
 &= \int_{k-1}^k du \int_l^{l+1} \psi_s^1(u,v) dudv \\
 &= -\psi_{k-1,l}^{s,1}.
 \end{aligned}$$

Similarly, we can deduce that

$$\psi_{k,-l}^{s,1} = \psi_{k,l-1}^{s,1}, \quad \psi_{-k,-l}^{s,1} = -\psi_{k-1,l-1}^{s,1}.$$

Therefore, we need to calculate $\psi_{k,l}^{s,1}$ only for all $k, l \geq 0$.

Let $\psi(r)$ be a odd function, and

$$\psi^1(u, v) = \psi(r) \cos \theta,$$

where

$$r = \sqrt{u^2 + v^2}, \quad \theta = \arctg \frac{v}{u}.$$

Let $\phi(x)$ be

$$\phi(x) = \int_{-\infty}^x \psi(r) dr.$$

Then $\phi(x)$ is an even function, since $\psi(x)$ is an odd function. It is easy to find that

$$\begin{aligned} \frac{\partial}{\partial u} \phi(\sqrt{u^2 + v^2}) &= \phi'(r) \frac{u}{r} = \psi^1(u, v), \\ \frac{\partial}{\partial v} \phi(\sqrt{u^2 + v^2}) &= \phi'(r) \frac{v}{r} = \psi^2(u, v). \end{aligned}$$

Hence

$$\begin{aligned} \psi_{k,l}^{s,1} &= \int_{\frac{k}{s}}^{\frac{(k+1)}{s}} du \int_{\frac{l}{s}}^{\frac{(l+1)}{s}} \frac{\partial}{\partial u} \phi \left(\sqrt{u^2 + v^2} \right) dv \\ &= \int_{\frac{l}{s}}^{\frac{(l+1)}{s}} \left[\phi \left(\sqrt{v^2 + \left(\frac{k+1}{s} \right)^2} \right) - \phi \left(\sqrt{v^2 + \left(\frac{k}{s} \right)^2} \right) \right] dv \\ &= \phi_{l,k+1}^s - \phi_{l+1,k+1}^s - \phi_{l,k}^s + \phi_{l+1,k}^s, \end{aligned}$$

where

$$\phi_{l,k}^s = \int_{\frac{l}{s}}^{\infty} \phi \left(\sqrt{v^2 + \left(\frac{k}{s} \right)^2} \right) dv.$$

As for the calculation of $W_s^2 f(n, m)$, similarly, we have

$$W_s^2 f(n, m) = \sum_{k,l} f(n-1-k, m-1-l) \psi_{k,l}^{s,2},$$

where

$$\psi_{k,l}^{s,2} = \int_{\frac{k}{s}}^{\frac{(k+1)}{s}} du \int_{\frac{l}{s}}^{\frac{(l+1)}{s}} \psi^2(u, v) dv$$

$$\begin{aligned}
&= \int_{\frac{k}{s}}^{\frac{(k+1)}{s}} \left[\phi \left(\sqrt{\left(\frac{l+1}{s} \right)^2 + u^2} \right) - \phi \left(\sqrt{\left(\frac{l}{s} \right)^2 + u^2} \right) \right] du \\
&= \phi_{k,l+1}^s - \phi_{k+1,l+1}^s - \phi_{k,l}^s + \phi_{k+1,l}^s \\
&= \psi_{l,k}^{s,1}.
\end{aligned}$$

Based on the above discussion, we can calculate the filtering coefficients $\{\phi_{k,l}^s\}$ only for all $k, l \geq 0$.

In this chapter, the quadratic spline function will be served as ψ , then

$$\phi(x) = \begin{cases} 8(x^3 - x^2) + \frac{4}{3} & 0 \leq x \leq 1/2; \\ -\frac{8}{3}(x-1)^3 & 1/2 < x \leq 1; \\ 0 & x > 1. \end{cases}$$

- (1) If $k^2 + l^2 \geq s^2$, we have $\phi_{k,l}^s = 0$.
- (2) If $s^2 \geq k^2 + l^2 \geq (\frac{s}{2})^2$, we have

$$\phi_{k,l}^s = -\frac{8}{3} \int_{\frac{k}{s}}^{\sqrt{1-(\frac{l}{s})^2}} \left(\sqrt{u^2 + \left(\frac{l}{s} \right)^2} - 1 \right)^3 du.$$

- (3) If $k^2 + l^2 \leq (\frac{s}{2})^2$, then

$$\begin{aligned}
\phi_{k,l}^s &= \int_{\frac{k}{s}}^{\sqrt{\frac{1}{4} - (\frac{l}{s})^2}} \left[8 \left(\left(u^2 + \left(\frac{l}{s} \right)^2 \right)^{\frac{3}{2}} - \left(u^2 + \left(\frac{l}{s} \right)^2 \right) \right) + \frac{4}{3} \right] du \\
&\quad - \frac{8}{3} \int_{\sqrt{\frac{1}{4} - (\frac{l}{s})^2}}^{\sqrt{1 - (\frac{l}{s})^2}} \left(\sqrt{u^2 + \left(\frac{l}{s} \right)^2} - 1 \right)^3 du. \tag{5.4}
\end{aligned}$$

In order to simplify the calculation, the following replacements will be done:

$$\begin{aligned}
\frac{k}{s} &\implies a, \\
\sqrt{\frac{1}{4} - (\frac{l}{s})^2} &\implies b, \\
\sqrt{\frac{1}{4} - \left(\frac{l}{s} \right)^2} &\implies c,
\end{aligned}$$

$$\begin{aligned}\sqrt{1 - \left(\frac{l}{s}\right)^2} &\implies d, \\ \frac{l}{s} &\implies t.\end{aligned}$$

Thus, Eq. 5.4 becomes

$$\phi_{k,l}^s = \int_a^b [8(u^2 + t^2)(\sqrt{u^2 + t^2} - 1) + \frac{4}{3}] du - \frac{8}{3} \int_c^d (\sqrt{u^2 + t^2} - 1)^3 du.$$

We denote

$$\begin{aligned}I(t, a, b) &= \int_a^b \left[8(u^2 + t^2)(\sqrt{u^2 + t^2} - 1) + \frac{4}{3} \right] du, \\ J(t, c, d) &= -\frac{8}{3} \int_c^d (\sqrt{u^2 + t^2} - 1)^3 du.\end{aligned}\quad (5.5)$$

To avoid the complicated calculation, a software called *Mathematica* is utilized to solve Eq. 5.5, and the following results can be achieved:

$$\begin{aligned}I(t, a, b) &= \frac{8}{3}(a^3 - b^3) - \frac{4}{3}(a - b)[1 - 6t^2] \\ &\quad - \sqrt{a^2 + t^2}[2a^3 + 5at^2] + \sqrt{b^2 + t^2}(2b^3 + 5bt^2) \\ &\quad - 3t^4 \log \frac{a + \sqrt{a^2 + t^2}}{b + \sqrt{b^2 + t^2}}, \\ J(t, c, d) &= -\frac{8}{3}(c - d)(c^2 + d^2 + cd + 3t^2 + 1) \\ &\quad + \frac{1}{3} \left[c(2c^2 + 5t^2 + 12)\sqrt{c^2 + t^2} - d(2d^2 + 5t^2 + 12)\sqrt{d^2 + t^2} \right] \\ &\quad + t^2(4 + t^2) \log \frac{c + \sqrt{c^2 + t^2}}{d + \sqrt{d^2 + t^2}}.\end{aligned}$$

Hence, for $\forall k, l \geq 0$, we have

$$\phi_{k,l}^s = \begin{cases} 0 & \text{if } k^2 + l^2 \geq s^2; \\ J\left(\frac{l}{s}, \frac{k}{s}, \sqrt{1 - \left(\frac{l}{s}\right)^2}\right) & \text{if } s^2 \geq k^2 + l^2 \geq \frac{1}{4}s^2; \\ I\left(\frac{l}{s}, \frac{k}{s}, \sqrt{\frac{1}{4} - \left(\frac{l}{s}\right)^2}\right) & \text{if } k^2 + l^2 \leq \frac{1}{4}s^2. \\ +J\left(\frac{l}{s}, \sqrt{\frac{1}{4} - \left(\frac{l}{s}\right)^2}, \sqrt{1 - \left(\frac{l}{s}\right)^2}\right) & \text{if } k^2 + l^2 \leq \frac{1}{4}s^2. \end{cases}\quad (5.6)$$

For $s = 2^j, j = 1, 2, 3$, the corresponding filtering coefficients $\{\psi_{k,l}^{s,1}\}$ in Eq. 5.3 are shown in Tables 5.2–5.4. For the given scale s , the complexity of computing $\{W_s^i f(n, m)\}_{n,m=0}^N$ is $O(N^2)$.

$l \setminus k$	$k=0$	$k=1$
$l=0$	-0.3485791385	-0.0351441652
$l=1$	-0.1097541898	-0.0085225009

Table 5.2 Filtering coefficients $\{\psi_{k,l}^{2^j,1}\}, (j = 1)$ if $\psi(r)$ is a quadic spline wavelet

$l \setminus k$	$k=0$	$k=1$
$l=0$	-0.0838140026	-0.0472041145
$l=1$	-0.1416904181	-0.0958706033
$l=2$	-0.0651057437	-0.0327749476
$l=3$	-0.0088690016	-0.0030044997
$l \setminus k$	$k=2$	$k=3$
$l=0$	-0.0129809398	-0.0012592132
$l=1$	-0.0196341425	-0.0012698699
$l=2$	-0.0063385521	-0.0000750836
$l=3$	-0.0001088651	

Table 5.3 Filtering coefficients $\{\psi_{k,l}^{2^j,1}\}, (j = 2)$ if $\psi(r)$ is a quadic spline wavelet

5.3 Wavelet Transform for Contour Extraction and Background Removal

In this section, the singularities will be studied by Lipschitz exponent. According to the mathematical analysis of the different geometric structures of the singularities using Lipschitz exponent. The singularities can be categorized into three basic geometric structures, namely, step-structure, roof-structure, and Dirac-structure. The characterization of these structures with wavelet transform is also studied in this section. A significant property will be derived, i.e. the wavelet transform of the step-structure of singularity is a non-zero constant which is independent on both the gradient direction and the scale of the wavelet transform. This property leads

$l \setminus k$	$k=0$	$k=1$
$l=0$	-0.0130508747	-0.0107604843
$l=1$	-0.0322880745	-0.0277145673
$l=2$	-0.0400620662	-0.0347912423
$l=3$	-0.0361453630	-0.0306917503
$l=4$	-0.0233931188	-0.0196621399
$l=5$	-0.0120095834	-0.0100409016
$l=6$	-0.0043909089	-0.0034857604
$l=7$	-0.0006066251	-0.0003857070
$l \setminus k$	$k=2$	$k=3$
$l=0$	-0.0080117621	-0.0051634177
$l=1$	-0.0208740719	-0.0181548643
$l=2$	-0.0258209687	-0.0154902851
$l=3$	-0.0216873437	-0.0128720086
$l=4$	-0.0139309634	-0.0082158195
$l=5$	-0.0069123111	-0.0037158530
$l=6$	-0.0020833646	-0.0007867273
$l=7$	-0.0001258312	-0.0000085766
$l \setminus k$	$k=4$	$k=5$
$l=0$	-0.0025986142	-0.0010915556
$l=1$	-0.0065528429	-0.0027379275
$l=2$	-0.0097385549	-0.0031414898
$l=3$	-0.0063897478	-0.0023643500
$l=4$	-0.0037783380	-0.0010769776
$l=5$	-0.0013164375	-0.0001667994
$l=6$	-0.0001085378	-0.0000003273
$l=7$		
$l \setminus k$	$k=4$	$k=7$
$l=0$	-0.0003376677	-0.0000403965
$l=1$	-0.0008041780	-0.0000769711
$l=2$	-0.0008010443	-0.0000415404
$l=3$	-0.0004234327	-0.0000038524
$l=4$	-0.0000748098	
$l=5$	-0.0000002738	
$l=6$		
$l=7$		

Table 5.4 Filtering coefficients $\{\psi_{k,l}^{2^j,1}\}$, ($j = 3$) if $\psi(r)$ is a quadric spline wavelet

to provide a simple and direct strategy for detecting this specific structure of singularities, including contours of the patterns. Thereafter, a novel algorithm which is referred to as scale-independent algorithm will be developed, and the modular-angle-separated (MAS) wavelet transform will be applied. As the applications of this new method, many examples will be presented, which deal with the data preprocessing in the recognition of the two-dimensional object and document processing.

Edges are typical singularities in images, meanwhile, contour is a specific edge, which belongs to the step-structure. Extraction of singularities, specifically, extraction of contours, plays an important role in the data preprocessing phase of pattern recognition including document processing. Research on extracting the contours is one of the most significant topics in this discipline. Many methods have been developed to analyze the properties of contours and detect them from various images. For example, some works can be found in [Chen and Yang, 1995; Deng and Lyengar, 1996; Law et al., 1996; Matalas et al., 1997; Tang et al., 1997a; Tang et al., 1997c; Thune et al., 1997].

The subjects of Lipschitz exponents and wavelet transform are remarkable mathematical tools to analyze the singularities including the edges, and further, to detect them effectively. A significant study related to these research topics has been done by Mallat, Hwang and Zhong, and published in *IEEE Trans. on Information Theory* [Mallat and Hwang, 1992], and *IEEE Trans. on Pattern Analysis and Machine Intelligence* [Mallat and Zhong, 1992]. Many important contributions have been made in these papers. It has been shown that the local maxima of the wavelet transform modulus can provide enough information for analyzing the singularities, and can detect *all* singularities. However, it may not identify different structures of singularities. This chapter will present a new method which can handle this problem. Now, we look at an example. We consider an image which is shown in Fig. 5.9(a). Two classes of singularities are embedded in this image: (i) a contour of the aircraft which belongs to the step-structure singularities, and (ii) some lines and texts which belong to the Dirac-structure singularities. Such an image is referred to as *multi-structure-singularities image*. A particular task is that we are required to extract the contour of the aircraft, and to remove all lines and texts. What result will arrive when the local maxima of the wavelet transform modulus is applied?

Look at the following experiments: Fig. 5.9(b) displays the modulus image, where the black pixels indicate zero values and the white ones cor-

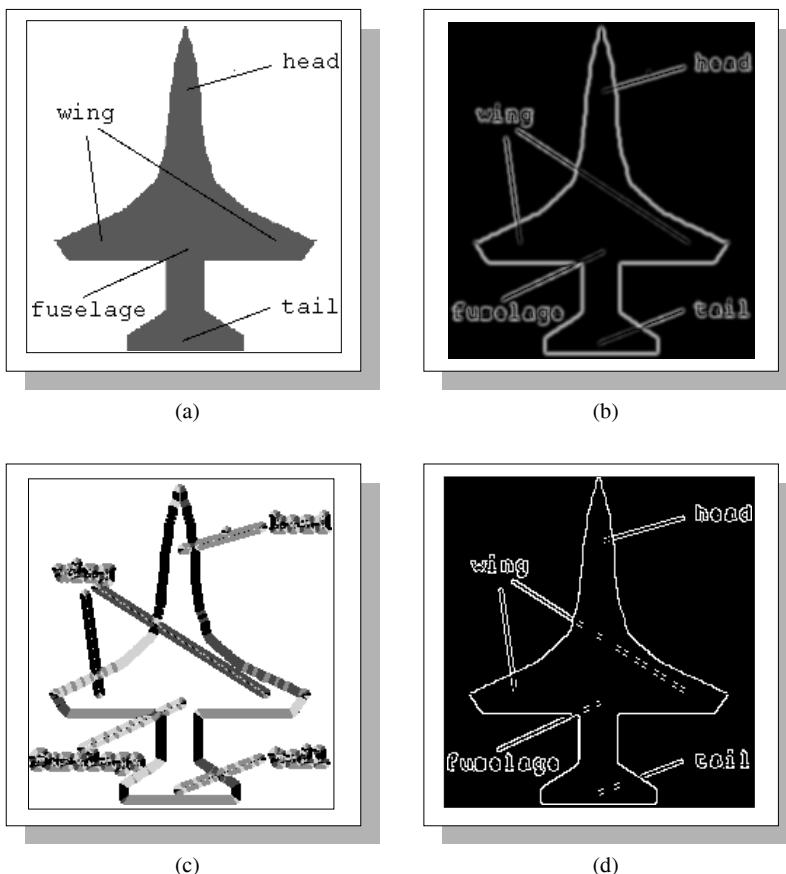


Fig. 5.9 Result of edge detection using the wavelet transform modulus maxima in data preprocessing phase of pattern recognition.

respond to the highest values. Fig. 5.9(c) gives the angle image, in which the angle values range from 0 (black) to 2π (white). Fig. 5.9(d) shows the modulus maxima. It is clear, from this result, that the method of the modulus maxima has detected all singularities without recognizing two different structures of singularities. Thus, the resulting image contains lines and texts which are required to be deleted from the image. Another example can be found in Figs. 5.10 and 5.11. A mailing address which was written on a paper with guide lines is shown in Fig. 5.10(a). In order to automatically process it by computers, a preprocessing is to remove the guide lines

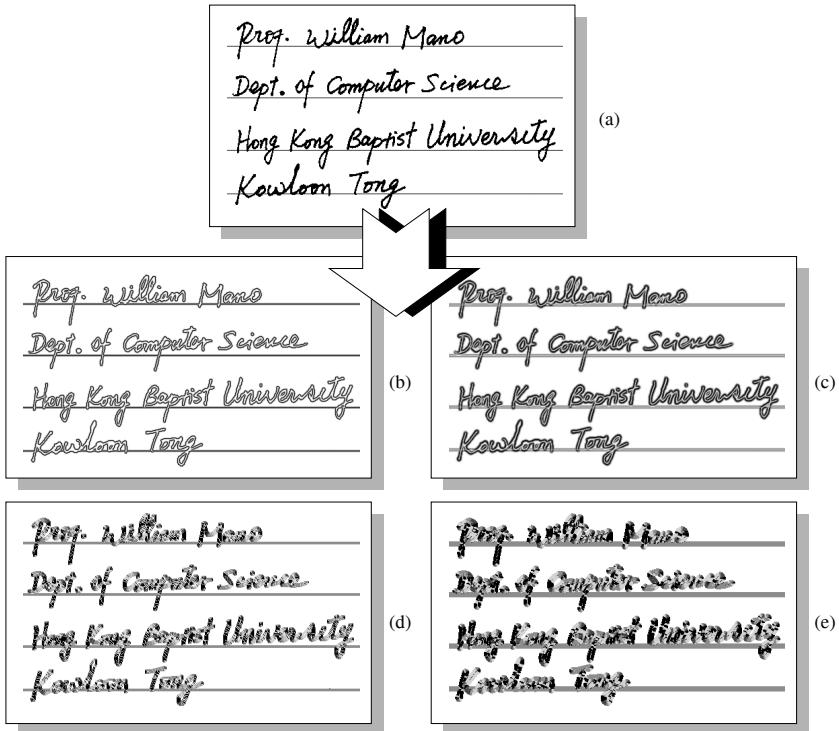


Fig. 5.10 Wavelet transform of a page of document with different scales ($s = 2, 4$).

and extract the contours of the handwriting English letters. Figs. 5.10(b) and (c) display the modulus images with scales of $s = 2, 4$ of the wavelet transform respectively. Figs. 5.10(d) and (e) state the angle images with transform scales of $s = 2, 4$ respectively. Obviously, it is still the same problem when different scales of the wavelet transforms are used, and the results are given in Figs. 5.11(a) and (b).

In order to improve the method proposed in [Mallat and Hwang, 1992; Mallat and Zhong, 1992], so that it can be used to detect the contour of an object which is a typical step-structure singularity and eliminate others, this section will carefully study the wavelet transform with respect to three basic geometric structures of singularities. A very important property will be proven that the wavelet transform of a step singularity is a non-zero constant which is independent on the scale of the wavelet transform. Based

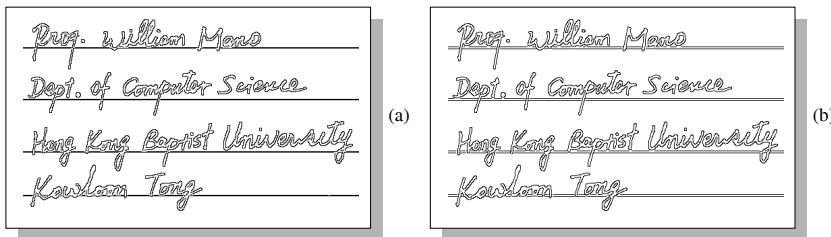


Fig. 5.11 Result of edge detection using the wavelet transform modulus maxima in document processing, (a) $s = 2$, (b) $s = 4$.

on this result, a scale-independent wavelet algorithm will be developed. This novel method can detect step-structure singularities from the multi-structure-singularity images.

This section is organized into the following sub-sections:

- (1) Basic Edge Structures
- (2) Analysis of the Basic Edge Structures with Wavelet Transform
- (3) Scale-Independent Algorithm
- (4) Experiments

5.3.1 Basic Edge Structures

The edge is one class of singularities in the one- and two-dimensional signals. Lipschitz exponent is a remarkable mathematical tool to analyze the singularities including the edges. According to the analysis with Lipschitz exponents, the edges can be categorized into three basic geometric structures [Tang et al., 1998e].

Let $0 \leq \alpha \leq 1$, function $f(x)$ is called uniformly Lipschitz α over (a, b) if there exists a constant K such that $\forall x_1, x_2 \in (a, b)$, we have

$$|f(x_1) - f(x_2)| \leq K|x_1 - x_2|^\alpha, \quad |x_1 - x_2| \rightarrow 0. \quad (5.7)$$

where, α denotes the regularity exponent. The value of α plays a key role in analyzing the singularity or regularity of a signal. We are interested in the case of $0 \leq \alpha \leq 1$, which corresponds to the different geometric structures of edges. The geometric structures of the edges can be analyzed by the value of α . The smaller the α is, the weaker is the regularity (i.e., smoothness), or the smaller the α , the stronger the singularity. Lipschitz exponent α can be extended to the range of $-1 \leq \alpha < 0$. $f(x)$ is called uniformly

Lipschitz α if its primitive function is uniformly Lipschitz $\alpha + 1$ ($-1 \leq \alpha < 0$). This extension permits $f(x)$ to be a distribution (i.e., a generalized function). For instance, when $\alpha = -1$, the well-known *Dirac* function $\delta(x)$ is a uniformly Lipschitz -1 in any interval containing 0. The theory of distribution can be referred to [Rudin, 1973]. In summary, three special values of α will be considered in the concrete applications, namely,

- (1) $\alpha = 0$: This corresponds to the edges with step-structure;
- (2) $\alpha = 1$: It corresponds to the edges with roof-structure.
- (3) $\alpha = -1$: It corresponds to the edges with Dirac-structure.

Any of the three structures of edges can be regarded as a superposition of an ideal edge $e(x)$, and a additional signal $f_0(x)$:

$$f(x) = e(x) + f_0(x).$$

It will be proved in the next subsection that the later is not important in this study. We, thus, will pay the attention to the former, ideal versions, which can be formulated below:

- (1). Ideal step edge

$$e(x) = \begin{cases} 1 & x \geq x_0; \\ 0 & x < x_0. \end{cases}$$

- (2). Ideal roof edge

$$e(x) = \begin{cases} m(x - x_0) + c_0 & x_0 - h_1 < x \leq x_0; \\ n(x - x_0) + c_0 & x_0 < x < x_0 + h_2. \end{cases}$$

- (3). Ideal Dirac edge

$$e(x) = \begin{cases} \infty & x = x_0; \\ 0 & x \neq x_0. \end{cases}$$

Similarly, the Lipschitz exponent of regularity for the two-dimensional images can be defined below:

Definition 5.1 Let $0 \leq \alpha \leq 1$, function $f(u, v)$ is called uniformly *Lipschitz* α on the interval of $(a, b) \times (c, d)$, if there exists a constant K , such that, $\forall (u_1, v_1) \in (a, b) \times (c, d)$, and

$$|f(u_1, v_1) - f(u_2, v_2)| \leq K(|u_1 - u_2|^2 + |v_1 - v_2|^2)^{\alpha/2}, \quad (5.8)$$

$$|u_1 - u_2| \rightarrow 0, \quad |v_1 - v_2| \rightarrow 0.$$

In 2-D signal, Lipschitz exponent α can be generalized to $-1 \leq \alpha < 0$ similarly to 1-D case. Tempered distribution $f(x, y)$ is an example in this generalization. Thus, $f(x, y)$ is called uniformly Lipschitz α over $(a, b) \times (c, d)$, if the primitive function of the distribution $f(x, y)$ is uniform Lipschitz $\alpha + 1$ ($-1 \leq \alpha < 0$) on $(a, b) \times (c, d)$.

The graphic descriptions of the basic structures of the edges are illustrated in Figs. 5.12, 5.13 and 5.14 respectively.

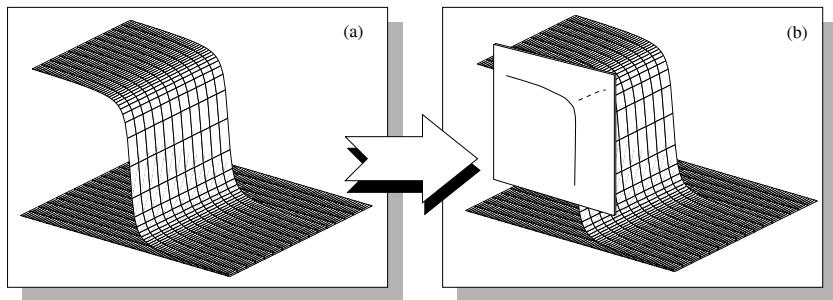


Fig. 5.12 Graphic description of the step-structure edge

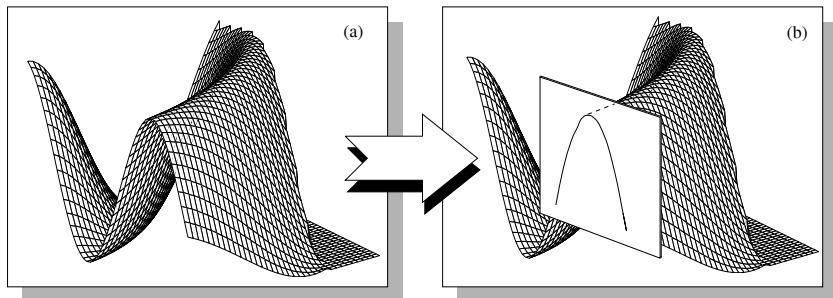


Fig. 5.13 Graphic description of the roof-structure edge

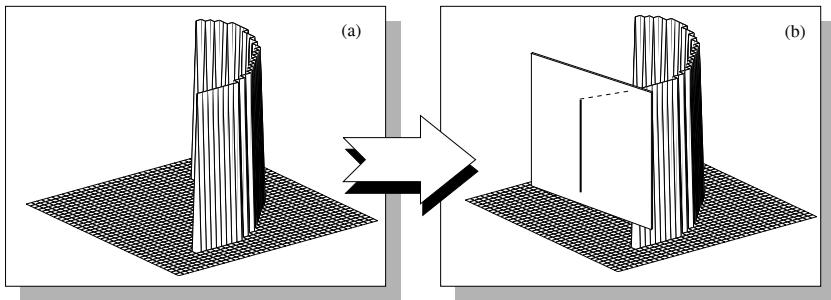


Fig. 5.14 Graphic description of the Dirac-structure edge

5.3.2 Analysis of the Basic Edge Structures with Wavelet Transform

A remarkable property of the wavelet transform is its ability to characterize the local regularity of functions. After characterizing the singularities with Lipschitz exponents and classifying the edges into the basic structures, the analysis of these structures with wavelet transform will be presented in this subsection.

Let $L^2(\mathbb{R})$ be a *Hilbert* space of all the square integrable functions on \mathbb{R} , and $\psi \in L^2(\mathbb{R})$ be a wavelet function, i.e., $\psi(x)$ decreases fast when x goes to infinity and satisfies

$$\int_{\mathbb{R}} \psi(x) dx = 0.$$

In the edge detection, the speed of the decrease of $\psi(x)$ effects the result of local detection. Usually, the faster of the speed is, the better is the result. In this chapter, a compactly supported wavelet function is chosen as $\psi(x)$. $\forall f(x) \in L^2(\mathbb{R})$ and $s > 0$, the wavelet transform of $f(x)$ with the scale s is defined as follows:

$$W_s f(x) := (f * \psi_s)(x) = \int_{\mathbb{R}} f(t) \frac{1}{s} \psi\left(\frac{x-t}{s}\right) dt,$$

where $*$ denotes the convolution operation, and $\psi_s(x) := \frac{1}{s} \psi\left(\frac{x}{s}\right)$. Particularly, $W_{2^j} f(x)$, ($j \in \mathbb{Z}$) is called dyadic wavelet transform, where \mathbb{Z} stands for the set of all integers.

The local Lipschitz singularity can be characterized with wavelet transform by the following Theorem [Meyer, 1990]:

Theorem 5.1 Let α be a real number, $\alpha \geq -1$, function $f(x)$ is said to be uniformly Lipschitz α , if and only if there exists a constant K such that

$$|W_s f(x)| \leq K s^\alpha, \quad \forall x \in (a, b). \quad (5.9)$$

This theorem is a theoretical basis for wavelet-based approach to the edge detection. Unfortunately, it gives only an inequality, meanwhile the constant K is unknown. Hence, it is difficult to apply this result to detect edges directly in practice. To contest such a deficiency, we will discuss the properties of wavelet transform with respect to the three basic structures of edges, thereafter, find a solution to handle some specific edges. After applying wavelet transform to these edges, the following main results can be obtained:

Theorem 5.2 Let $\psi(x)$ be an odd wavelet function, which is supported on $[-\delta, \delta]$, satisfying $\int_0^\infty \psi(x) dx \neq 0$, and x_0 be the coordinate of an ideal edge.

(1). The wavelet transform of a step edge is described by

$$W_s e(x) = \int_{-\infty}^{-|x-x_0|s^{-1}} \psi(t) dt. \quad (5.10)$$

(2). The wavelet transform of a Dirac edge can be obtained as follows:

$$W_s e(x) = \frac{1}{s} \psi\left(\frac{x-x_0}{s}\right). \quad (5.11)$$

(3). If $|x - x_0| + s\delta \leq h$, the wavelet transform of a roof edge can be computed in the following cases:

(i). if $|x - x_0| \leq s\delta$, then

$$\begin{aligned} W_s e(x) &= (x - x_0)(m - n) \int_{|x-x_0|s^{-1}}^{\infty} \psi(t) dt \\ &\quad -s(m + n) \int_{|x-x_0|s^{-1}}^{\infty} t\psi(t) dt \\ &\quad -s \int_{-|x-x_0|s^{-1}}^{|x-x_0|s^{-1}} t\psi(t) dt \cdot \begin{cases} n & \text{if } x \geq x_0; \\ m & \text{if } x < x_0. \end{cases} \quad (5.12) \end{aligned}$$

(ii). if $x - x_0 > s\delta$, then

$$W_s e(x) = -2ns \int_0^\infty t\psi(t)dt. \quad (5.13)$$

(iii). if $x - x_0 > -s\delta$, then

$$W_s e(x) = -2ms \int_0^\infty t\psi(t)dt. \quad (5.14)$$

Particularly, at point x_0 , the wavelet transforms of these basic edge structures become

$$W_s e(x_0) = \begin{cases} - \int_0^\infty \psi(t)dt & \text{if } x_0 \text{ is a step edge;} \\ 0 & \text{if } x_0 \text{ is a Dirac edge;} \\ -s(m+n) \int_0^\infty t\psi(t)dt & \text{if } x_0 \text{ is a roof edge.} \end{cases} \quad (5.15)$$

Proof. (1). If x_0 is the coordinate of a step edge, the wavelet transform becomes

$$\begin{aligned} W_s f(x) &= \int f(t)\psi_s(x-t)dt \\ &= \int_{x_0}^\infty \psi_s(x-t)dt \\ &= \int_{-\infty}^{(x-x_0)s^{-1}} \psi(t)dt \\ &= \int_{-\infty}^{-|x-x_0|s^{-1}} \psi(t)dt + \int_{-|x-x_0|s^{-1}}^{(x-x_0)s^{-1}} \psi(t)dt \\ &= \int_{-\infty}^{-|x-x_0|s^{-1}} \psi(t)dt. \end{aligned}$$

(2). If x_0 is the coordinate of a Dirac edge, the wavelet transform can be produced by

$$\begin{aligned} W_s f(x) &= (f * \psi_s)(x) \\ &= \psi_s(x - x_0) \\ &= \frac{1}{s} \psi\left(\frac{x - x_0}{s}\right). \end{aligned}$$

(3). If x_0 is the coordinate of a roof edge, it is easy to find that $x - st \in [x_0 - h, x_0 + h]$ for $\forall x : |x - x_0| \leq h - s\delta$. Therefore, we have (i). If

$|x - x_0| \leq \delta s$, i.e. $-\delta \leq (x - x_0)s^{-1} \leq \delta$, the wavelet transform is computed below:

$$\begin{aligned}
W_s e(x) &= \int_{-\infty}^{\infty} e(x-t)\psi_s(t)dt \\
&= \int_{-\delta}^{(x-x_0)s^{-1}} e(x-st)\psi(t)dt + \int_{(x-x_0)s^{-1}}^{\delta} e(x-st)\psi(t)dt \\
&= \int_{-\delta}^{(x-x_0)s^{-1}} [n(x - x_0 - st) + c_0]\psi(t)dt \\
&\quad + \int_{(x-x_0)s^{-1}}^{\delta} [m(x - x_0 - st) + c_0]\psi(t)dt \\
&= n(x - x_0) \int_{-\delta}^{(x-x_0)s^{-1}} \psi(t)dt - ns \int_{-\delta}^{(x-x_0)s^{-1}} t\psi(t)dt \\
&\quad + m(x - x_0) \int_{(x-x_0)s^{-1}}^{\delta} \psi(t)dt - ms \int_{(x-x_0)s^{-1}}^{\delta} t\psi(t)dt \\
&= (m - n)(x - x_0) \int_{(x-x_0)s^{-1}}^{\delta} \psi(t)dt - ns \int_{-\delta}^{-|x-x_0|s^{-1}} t\psi(t)dt \\
&\quad - ns \int_{-|x-x_0|s^{-1}}^{(x-x_0)s^{-1}} t\psi(t)dt - ms \int_{(x-x_0)s^{-1}}^{|x-x_0|s^{-1}} t\psi(t)dt \\
&\quad - ms \int_{|x-x_0|s^{-1}}^{\delta} t\psi(t)dt \\
&= (m - n)(x - x_0) \int_{|x-x_0|s^{-1}}^{\delta} \psi(t)dt - (m + n)s \int_{|x-x_0|s^{-1}}^{\delta} t\psi(t)dt \\
&\quad - s \left(\int_{-|x-x_0|s^{-1}}^{|x-x_0|s^{-1}} t\psi(t)dt \right) \cdot \begin{cases} n & \text{if } x \geq x_0 \\ m & \text{if } x < x_0 \end{cases} \\
&= (m - n)(x - x_0) \int_{|x-x_0|s^{-1}}^{\infty} \psi(t)dt - (m + n)s \int_{|x-x_0|s^{-1}}^{\infty} t\psi(t)dt \\
&\quad - s \left(\int_{-|x-x_0|s^{-1}}^{|x-x_0|s^{-1}} t\psi(t)dt \right) \cdot \begin{cases} n & \text{if } x \geq x_0; \\ m & \text{if } x < x_0. \end{cases}
\end{aligned}$$

(ii). If $x - x_0 > s\delta$, then

$$W_s e(x) = \int_{-\delta}^{\delta} e(x-st)\psi(t)dt$$

$$\begin{aligned}
&= \int_{-\delta}^{\delta} [n(x - st - x_0) + c_0] \psi(t) dt \\
&= -ns \int_{-\delta}^{\delta} t \psi(t) dt \\
&= -2ns \int_0^{\infty} t \psi(t) dt.
\end{aligned}$$

(iii). If $x - x_0 < s\delta$, then

$$\begin{aligned}
W_s e(x) &= \int_{-\delta}^{\delta} e(x - st) \psi(t) dt \\
&= \int_{-\delta}^{\delta} [m(x - st - x_0) + c_0] \psi(t) dt \\
&= -2ms \int_0^{\infty} t \psi(t) dt.
\end{aligned}$$

It is easy to understand the following facts from Eq. 5.15 and the above analyses:

- The wavelet transform $W_s e(x_0)$ of the step-structure edge is a nonzero constant which is independent on the scale of the wavelet transform. It is sign-preserving at the both sides of the neighborhood of x_0 and the extremum is reached at x_0 .
- For the Dirac-structure edge x_0 , its wavelet transform becomes $W_s e(x_0) = 0$ and $W_s e(x)$ has the opposite signs when x belongs to the left or right side of the neighborhood of x_0 . Moreover, the extrema are located at both sides of the neighborhood of x_0 . When the local extremum is used to detect the Dirac edge, the ambiguity may appear, since two extrema are obtained.
- With respect to the roof-structure edge x_0 , if $m = -n$, it can be concluded that

$$W_s e(x_0) = 0,$$

$$W_s e(x) \approx 2m(x - x_0) \int_{s^{-1}|x-x_0|}^{\infty} \psi(t) dt.$$

According to the above results, one can deduce that the sign of $W_s e(x)$ at the left side of the neighborhood of x_0 is opposite to

that at the right side. It can easily be found that if $m \neq -n$,

$$W_s e(x_0) = -s(m+n) \int_0^\infty t\psi(t)dt,$$

which is dependent on the scale s .

From the above theorem, the following conclusion will be achieved,

Corollary 5.1 x_0 is the coordinate of a step edge, if and only if the wavelet transform of $e(x)$, $W_s e(x_0)$, is a nonzero constant which is independent on the scale of the wavelet transform.

This significant property will be utilized to extract the step-structure edges from a signal with multi-structure edges. The details will be presented in the next subsection.

Similarly, for the two-dimensional images, the characterization of the basic geometric structures of the edges with wavelet transform can also be established. However, it is more complicated to do so, because the two-dimensional image has multi-directions and therefore two wavelet functions have to be applied. In this chapter, let us consider a special class of wavelet functions, $\psi^1(u, v)$ and $\psi^2(u, v)$, whose moduli and angles can be separated, such that

$$\psi^1(u, v) = \psi(r)k_1(\theta), \quad \psi^2(u, v) = \psi(r)k_2(\theta),$$

where $\psi(r)$ stands for the function with respect to the modulus $r = \sqrt{u^2 + v^2}$; and $k_1(\theta)$ and $k_2(\theta)$ denote the functions with respect to the angle $\theta = \arctan(\frac{v}{u})$. They are referred to as *modulus-angle-separated wavelets (MASW)*.

In the remainder of this subsection, we will prove that the modulus-angle-separated wavelets possess two significant properties, namely,

- The wavelet transforms of the edges are independent on the gradient directions of edges;
- The wavelet transform of the step-structure edge is independent on the scales of the transform.

The first characteristic provides an ability to detect edges in different directions, while the second one can be devoted to identify the step-structure edges from others. These are important in the concrete applications.

The two-dimensional MASW transform is defined as follows.

$$W_s^i f(x, y) = \int \int f(x - u, y - v) \psi_s^i(u, v) du dv, \quad (i = 1, 2),$$

where

$$\psi_s^i(u, v) = \frac{1}{s^2} \psi^i\left(\frac{u}{s}, \frac{v}{s}\right) \quad (i = 1, 2).$$

Particularly, if $s = 2^j$ ($j \in \mathbb{Z}$), they are called the dyadic wavelets.

It is easy to show that both ψ^1 and ψ^2 are two-dimensional wavelets, if $\int_0^{2\pi} k_1(\theta) d\theta = \int_0^{2\pi} k_2(\theta) d\theta = 0$. To simplify the derivation without losing the generality, a special case will be applied in this study. In this way, $k_1(\theta)$ and $k_2(\theta)$ now become

$$k_1(\theta) = \cos \theta, \quad k_2(\theta) = \sin \theta.$$

Let

$$\nabla W_s f(x, y) := \begin{pmatrix} W_s^1 f(x, y) \\ W_s^2 f(x, y) \end{pmatrix},$$

and

$$|\nabla W_s f(x, y)| := \sqrt{|W_s^1 f(x, y)|^2 + |W_s^2 f(x, y)|^2},$$

Similarly, Theorem 5.1 can be extended to the two-dimensional images [Mallat and Hwang, 1992], and presented below.

Theorem 5.3 Let α be a real number, $0 < \alpha < 1$, then function $f(x, y)$ is uniformly Lipschitz α over $(a, b) \times (c, d)$, if and only if there exists a constant K such that, $\forall (x, y) \in (a, b) \times (c, d)$, we have

$$|\nabla W_s f(x)| \leq K s^\alpha. \quad (5.16)$$

For two-dimensional images, the geometric structures of edges are much more complicated than that of the one-dimensional signals. Nevertheless, the step- and Dirac-structures of edges, in practice, are considered as contours in images. Thus, we will emphasize to discuss these two structures in this subsection. The edge shown in Fig. 5.15(a) is said to be an ideal step-structure edge, and contours of patterns are typical examples of it. The edge illustrated in Fig. 5.15(b) is called an ideal curve-Dirac edge which consists of many impulses at the gradient direction. The drawing lines in

the image are the examples of this edge. Another type of Dirac edge is shown in Fig. 5.15(c) which is called an ideal single-point-Dirac edge, and it possesses an impulse at any direction. It is a special case of the curve-Dirac edges, and write noise belongs to this edge. The edge can be regarded

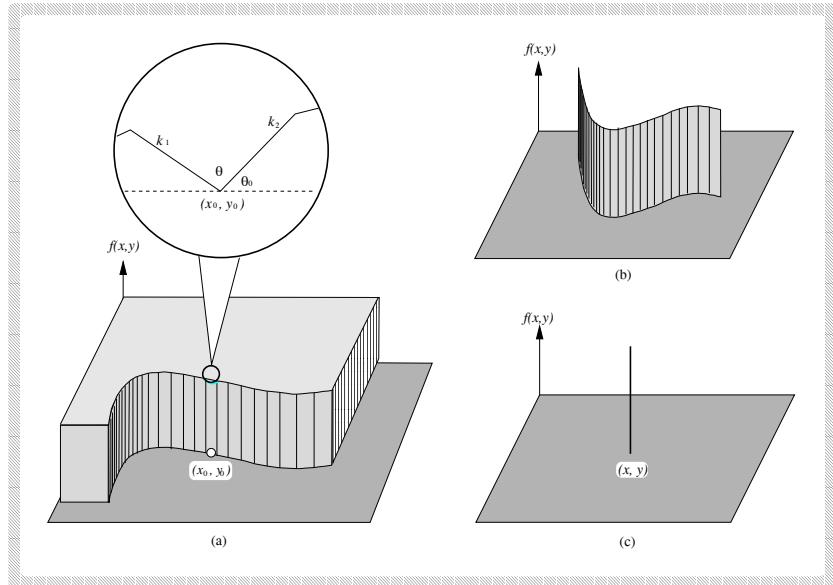


Fig. 5.15 Two dimensional ideal edges

as a polygonal function, which consists of many segments. Suppose θ is an angle between any two adjacent segments as shown in Fig. 5.15(a). If angle $\theta = \pi$, these two segments become a straight line on the edge. The main result in this subsection is as follows:

Theorem 5.4 Let $\psi(r)$ be a compactly supported function on the interval of $[0, \infty)$. Then $\psi^1(u, v)$ and $\psi^2(u, v)$ are two 2-D MASW with the form of

$$\psi^1(u, v) = \psi(r) \cos \theta, \quad \psi^2(u, v) = \psi(r) \sin \theta, \quad (5.17)$$

where $r = \sqrt{u^2 + v^2}$, and $\theta = \arctan \frac{v}{u}$. Furthermore, we have

- (1) if (x_0, y_0) is the coordinate of a point on the ideal step edge with

angle θ , $0 < \theta < 2\pi$, then

$$|\nabla W_s e(x_0, y_0)| = 2 \left| \sin \frac{\theta}{2} \right| \cdot \left| \int_0^\infty r\psi(r)dr \right|. \quad (5.18)$$

(2) if (x_0, y_0) is the coordinate of a point on the Dirac edge, then

$$|\nabla W_s e(x_0, y_0)| = 0. \quad (5.19)$$

(3) if (x_0, y_0) is the coordinate of a point on the roof edge, then

$$|\nabla W_s e(x_0, y_0)| \rightarrow 0, \quad (s \rightarrow 0). \quad (5.20)$$

Proof. (1) First, the step-structure edges are considered. Suppose that two polygonal lines shown in Fig. 5.15(a) are met at the coordinate (x_0, y_0) . Let θ be the angle between these two lines, k_1 and k_2 be the slopes of them, and θ_0 be the angle between x-axis and one side. It can be computed that

$$\begin{aligned} W_s^1 f(x_0, y_0) &= \iint_{v \leq k_1 u \text{ and } v \leq k_2 u} \psi_s^1(u, v)dudv \\ &= \int_0^\infty rdr \int_{\theta_0}^{\theta_0 + \theta} \frac{1}{s^2} \psi\left(\frac{r}{s}\right) \cos \alpha d\alpha \\ &= \int_0^\infty rdr \int_{\theta_0}^{\theta_0 + \theta} \psi(r) \cos \alpha d\alpha \\ &= \left(\int_0^\infty r\psi(r)dr \right) [\sin(\theta_0 + \theta) - \sin \theta_0] \\ &= \left(\int_0^\infty r\psi(r)dr \right) 2 \cos\left(\theta_0 + \frac{\theta}{2}\right) \sin \frac{\theta}{2}. \end{aligned}$$

Similarly,

$$\begin{aligned} W_s^2 f(x_0, y_0) &= \iint_{v \leq k_1 u, v \leq k_2 u} \psi_s^2(u, v)dudv \\ &= \left(\int_0^\infty r\psi(r)dr \right) 2 \sin\left(\theta_0 + \frac{\theta}{2}\right) \sin \frac{\theta}{2}. \end{aligned}$$

Hence, we obtain

$$|\nabla W_s e(x_0, y_0)| = 2 \left| \sin \frac{\theta}{2} \right| \left| \int_0^\infty r\psi(r)dr \right|.$$

(2) A curve-Dirac edge can be approximated by a sequence of segments which are straight lines, if the curve is smooth. We can only consider the straight lines when the wavelet ψ is compactly supported. Let the parameter equation of a straight line be the form of

$$\begin{cases} u = x_0 - at; \\ v = y_0 - bt. \end{cases} \quad t \in (-\infty, \infty)$$

Then

$$\begin{aligned} W_s^1 e_l(x_0, y_0) &= \int_{-\infty}^{\infty} \psi_s^1(at, bt) dt \\ &= \int_{-\infty}^{\infty} \psi\left(\frac{\sqrt{a^2 + b^2}}{s}|t|\right) \frac{at}{\sqrt{a^2 + b^2}} dt \\ &= 0. \end{aligned}$$

We can also deduce that

$$W_s^2 e_l(x_0, y_0) = 0.$$

Therefore, we have

$$|\nabla W_s e(x_0, y_0)| = 0.$$

Similarly, the result for the single-point-Dirac edges can be obtained.

(3) For the roof-structure edges, since $\alpha = 1$, we have

$$|\nabla W_s e(x_0, y_0)| \leq Ks. \quad (5.21)$$

It is clear that when $s \rightarrow 0$, then $Ks \rightarrow 0$. Thus, $|\nabla W_s e(x_0, y_0)| \rightarrow 0$.

It is easily seen that the modulus $|\nabla W_s e(x_0, y_0)|$ of the MASW transform with respect to a ideal step-structure edge is a nonzero constant which is independent on both the gradient direction of the edge and the scale of the transform, if $\int_0^\infty r\psi(r)dr \neq 0$. Based on this property, the step-structure edges with various orientations can be detected effectively. The modulus $|\nabla W_s e(x_0, y_0)|$ reaches the maxima, if $\theta = \pi$. Therefore, the closer to π the angle is, the more effective the detection will be.

In practice, signals are usually much more complicated than the ideal ones. They can be regarded as the overlaying of the ideal edges and a background which is a smooth function, or a function with less singularity.

Thus, a practical signal $f(x)$ can be written as

$$f(x) = f_0(x) + \sum_i e_i(x),$$

where $e_i(x)$ denotes an ideal edge and $f_0(x)$ is a smooth signal which has less singularity or even no any singularity.

By applying wavelet transform to $f_0(x) + \sum_i e_i(x)$, we have

$$W_s f(x) = W_s f_0(x) + \sum_i W_s e_i(x).$$

Thus,

$$|W_s f(x) - \sum_i W_s e_i(x)| = |W_s f_0(x)|.$$

Theorem 5.1 shows that $|W_s f_0(x)|$ at the edge x_0 is usually much smaller than $|W_s e_i(x)|$, since $f_0(x)$ has smaller singularity. Hence, the following important conclusion can be obtained:

$$W_s f(x) \approx \sum_i W_s e_i(x).$$

Consequently, all results derived from the ideal case in this subsection can be utilized in the practical signals.

Consider a signal with a noisy function $n(x)$ which is a real, wide sense stationary white noise of variance σ^2 . Thus, its correlation function is

$$E(n(u)n(v)) = \sigma^2 \delta(u - v) = \begin{cases} \sigma^2, & \text{for } u = v; \\ 0, & \text{otherwise.} \end{cases}$$

Since

$$\begin{aligned} |W_s n(x)|^2 &= W_s n(x) \overline{W_s n(x)} \\ &= \int \int n(u) \overline{n(v)} \psi_s(x-u) \overline{\psi}_s(x-v) du dv, \end{aligned}$$

we have

$$\begin{aligned} E|W_s n(x)|^2 &= \int \int \sigma^2 \delta(u - v) \psi_s(x-u) \overline{\psi}_s(x-v) du dv \\ &= \sigma^2 \int |\psi_s(x-u)|^2 du \\ &= \sigma^2 \frac{1}{s} \int |\psi_u|^2 du \end{aligned}$$

$$= \frac{1}{s} \sigma^2 \|\psi\|_2.$$

which shows that the variance of the wavelet transform of the noise is dependent on the scale. The larger the scale is, the less is the variance of the wavelet transform of the noise. Therefore, we can synthesize several scales to detect edges and to remove the noise.

5.3.3 Scale-Independent Algorithm

Based on the above analyses, the concrete algorithm can be developed. This algorithm possesses two objectives: (1) to extract the step-structure edges for one- and two-dimensional signals, (2) to eliminate other edges such as drawing lines, noise, texts, etc.

In view of Corollary 5.1, this implies that the wavelet transform of the step-structure edge is a nonzero constant without considering the scales s . In other words, for different values of scales s_1, s_2, \dots, s_J , their wavelet transforms with respect to the step-structure edges, $W_{s_1}f(x), W_{s_2}f(x), \dots, W_{s_J}f(x)$, are nonzero constants, and are equal each to each. At the points on the step-structure edge, it is clear that

$$\frac{W_{s_j}f(x)}{W_{s_l}f(x)} = 1, \quad (j, l = 1, 2, \dots, J). \quad (5.22)$$

In practice, the sign of equality in Eq. (5.22) may not be satisfied due to the following reasons:

- The edge is not an ideal one as mentioned in the previous subsection;
- The image to be processed is distorted by noise;
- The background of the image may bring a small error in the calculation of the wavelet transforms;
- Other factors may influence the result, for example, the accuracy of the computer system may have an effect on the computation.

Therefore, $W_{s_i}f(x), i = 1, 2, \dots, J$ may not keep the same value for different scales. Consequently, the equality sign in Eq. (5.22) should be replaced by the approximate sign, and Eq. (5.22) now becomes

$$\frac{W_{s_j}f(x)}{W_{s_l}f(x)} \approx 1, \quad (j, l = 1, 2, \dots, J). \quad (5.23)$$

As a matter of fact, the points which satisfy Eq. (5.23) are considered being a step-structure edge.

However, the approximate sign makes uncertainty of extracting edge points. In order to avoid this uncertainty, and to facilitate the design of algorithm detecting the step-structure edges, we utilize a real number R , instead of the integer number 1. Let R be very close to 1 and greater than 1. In replacing sign “ \approx ” by sign “ \leq ”, Eq. (5.23) is equivalent to

$$\frac{W_{s_j}f(x)}{W_{s_l}f(x)} \leq R, \text{ and } \frac{W_{s_l}f(x)}{W_{s_j}f(x)} \leq R, \quad (j, l = 1, 2, \dots, J). \quad (5.24)$$

The closer to 1 the real number R is, the closer to 1 is $\frac{W_{s_j}f(x)}{W_{s_l}f(x)}$, i.e.

$$R \rightarrow 1 \iff \frac{W_{s_j}f(x)}{W_{s_l}f(x)} \rightarrow 1.$$

Moreover, both $W_{s_j}f(x)$ and $W_{s_l}f(x)$ are positive values or negative ones at the same time. In other words, the following is held:

$$\frac{W_{s_j}f(x)}{W_{s_l}f(x)} > 0, \text{ and } \frac{W_{s_l}f(x)}{W_{s_j}f(x)} > 0, \quad (j, l = 1, 2, \dots, J). \quad (5.25)$$

Combining Eqs. (5.24) and (5.25) produces

$$0 < \frac{W_{s_j}f(x)}{W_{s_l}f(x)} \leq R, \text{ and } 0 < \frac{W_{s_l}f(x)}{W_{s_j}f(x)} \leq R, \quad (j, l = 1, 2, \dots, J), \quad (5.26)$$

which is equivalent to

$$\frac{1}{R} \leq \frac{W_{s_j}f(x)}{W_{s_l}f(x)} \leq R, \quad (j, l = 1, 2, \dots, J), \quad (5.27)$$

in the view of mathematics. The real number R in Eq. (5.27) is referred to as *Proportional Threshold*.

Because the edges are high frequency signals, the wavelet transform of them may have large values, thus only the points, which have large values are considered being edges. Therefore, a threshold, T , is established, such that if the points satisfy

$$|W_{s_j}f(x)| \geq T, \quad (5.28)$$

they are on the edge. Number T in Eq. (5.28) is referred to as *Peak Threshold*.

Finally, we can summarize that the points on the step-structure edges have to satisfy

$$|W_{s_j} f(x)| \geq T, \text{ and } \frac{1}{R} \leq \frac{W_{s_j} f(x)}{W_{s_l} f(x)} \leq R, \quad (j, l = 1, 2, \dots, J). \quad (5.29)$$

In view of the discussion in the above, the algorithm can be designed below. Since the wavelet transform of the step-structure edges is scale-independent, this algorithm is referred to as *Scale-Independent Algorithm*.

Algorithm 5.3 Detecting step-structure edges

(1) Given one-dimensional signal $f(x)$:

Step 1: Take different scales s_1, \dots, s_J , and calculate $W_{s_j} f(x)$, ($1 \leq j \leq J$) based on Eq. (5.2) in the previous subsection.

Step 2: Select peak-threshold T , such that

$$|W_{s_j} f(x)| \geq T.$$

Step 3: Select proportional threshold R , such that

$$\frac{1}{R} \leq \frac{W_{s_j} f(x)}{W_{s_l} f(x)} \leq R, \quad (1 \leq j \leq J).$$

Then x is detected as a pixel of the step edge.

(2) For two-dimensional signal $f(x, y)$:

Step 1: Take different scales s_1, \dots, s_J , and calculate $W_{s_j} f(x, y)$, ($1 \leq j \leq J$) based on Eq. (5.3) in the previous subsection.

Step 2: Select peak-threshold T , such that

$$|\nabla W_{s_j} f(x, y)| \geq T.$$

Step 3: Select proportional threshold R , such that

$$\frac{1}{R} \leq \frac{|\nabla W_{s_j} f(x, y)|}{|\nabla W_{s_l} f(x, y)|} \leq R, \quad (1 \leq j \leq J).$$

Then (x, y) is detected as a pixel of the step edge.

It should be pointed out that there is no deviation when the step edges are detected by the above schemes.

5.3.4 Experiments

We focus this subsection on the verification of the effectiveness of the proposed novel wavelet-based method by experiments. Three types of the experiments have been done, namely, (1) extraction of contours of 2-D objects, (2) removal of background in images, and (3) preprocessing in document analysis and recognition. Seven examples will be presented in this subsection.

In these experiments, Eqs. (5.3) and (5.6) have been applied to compute the filtering coefficients $\{\psi_{k,l}^{s,1}\}$ in the MASW transform. These coefficients have been used in the following examples.

First, the experiments of detecting boundaries of 2-D objects are presented. Let us look back at Fig. 5.9. The particular task is that we are required to extract the contour of the aircraft, and to remove all lines and texts. Unfortunately, the algorithm based on the modulus maxima of the wavelet transform [Mallat and Hwang, 1992; Mallat and Zhong, 1992] has detected all edges without identifying different structures of edges. Thus, the resulting image contains lines and texts which are required to be deleted from the image. The new method developed in this subsection possesses an important property, i.e. the wavelet transform of a step-structure edge is scale-independent. It can improve the method of the modulus maxima, and new result can be found in Fig. 5.16. The original image is shown in Fig. 5.16(a) containing a planner object, aircraft, with several drawing lines and texts. Figs. 5.16(b), and (e) display the modulus and angle images undergone by the MASW transform with scale of $s = 2$ respectively. Figs. 5.16(c) and (f) give the modulus and angle images by the MASW transform with scale of $s = 4$ respectively. Figs. 5.16(d) and (g) provide the modulus and angle ones with scale of $s = 8$ respectively. After applying Eq. (5.29) to these images, the resulting image is obtained, and shown in Fig. 5.16(h). It is clear that only the contour of the aircraft has been extracted, while all other edges including drawing lines and texts have been eliminated.

The second example of extracting the contour of the planner objects is illustrated in Fig. 5.17. The original image is given in Fig. 5.17(a). We take $s = 2, 4, 8$, thereafter, the modulus images are obtained in Figs. 5.17(b)-(d), the angle ones are produced in Figs. 5.17(e)-(g). The result is presented in Fig. 5.17(h), where the contour of the 2-D object has been detected, while the others have been removed.

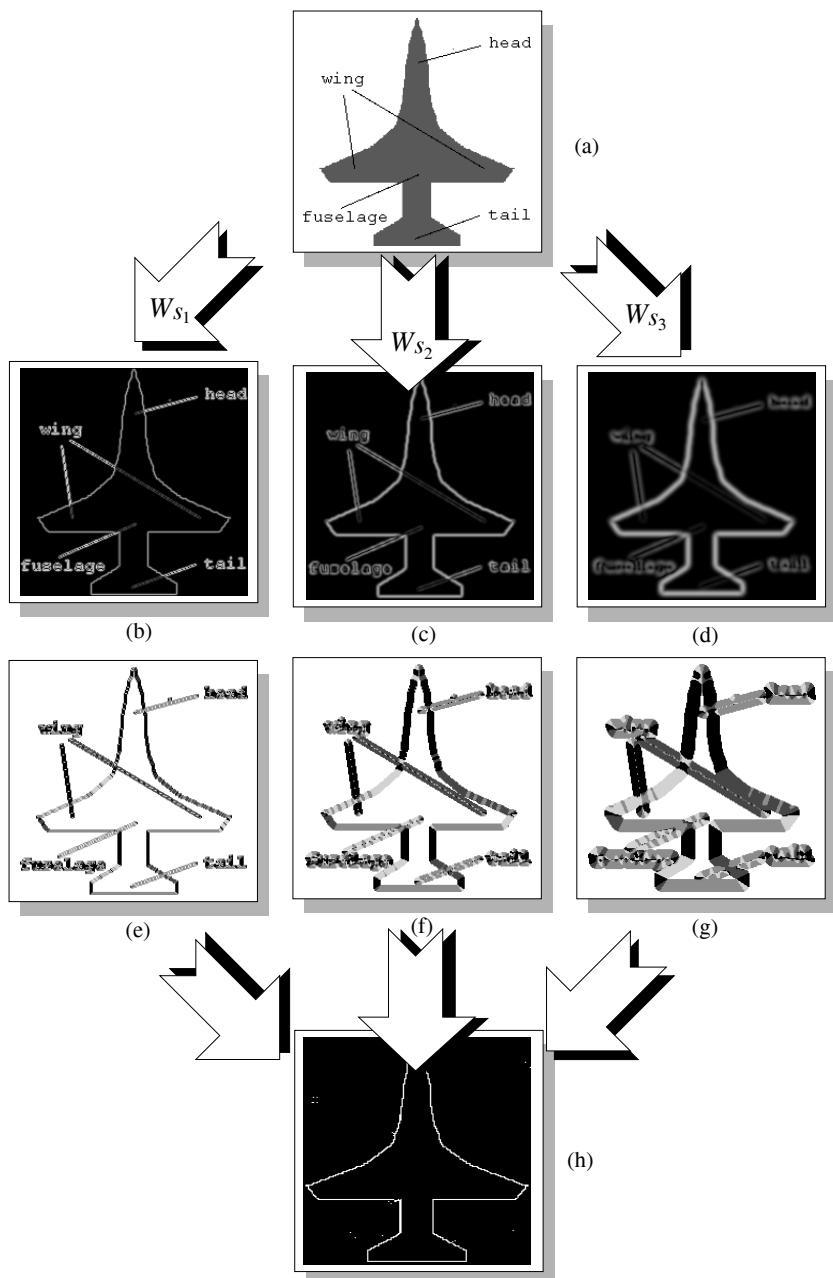


Fig. 5.16 Example 1: the contour of the aircraft is detected using the proposed algorithm.

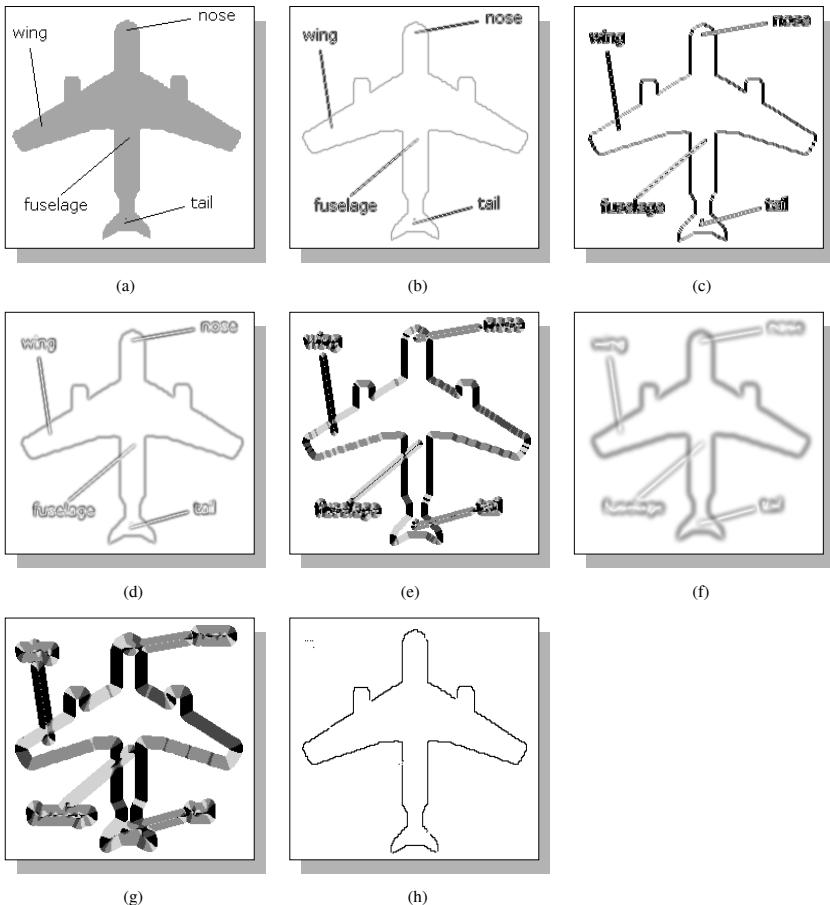


Fig. 5.17 Example 2: the contour of the aircraft is detected using the proposed algorithm.

Now, we turn to other examples shown in Figs. 5.18 and 5.19. We are required to extract the boundaries of the characters, for example, English printed character “A” and Chinese handwriting “book”, and to delete the drawing lines and noise. The same problem was arisen when different scales, for example, $s = 2, 4, 8$ were used in the modulus maxima of the wavelet transform. To contest this deficiency, the scale-independent wavelet method is applied to these examples. For the English printed character, the original image is shown in Fig. 5.18(a) containing a character, English letter “A”,

with several drawing lines and white noise. Figs. 5.18(b)-(d) display the modulus images of the MASW transforms with scales $s = 2, 4, 8$ respectively. Figs. 5.18(e)-(g) show the angle images of the MASW transforms of $s = 2, 4, 8$ respectively. After applying Eq. (5.29) to these images, the resulting image is obtained, and shown in Fig. 5.18(h). Only the boundary of the character has been extracted, and all other edges including drawing lines and noise have been eliminated.

Detection of the contour from a Chinese handwriting using the scale-independent algorithm can be found in Fig. 5.19, where, the original image is depicted in Fig. 5.19(a), and the resulting one is illustrated in Fig. 5.19(h).

The proposed method can be applied to document processing. An example is depicted in Fig. 5.20, which is an improvement of the results (see Figs. 5.10 and 5.11) produced by the modulus maxima. The original image is shown in Fig. 5.20(a). It is a mailing address that is written on a paper with guide lines. In order to automatically process it by computers, a preprocessing is to remove the guide lines and extract the contours of the handwriting English letters. Figs. 5.20(b)-(c) indicate the modulus images of MASW transform with scales $s = 2$ and $s = 4$ respectively. Figs. 5.20(d)-(e) display the angle images with wavelet transform of scales $s = 2$ and $s = 4$ respectively. Using the algorithm mentioned in previous subsection produces an image in Fig. 5.20(f), that is what we need.

The sixth example is shown in Fig. 5.21, where a Chinese article is written on a paper with guide-mesh. Fig. 5.21(a) gives the original image. Figs. 5.21(b) and (c) display the modulus and angle images undergone by the MASW transform with scale of $s = 2$ respectively. Figs. 5.21(d) and (e) give the modulus and angle images by the MASW transform with scale of $s = 4$ respectively. After applying Eq. (5.29) to these images, the resulting image is obtained, and depicted in Fig. 5.21(h).

The seventh example is depicted in Fig. 5.22. The original image is illustrated in Fig. 5.22(a), and the resulting ones in accordance with the modulus maxima are depicted in Figs. 5.22(f) and (g), which are not the desired images, since they include not only the Chinese handwriting, but the guide lines as well. Applying our method produces a good result and is displayed in Fig. 5.22(h), in which, only Chinese handwritten characters are kept, and others are removed.

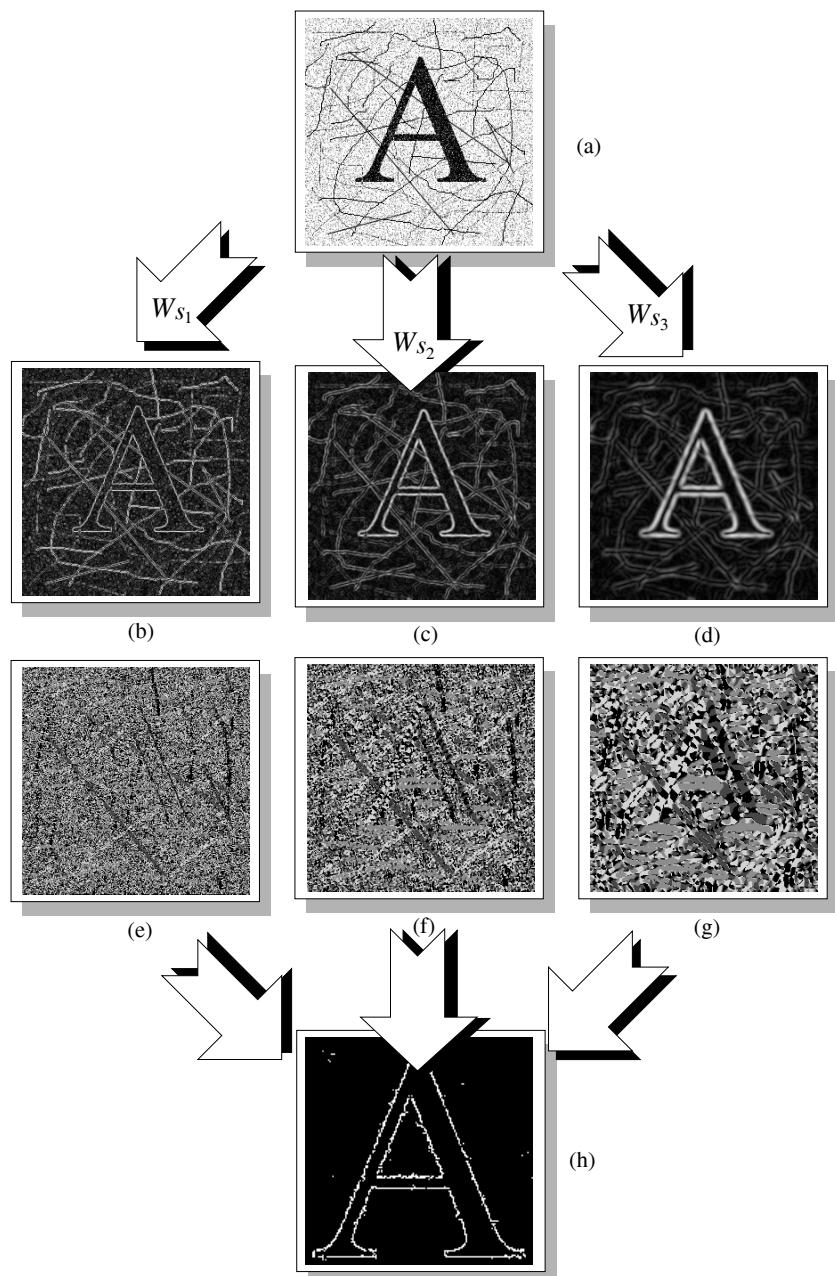


Fig. 5.18 Example 3: The boundary of English letter "A" is extracted using the proposed algorithm.

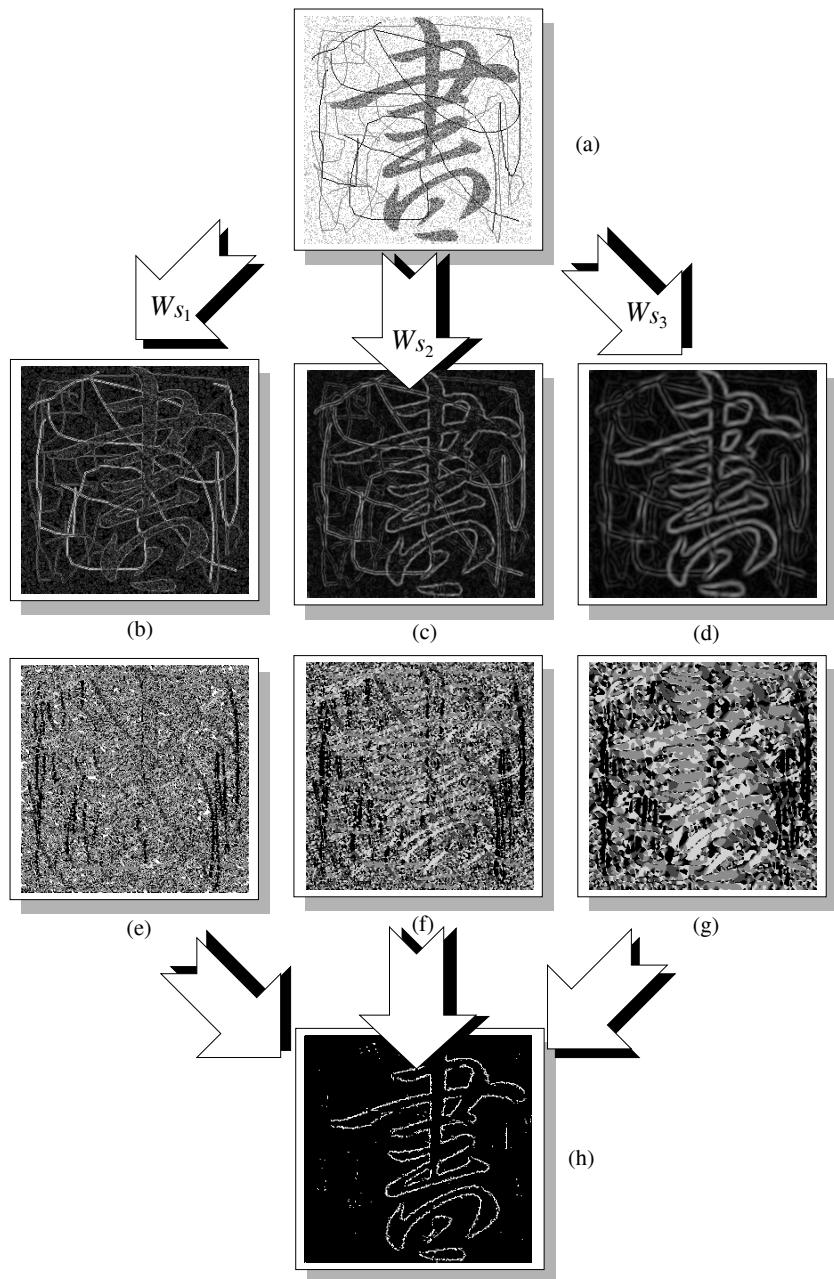


Fig. 5.19 Example 4: The contours of Chinese handwriting "book" are extracted using the proposed algorithm.

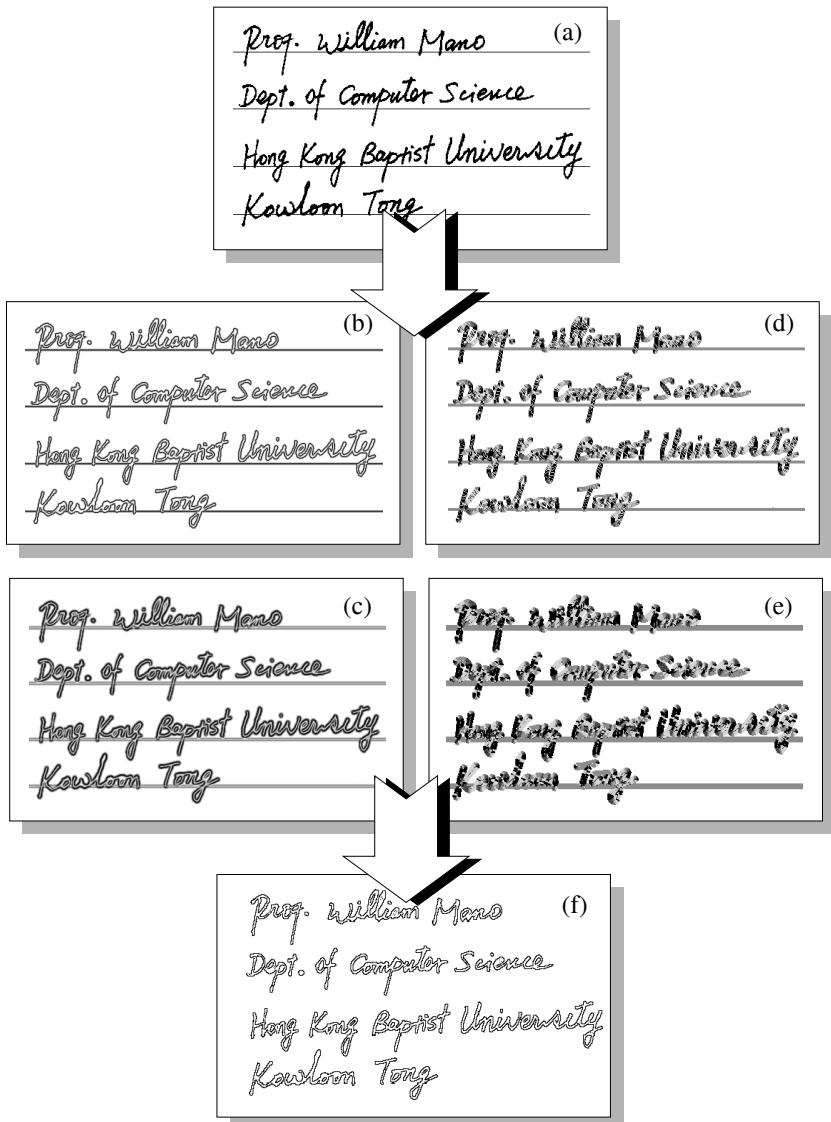


Fig. 5.20 Example 5: The boundaries of English letters are extracted using the proposed algorithm.

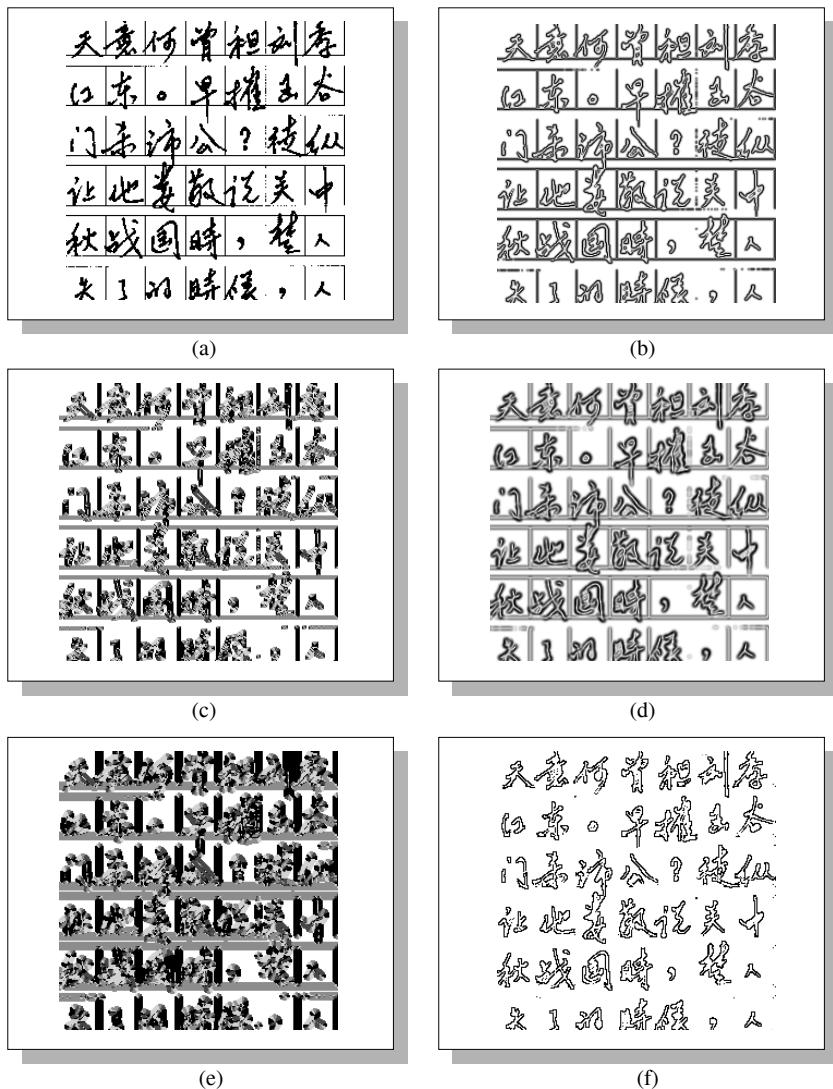


Fig. 5.21 Example 6: The contours of handwriting Chinese characters are extracted using the proposed algorithm.

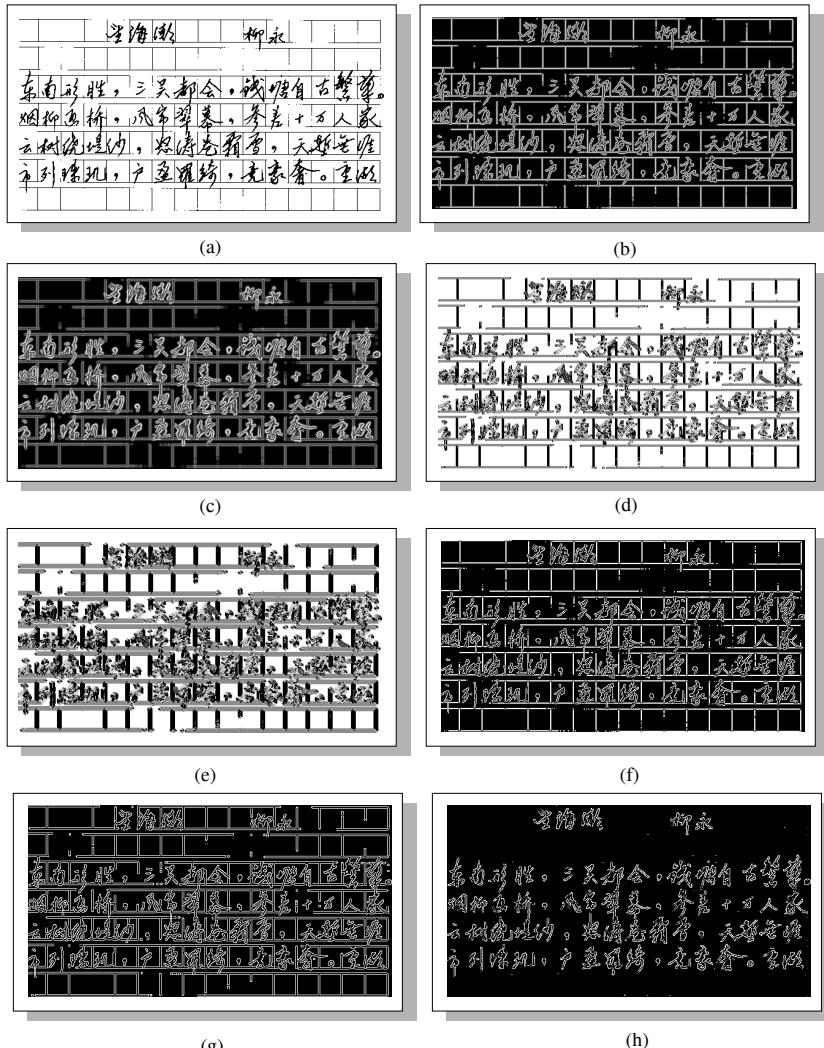


Fig. 5.22 Example 7: The boundaries of handwriting Chinese characters are extracted using the proposed algorithm.

Chapter 6

Characterization of Dirac-Edges with Quadratic Spline Wavelet Transform

This chapter aims at studying the characterization of Dirac-structure edges with wavelet transform, and selecting the suitable wavelet functions to detect them. Three significant characteristics of the local maximum modulus of the wavelet transform with respect to the Dirac-structure edges are presented, namely: (1) slope invariant: the local maximum modulus of the wavelet transform of a Dirac-structure edge is independent on the slope of the edge. (2) grey-level invariant: the local maximum modulus of the wavelet transform with respect to a Dirac-structure edge takes place at the same points when the images with different grey-levels are processed. (3) width light-dependent: for various widths of the Dirac-structure edge images, the location of maximum modulus of the wavelet transform varies lightly under the certain circumscription that the scale of the wavelet transform is larger than the width of the Dirac-structure edges. It is important, in practice, to select the suitable wavelet functions, according to the structures of edges. For example, Haar wavelet is better to represent brick-like images than other wavelets. A mapping technique is applied in this chapter to construct such a wavelet function. In this way, a low-pass function is mapped onto a wavelet function by a derivation operation. In this chapter, the quadratic spline wavelet is utilized to characterize the Dirac-structure edges and an algorithm to extract the Dirac-structure edges by wavelet transform is also developed.

A mathematical characterization of three basic geometric structures of edges (i.e. step-structure, roof-structure, and Dirac-structure) with Lipschitz exponents has been presented in Chapter 5. A significant property has been proved that the modulus of wavelet transform at each point of

the step edge is a non-zero constant which is independent on both the gradient direction and the scale of the wavelet transform. This property led to provide a simple and direct strategy for detecting a specific structure of edges, the step-structure edges. Thus, an algorithm called scale-independent algorithm has been developed. The method developed in Chapter 5 possesses an important property, i.e. the wavelet transform of a step-structure edge is scale-independent. It can improve the method proposed in [Mallat and Hwang, 1992; Mallat and Zhong, 1992], and the result can be found in Fig. 5.17 of Chapter 5, where the modulus-angle-separated-wavelet (MASW) has been used. The precise definition of the MASW can be found in Chapter 5 and [Tang et al., 1998c]. The original image is shown in Fig. 5.17(a) containing a planner object, aircraft, with several drawing lines and texts. Figs. 5.17(b) and (c) display the modulus and angle images undergone by the MASW transform with scale of $s = 2$ respectively. Figs. 5.17(d) and (e) give the modulus and angle images by the MASW transform with scale of $s = 4$ respectively. Figs. 5.17(f) and (g) provide the modulus and angle images with scale of $s = 8$ respectively. After applying the scale-independent algorithm to these images, the resulting image is obtained, and shown in Fig. 5.17(h). It is clear that only the contour of the aircraft has been extracted, while all other edges including drawing lines and texts have been eliminated.

On the other hand, as a complement of Chapter 5, the purpose of this chapter is to develop a method which can identify different structures of edges, thereafter, detect the Dirac-structure edges such as the drawing lines and texts in Figs. 5.17(a) and eliminate the step-structure edges such as the contour of the aircraft in Figs. 5.17(a).

6.1 Selection of Wavelet Functions by Derivation

In practice, the selection of a suitable wavelet function in accordance with the structure of the edges is an important topic in the application of wavelet transform to image processing and pattern recognition. In this section, a method of selection of wavelet function by derivation of the low-pass function will be presented.

6.1.1 Scale Wavelet Transform

Let $L^2(R^2)$ be the Hilbert space of all the square-integrable 2-D functions on plane R^2 , $\psi \in L^2(R^2)$ is called a wavelet function, if

$$\int_R \int_R \psi(x, y) dx dy = 0, \quad (6.1)$$

For $f \in L^2(R^2)$ and scale $s > 0$, the scale wavelet transform of $f(x, y)$ is defined by

$$\begin{aligned} W_s f(x, y) &:= (f * \psi_s)(x, y) \\ &= \int_R \int_R f(u, v) \frac{1}{s^2} \psi\left(\frac{x-u}{s}, \frac{y-v}{s}\right) du dv, \end{aligned} \quad (6.2)$$

where $*$ denotes the convolution operator, and

$$\psi_s(u, v) := \psi\left(\frac{u}{s}, \frac{v}{s}\right).$$

The theory dealing with the scale wavelet transform can be found in many articles, such as [Chui, 1992; Daubechies, 1992]. In practice, the wavelet transform can be calculated discretely using the following formula:

$$\begin{aligned} W_s f(n, m) &= \int \int f(u, v) \psi_s(n - u, m - v) du dv \\ &= \sum_{k,l} f(k, l) \int_k^{k+1} \int_l^{l+1} \psi_s(n - u, m - v) du dv \\ &= \sum_{k,l} f(k, l) \int_{n-k-1}^{n-k} \int_{m-l-1}^{m-l} \psi_s(u, v) du dv \\ &= \sum_{k,l} f(n - k - 1, m - l - 1) \psi_{k,l}^s, \end{aligned}$$

where

$$\psi_{k,l}^s = \int_k^{k+1} \int_l^{l+1} \psi_s(u, v) du dv = \int_{k/s}^{(k+1)/s} \int_{l/s}^{(l+1)/s} \psi(u, v) du dv.$$

Obviously, the scale wavelet transform described in Eq. (6.2) is a filter in essence. We can conclude that its Fourier transform defined by

$$\hat{\psi}(\xi, \eta) := \int_R \int_R \psi(x, y) e^{-i(\xi x + \eta y)} dx dy$$

satisfies the condition of $\hat{\psi} \in L^2(R^2)$, since $\psi \in L^2(R^2)$. Thus, both functions ψ and $\hat{\psi}$ decrease at infinity. Usually, ψ is chosen to meet the following two conditions at least: First, $\psi \in L^2(R^2)$ has to be held; Second, both ψ and $\hat{\psi}$ decrease with exponents at infinity. According to Eq. (6.1) we can see that $\hat{\psi}(0, 0) = 0$, which implies that the scale wavelet transform $W_s f(x, y)$ corresponds to a band-pass filter essentially. Moreover, $W_s f(x, y)$ characterizes the local properties of $f(x, y)$ on both the time and frequency domains, namely:

- On the time domain, it is easy to see that

$$W_s f(x, y) := (f * \psi_s)(x, y) = \int_R \int_R f(x - u, y - v) \psi_s(u, v) du dv$$

characterizes the local property of f around the point (x, y) . The smaller the scale s , the more narrow the time-window is.

- On the frequency domain, based on the basic properties of Fourier transform, it can be conclude that

$$\begin{aligned} (W_s f)(x, y) &= \left(\frac{1}{2\pi}\right)^2 \int_R \int_R (W_s f)^\gamma(\xi, \eta) e^{i(\xi x + \eta y)} d\xi d\eta \\ &= \left(\frac{1}{2\pi}\right)^2 \int_R \int_R \hat{f}(\xi, \eta) (\psi_s)^\gamma(\xi, \eta) e^{i(\xi x + \eta y)} d\xi d\eta \\ &= \left(\frac{1}{2\pi}\right)^2 \int_R \int_R \hat{f}(\xi, \eta) \hat{\psi}(s\xi, s\eta) e^{i(\xi x + \eta y)} d\xi d\eta. \end{aligned}$$

Since $\hat{\psi}(s\xi, s\eta) e^{i(\xi x + \eta y)}$ decreases as $\hat{\psi}(s\xi, s\eta)$ does, thus $(W_s f)(x, y)$ characterizes the local property of \hat{f} . The position of the locality depends on the scale s and function ψ . And the larger the scale s , the narrow the frequency-window is.

The reduction of ψ and $\hat{\psi}$ at infinity plays an important role for characterizing the local properties of $f(x, y)$ on both the time and frequency domains. The faster the decrement of ψ and $\hat{\psi}$, the better the characterization of the local properties will be. In other words, the shorter the support of ψ and $\hat{\psi}$, the higher the quality of the localization will be obtained.

6.1.2 Construction of Wavelet Function by Derivation of the Low-Pass Function

The Dirac-structure edges such as curves are transient components with high frequencies in an image. They are highly localized in spatial positions. Such components do not resemble any of the wave basis functions, such as Fourier basis function. This makes the Fourier and other wave transforms less than optimal representations for analyzing and processing the Dirac-structure edges in the images. To combat such a deficiency, wavelet analysis can be utilized. Wavelet theory is a good mathematical tool primarily used for representing such transient components more efficiently. In fact, the Dirac-structure edges can be characterized by the wavelet transforms.

However, one of the key factors which obstruct the application of wavelet transform to process these Dirac-structure edges is that it is still difficult, in practice, to select the suitable wavelet functions, which possess a perfect characteristic of localization.

Theoretically, the detection of the Dirac-structure edges by wavelet transforms can be regarded as a particular filtering operation. The derivative function of a smooth low-pass function which decreases at infinite can become a candidate of the wavelet function. It can be considered to be a mapping, i.e. a low-pass function can be mapped onto a wavelet function by the operation of derivation, which can also be described by the following:

$$\underbrace{\text{Low - Pass Function}}_{\text{Derivation}} \xrightarrow{\text{Mapping}} \text{Wavelet Function} \quad (6.3)$$

According to Eq. (6.3), we can use such a method to produce a wavelet function in the following steps: (1) selecting a low-pass function $\theta(x, y)$, (2) deriving this low-pass function to produce the wavelet functions $\psi^1(x, y)$ and $\psi^2(x, y)$.

Let $\theta(x, y)$ be a real function, and satisfies:

- $\theta(x, y)$ fast decreases at infinity;
- $\theta(x, y)$ is even function on both x and y ;
- $\hat{\theta}(0, 0) = 1$.

Consider

$$\psi^1(x, y) := \frac{\partial}{\partial x} \theta(x, y),$$

it is easy to see that $\int \int \psi^1(x, y) dx dy = 0$, which indicates that $\psi^1(x, y)$ is a 2-D wavelet. Then its scale wavelet transform is

$$\begin{aligned} W_s^1 f(x, y) &= (f * \psi_s^1)(x, y) \\ &= (f * s \frac{\partial}{\partial x} \theta_s)(x, y) \\ &= s \frac{\partial}{\partial x} (f * \theta_s)(x, y) \end{aligned}$$

where $\theta_s(x, y) := \frac{1}{s^2} \theta(\frac{x}{s}, \frac{y}{s})$. Based on this formula, it is clear that $f * \theta_s$ is a smooth operation with scale s , if θ is a smooth function with fast decreasing at infinity and $\hat{\theta}(0, 0) = 1$. When a quadratic spline function is utilized as the primitive function $\theta(x, y)$, the graphical description of the derivative function of the function $\theta(x, y)$ with respect to the horizontal axis x is illustrated in Fig. 6.1.

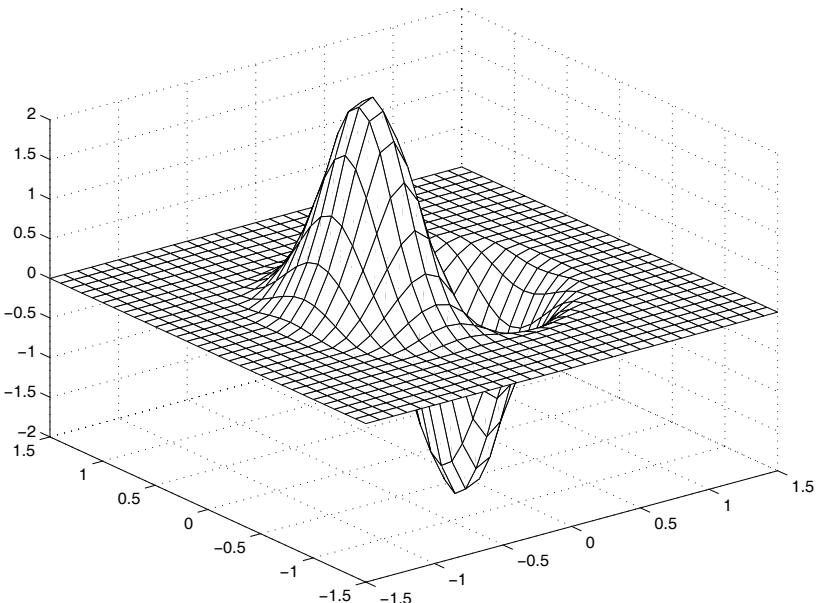


Fig. 6.1 The graphical description of derivative function $\psi^1(x, y) = \frac{\partial}{\partial x} \theta(x, y)$.

Since $W_s^1 f(x, y)$ is the derivative function of function $\theta(x, y)$ along the horizontal axis, the characteristic of the local maxima modulus of the

wavelet transform mainly influences the transient components of an image along the horizontal axis.

Similarly, let

$$\psi^2(x, y) := \frac{\partial}{\partial y} \theta(x, y).$$

It is also a 2-D wavelet, and its scale wavelet transform is

$$W_s^2 f(x, y) = (f * \psi_s^2)(x, y)$$

or

$$W_s^2 f(x, y) = s \frac{\partial}{\partial y} (f * \theta_s)(x, y)$$

which is the derivative function of function $\theta(x, y)$ along the vertical axis. Therefore, the characteristic of the local maxima modulus of the wavelet transform mainly influences the transient components of an image along the vertical axis. The graphical description of the derivative function of the function $\theta(x, y)$ with respect to the vertical axis y is illustrated in Fig. 6.2, where the primitive function is a quadratic spline function .

We denote the gradient direction and the amplitude of the wavelet transform respectively by

$$\nabla W_s f(x, y) := \begin{pmatrix} W_s^1 f(x, y) \\ W_s^2 f(x, y) \end{pmatrix} \quad (6.4)$$

and

$$|\nabla W_s f(x, y)| := \sqrt{|W_s^1 f(x, y)|^2 + |W_s^2 f(x, y)|^2}. \quad (6.5)$$

Locating the local maxima of $|\nabla W_s f(x, y)|$ along the direction of $\nabla W_s f(x, y)$ can detect the Dirac-structure edges including the curves of images.

In this chapter, function $\theta(x, y)$ is defined by

$$\theta(x, y) = \phi(\sqrt{x^2 + y^2}), \quad (6.6)$$

where ϕ is selected to be the quadratic spline function as follows:

$$\phi(r) = \begin{cases} 8r^2(r - 1) + \frac{4}{3} & 0 \leq r < \frac{1}{2} \\ -\frac{8}{3}(r - 1)^3 & \frac{1}{2} \leq r < 1 \\ 0 & r \geq 1. \end{cases} \quad (6.7)$$

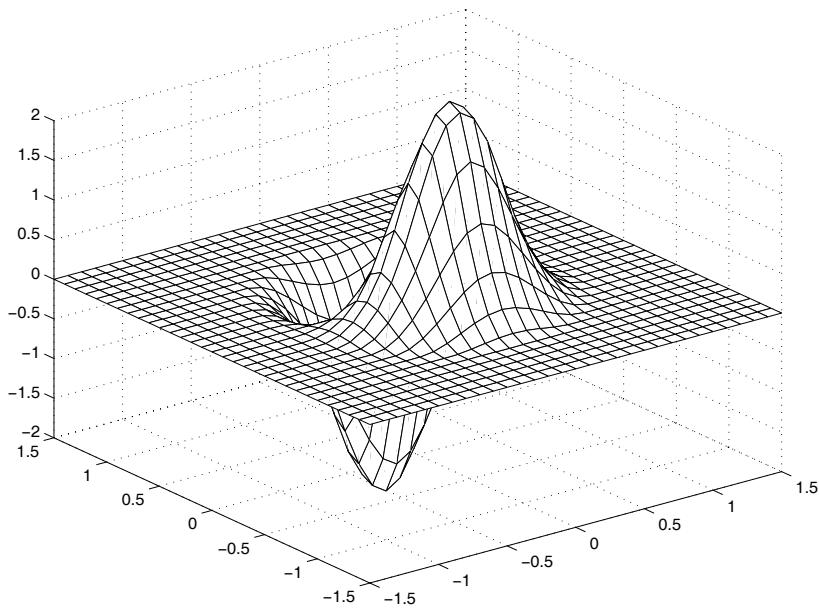


Fig. 6.2 The graphical description of derivative function $\psi^2(x, y) = \frac{\partial}{\partial y}\theta(x, y)$.

Since

$$\psi^1(x, y) = \frac{\partial}{\partial x}\theta(x, y), \quad \psi^2(x, y) = \frac{\partial}{\partial y}\theta(x, y), \quad (6.8)$$

The wavelet function ψ^1 , ψ^2 , which are called the quadratic spline wavelets, can be obtained as follows:

$$\begin{cases} \psi^1(x, y) = \phi'(\sqrt{x^2 + y^2}) \frac{x}{\sqrt{x^2 + y^2}} \\ \psi^2(x, y) = \phi'(\sqrt{x^2 + y^2}) \frac{y}{\sqrt{x^2 + y^2}}. \end{cases} \quad (6.9)$$

which can be represented graphically in Figs. 6.1 and 6.2.

6.2 Characterization of Dirac-Structure Edges by Wavelet Transform

In this section, three significant characteristics of the local maximum modulus of the wavelet transform with respect to the Dirac-structure edges in

the images will be presented, namely:

- Slope invariant: the local maximum modulus of the wavelet transform of a Dirac-structure edge is independent on the slope of the edge.
- Grey-level invariant: the local maximum modulus of the wavelet transforms with respect to a Dirac-structure edge takes place at the same points when the images with different grey-levels are to be processed.
- Width light-dependent: for various widths of images of the Dirac-structure edges, the location of maximum modulus of the wavelet transform varies lightly under the certain circumscription.

Curve is a display of the Dirac-structure edge in a two-dimensional image. To simplify the theoretical analysis, in this chapter, a segment of the curve will be considered. In the remainder of this book, we will not identify the notations of the Dirac-structure edge, curve, and segment of curve.

Before mathematically analyzing the characterization of curves by wavelet transform, we look at an example as shown in Fig. 6.3. In practice, we should consider the width of the curve, especially in the application of image processing. Fig. 6.3(a) gives a segment, which is marked by letter "A", of a curve with certain width. Suppose the quadratic spline function is selected to construct the wavelet functions. The distribution of the modulus of wavelet transform is illustrated in Fig. 6.3(b). Thus, the maximum modulus of wavelet transform will occur on two parallel lines around the curve image.

Now, we turn to the mathematical analysis. Suppose that the parameter equation of a curve l_d can be written in form of

$$l_d : \begin{cases} u = u(t) \\ v = v(t) \end{cases} \quad t \in [a, b], \quad (6.10)$$

where, $[a, b]$ denotes the interval of the curve. Let d be the width, and l be the skeleton of the curve, i.e. the central line of the curve image. A graphical description of such a curve is presented in Fig. 6.4. Note that, from the mathematical point of view, a curve does not have any width with it. Thus, we did not consider the width of the curve in Eq. (6.10). Therefore, we do not identify the notations of l and l_d . In Fig. 6.4, the lines labeled by l_λ are parallel lines, which are around the curve image, and

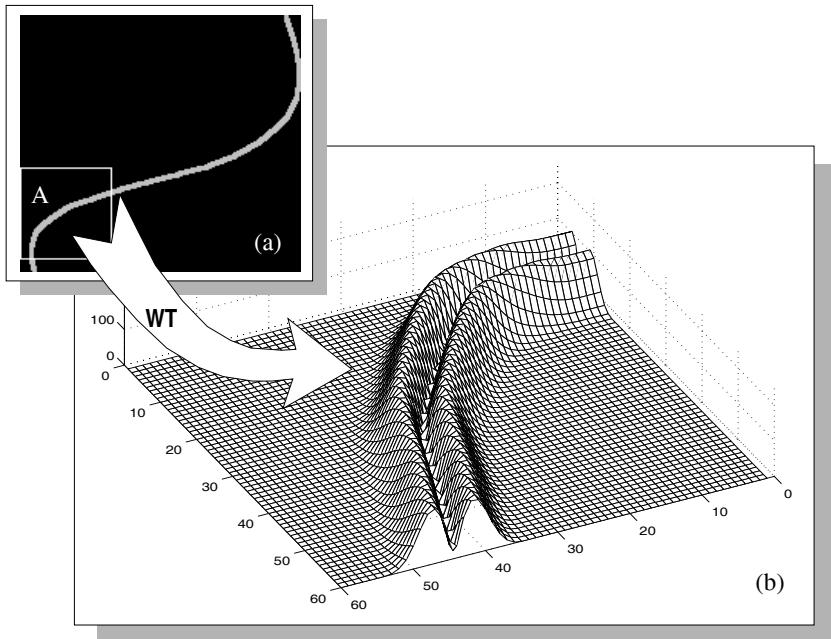


Fig. 6.3 An example of the distribution of the modulus of wavelet transform (WT) with respect to a curve.

the maximum modulus of the wavelet transform occurs on these lines. The normal unit vector of the curve l_d is described by

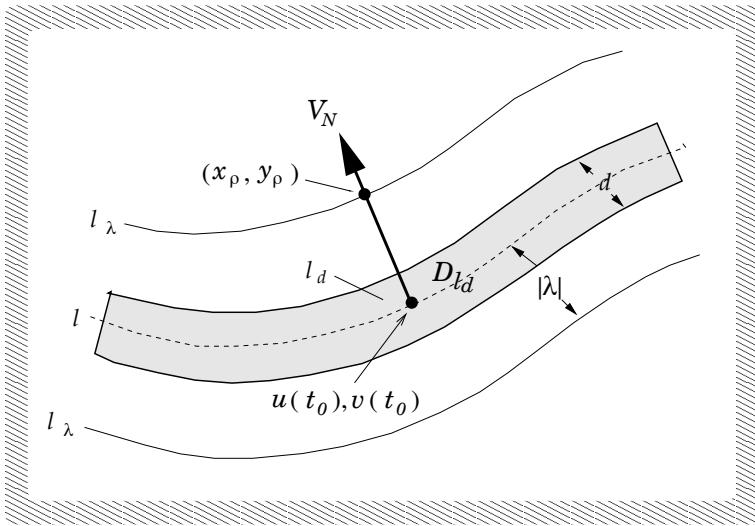
$$V_N = \left(\frac{v'(t)}{\sqrt{u'(t)^2 + v'(t)^2}}, -\frac{u'(t)}{\sqrt{u'(t)^2 + v'(t)^2}} \right).$$

Therefore, the equations of the parallel line l_λ can be represented as follows:

$$l_\lambda : \begin{cases} u_\lambda(t) = u(t) + \frac{\lambda v'(t)}{\sqrt{u'(t)^2 + v'(t)^2}} \\ v_\lambda(t) = v(t) - \frac{\lambda u'(t)}{\sqrt{u'(t)^2 + v'(t)^2}} \end{cases} \quad t \in [a, b],$$

where $|\lambda|$ is the distance between the parallel line l_λ and the central line l of the curve l_d , and the sign of λ determines which side of l_d the parallel line l_λ will be located. A curve image l_d can be described by

$$f_{l_d}(x, y) = c_f \chi_{D_{l_d}}(x, y)$$

Fig. 6.4 Curve l_d and its parallel line l_λ .

where c_f stands for the *grey level* of the curve and $\chi_{D_{l_d}}(x, y)$ denotes the characteristic function of the area D_{l_d} which is defined by

$$D_{l_d} := \{(u_\lambda(t), v_\lambda(t)) | t \in [a, b], \lambda \in [-d/2, d/2]\}.$$

$\forall g(x, y)$, we have

$$\begin{aligned} \int_R \int_R f_{l_d}(u, v) g(u, v) du dv &= c_f \int \int_{D_{l_d}} g(u, v) du dv \\ &= c_f \int_a^b dt \int_{-d/2}^{d/2} g(u_\lambda(t), v_\lambda(t)) J(t, \lambda) d\lambda, \end{aligned}$$

where $J(t, \lambda)$ is the Jacobi determinant of the coordinate transform

$$\begin{cases} u = u(t, \lambda) \\ v = v(t, \lambda) \end{cases},$$

i.e.

$$J(t, \lambda) = \left| \begin{array}{cc} \frac{\partial u}{\partial t} & \frac{\partial u}{\partial \lambda} \\ \frac{\partial v}{\partial t} & \frac{\partial v}{\partial \lambda} \end{array} \right| = -\sqrt{u'(t)^2 + v'(t)^2} - \lambda \frac{u'(t)v''(t) - v'(t)u''(t)}{u'(t)^2 + v'(t)^2}.$$

Hence

$$\begin{aligned} & \int_R \int_R f_{l_d}(u, v) g(u, v) dudv = \\ & -c_f \int_a^b dt \int_{-d/2}^{d/2} \left[\sqrt{u'(t)^2 + v'(t)^2} + \lambda \frac{u'(t)v''(t) - v'(t)u''(t)}{u'(t)^2 + v'(t)^2} \right] . \\ & g \left(u(t) + \frac{\lambda v'(t)}{\sqrt{u'(t)^2 + v'(t)^2}}, v(t) - \frac{\lambda u'(t)}{\sqrt{u'(t)^2 + v'(t)^2}} \right) d\lambda. \end{aligned}$$

Let $g(u, v) = \psi_s(u, v)$, we have

$$\begin{aligned} & \int_R \int_R f_{l_d}(u, v) \psi_s(u, v) dudv = \\ & -c_f \int_a^b dt \int_{-d/2}^{d/2} \left[\sqrt{u'(t)^2 + v'(t)^2} + \lambda \frac{u'(t)v''(t) - v'(t)u''(t)}{u'(t)^2 + v'(t)^2} \right] . \\ & \psi_s \left(u(t) + \frac{\lambda v'(t)}{\sqrt{u'(t)^2 + v'(t)^2}}, v(t) - \frac{\lambda u'(t)}{\sqrt{u'(t)^2 + v'(t)^2}} \right) d\lambda. \end{aligned}$$

When a segment of curve l_d is short enough, it can be considered to be a straight line, thereafter, the parameter equation of its central line can be represented as follows:

$$l : \begin{cases} u = u(t) + k_1(t - t_0) & t \in [a, b], \\ v = v(t) + k_2(t - t_0) \end{cases}$$

where k_1, k_2 denote the slope of l satisfying $k_1^2 + k_2^2 = 1$. Thus, the wavelet transform of this curve can be obtained

$$\begin{aligned} W_s f_{l_d}(x, y) = & -c_f \int_a^b dt \int_{-d/2}^{d/2} \psi_s(x - u(t_0) - k_1(t - t_0) - \\ & \lambda k_2, y - v(t_0) - k_2(t - t_0) + \lambda k_1) d\lambda. \end{aligned}$$

$\forall (u(t_0), v(t_0)) \in l$, we consider the point (x_ρ, y_ρ) in the normal line:

$$\begin{cases} x_\rho = u(t_0) + k_2 \rho \\ y_\rho = v(t_0) + k_1 \rho \end{cases},$$

where ρ denotes a parameter such that its absolute value $|\rho|$ is the just distance between (x_ρ, y_ρ) and $(u(t_0), v(t_0))$, i.e. $|\lambda| = |\rho|$. The local maximum modulus of wavelet transform of the point $(u(t_0), v(t_0))$ in the curve takes place at point (x_ρ, y_ρ) in the normal line. As the parameter ρ is determined,

the point (x_ρ, y_ρ) will be fund. Therefore, the calculation of parameter ρ plays a key role in finding the location where the maximum modulus of wavelet transform occurs. In this case, the wavelet transform becomes

$$\begin{aligned}
 W_s f_{l_d}(x, y) &= -c_f \int_a^b dt \int_{-d/2}^{d/2} \psi_s(k_2\rho - k_1(t - t_0) - \lambda k_2, \\
 &\quad -k_1\rho - k_2(t - t_0) + \lambda k_1) d\lambda \\
 &= -c_f \int_a^b dt \int_{-d/2}^{d/2} \psi_s(-k_1(t - t_0) - k_2(\lambda - \rho), \\
 &\quad -k_2(t - t_0) + k_1(\lambda - \rho)) d\lambda \\
 &= -c_f \int_{t_0-b}^{t_0-a} dt \int_{-d/2-\rho}^{d/2-\rho} \psi_s(k_1t - k_2\lambda, k_2t + k_1\lambda) d\lambda \\
 &= -c_f \int_{t_0-b}^{t_0-a} dt \int_{\rho-d/2}^{\rho+d/2} \psi_s(k_1t + k_2\lambda, k_2t - k_1\lambda) d\lambda \\
 &= -c_f \int_{(t_0-b)/s}^{(t_0-a)/s} dt \int_{(\rho-d/2)/s}^{(\rho+d/2)/s} \psi(k_1t + k_2\lambda, k_2t - k_1\lambda) d\lambda.
 \end{aligned}$$

6.2.1 Slope Invariant

In this sub-section, we will prove that the local maximum modulus of the wavelet transform of the curve are independent on the slope of that curve.

For the quadratic spline wavelets ψ^1, ψ^2 , since

$$\psi^1(x, y) = \frac{\partial}{\partial x} \theta(x, y), \quad \psi^2(x, y) = \frac{\partial}{\partial y} \theta(x, y),$$

where $\theta(x, y) = \phi(\sqrt{x^2 + y^2})$, and

$$\phi(r) = \begin{cases} 8r^2(r-1) + \frac{4}{3} & 0 \leq r < \frac{1}{2} \\ -\frac{8}{3}(r-1)^3 & \frac{1}{2} \leq r < 1 \\ 0 & r \geq 1, \end{cases}$$

therefore, ψ^1, ψ^2 can be represented as

$$\begin{cases} \psi^1(x, y) = \phi'(\sqrt{x^2 + y^2}) \frac{x}{\sqrt{x^2 + y^2}} \\ \psi^2(x, y) = \phi'(\sqrt{x^2 + y^2}) \frac{y}{\sqrt{x^2 + y^2}}. \end{cases}$$

The wavelet transforms using the quadratic spline wavelets ψ^1 , ψ^2 are

$$W_s^1 f(x_\rho, y_\rho) = -c_f \int_{(t_0-b)/s}^{(t_0-a)/s} dt \int_{(\rho-\frac{d}{2})/s}^{(\rho+\frac{d}{2})/s} \phi'(\sqrt{t^2 + \lambda^2}) \frac{k_1 t + k_2 \lambda}{\sqrt{t^2 + \lambda^2}} d\lambda \quad (6.11)$$

$$W_s^2 f(x_\rho, y_\rho) = -c_f \int_{(t_0-b)/s}^{(t_0-a)/s} dt \int_{(\rho-\frac{d}{2})/s}^{(\rho+\frac{d}{2})/s} \phi'(\sqrt{t^2 + \lambda^2}) \frac{k_2 t - k_1 \lambda}{\sqrt{t^2 + \lambda^2}} d\lambda. \quad (6.12)$$

It is easy to rewrite Eqs. (6.11) and (6.12) in following forms

$$\begin{aligned} W_s^1 f(x_\rho, y_\rho) &= -c_f(k_1 w_1 + k_2 w_2) \\ W_s^2 f(x_\rho, y_\rho) &= -c_f(k_2 w_1 - k_1 w_2), \end{aligned}$$

by denoting

$$\begin{aligned} w_1 &= \int_{(t_0-b)/s}^{(t_0-a)/s} dt \int_{(\rho-\frac{d}{2})/s}^{(\rho+\frac{d}{2})/s} \phi'(\sqrt{t^2 + \lambda^2}) \frac{t}{\sqrt{t^2 + \lambda^2}} d\lambda \\ w_2 &= \int_{(t_0-b)/s}^{(t_0-a)/s} dt \int_{(\rho-\frac{d}{2})/s}^{(\rho+\frac{d}{2})/s} \phi'(\sqrt{t^2 + \lambda^2}) \frac{\lambda}{\sqrt{t^2 + \lambda^2}} d\lambda. \end{aligned}$$

Hence, the second power of the amplitude of the wavelet transform can be obtained

$$\begin{aligned} |\nabla W_s f_{l_d}(x_\rho, y_\rho)|^2 &= |W_s^1 f_{l_d}(x_\rho, y_\rho)|^2 + |W_s^2 f_{l_d}(x_\rho, y_\rho)|^2 \\ &= c_f^2 [(k_1 w_1 + k_2 w_2)^2 + (k_2 w_1 - k_1 w_2)^2] \\ &= c_f^2 (w_1^2 + w_2^2). \end{aligned}$$

Consequently, the amplitude of the wavelet transform becomes

$$|\nabla W_s f_{l_d}(x_\rho, y_\rho)| = \sqrt{c_f^2 (w_1^2 + w_2^2)}. \quad (6.13)$$

In the final result of Eq. (6.13), the slope factors k_1 and k_2 disappear. That is, of course, the amplitude of the wavelet transform has the property of slope free. Consequently, the local maximum modulus of the wavelet transform also possesses this characteristic. Some graphical examples are illustrated in Fig. 6.5. Four curves, particularly, which are straight lines, with different orientations are analyzed. The original images are shown in Fig. 6.5(a). After applying the wavelet transforms to each one, the distributions of their modulus of transformations are given in Fig. 6.5(b),

and the corresponding 3-D graphical displays are drawn in Fig. 6.5(d). From this figure, it is clear that the highest peaks of all distributions are remained in same values. In other words, the local maximum modulus of the wavelet transform of the curve is independent on the slope of that curve. Fig. 6.5(c) presents the maxima of wavelet modulus for these four curves.

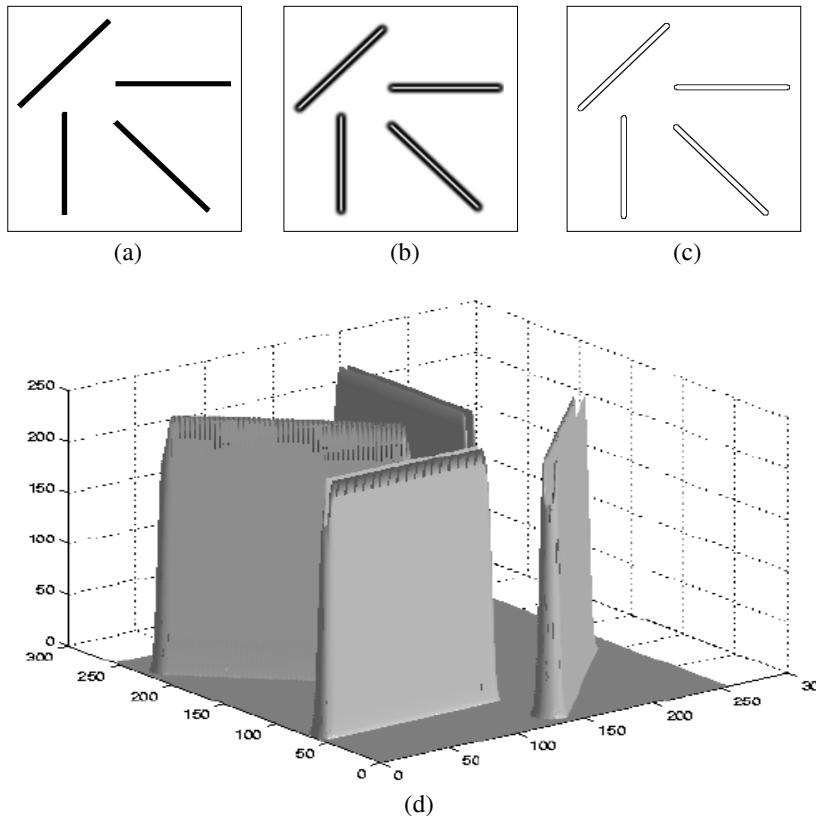


Fig. 6.5 Graphical examples of the property of the slope-invariant.

6.2.2 Grey-Level Invariant

In this sub-section, the characterization of the grey-level invariant will be verified. That means, we will prove that the maximum modulus of the wavelet transforms with respect to the curves take place at the same points

when the images with different grey-levels will be processed.

In fact, it is obviously to see that the wavelet transform $\nabla W_s f(x, y)$ is a linear system. Then, if the input function f is scaled, the output is scaled the same. That is,

$$|\nabla W_s(cf(x, y))| := |c||\nabla W_s f(x, y)|.$$

Therefore, the factor of the grey level c does not influence the location of the maximum modulus of the wavelet transform. This is the characteristic of grey-level invariant.

Some graphical examples are illustrated in Fig. 6.6. The left side of Fig. 6.6 contains two original images, which are circles. The grey-levels

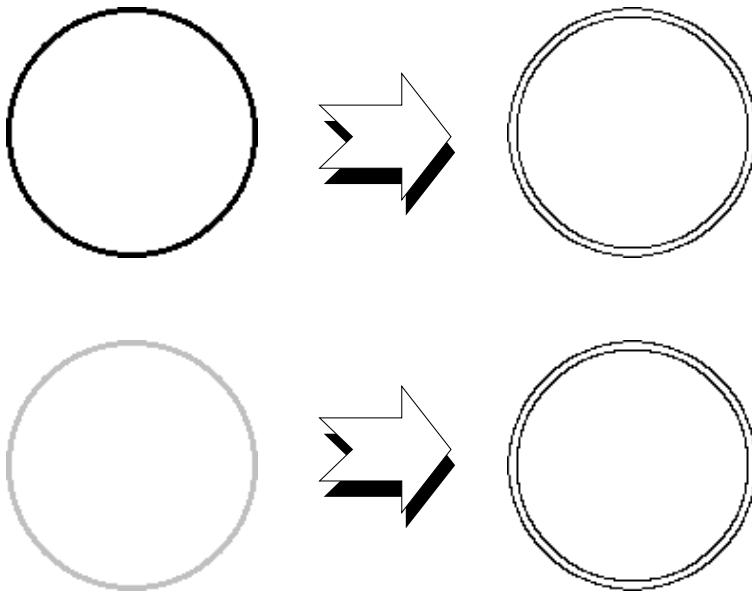


Fig. 6.6 Graphical examples of the property of the grey-level-invariant.

of them differ from each other. The same wavelet transforms are applied to these circles. The resulting maximum moduli of the wavelet transforms with respect to them are graphically displayed in the right side of Fig. 6.6. It is obvious that they are same, i.e. the double-line circles.

6.2.3 Width Light-Dependent

The property of the width light-dependent with respect to the local maximum modulus of the wavelet transforms of curves will be presented in this sub-section. It will be proved that for various widths of curves, d 's, the location of maximum modulus of the wavelet transform varies lightly. That means parameter ρ retains the same value approximately when $\rho > d$.

Note that, the quadratic spline function $\phi(r)$ is compactly supported, moreover the support is very short, so that for the point t_0 , we have

$$\begin{aligned} w_1 &= \int_{(t_0-b)/s}^{(t_0-a)/s} dt \int_{(\rho-\frac{d}{2})/s}^{(\rho+\frac{d}{2})/s} \phi'(\sqrt{t^2 + \lambda^2}) \frac{t}{\sqrt{t^2 + \lambda^2}} d\lambda \\ &= \int_{-\infty}^{\infty} dt \int_{(\rho-\frac{d}{2})/s}^{(\rho+\frac{d}{2})/s} \phi'(\sqrt{t^2 + \lambda^2}) \frac{t}{\sqrt{t^2 + \lambda^2}} d\lambda \\ &= 0. \end{aligned} \quad (6.14)$$

Substituting Eq. (6.14) into Eq. (6.13) yields

$$\begin{aligned} |\nabla W_s f_{l_d}(x_\rho, y_\rho)| &= \sqrt{c_f^2 w_2^2} = |c_f w_2| \\ &= |c_f| \left| \int_{-\infty}^{\infty} dt \int_{(\rho-\frac{d}{2})/s}^{(\rho+\frac{d}{2})/s} \phi'(\sqrt{t^2 + \lambda^2}) \frac{\lambda}{\sqrt{t^2 + \lambda^2}} d\lambda \right| \\ &= 2|c_f| \left| \int_0^{\infty} dt \int_{(\sqrt{t^2 + (\frac{\rho-d/2}{s})^2})}^{(\sqrt{t^2 + (\frac{\rho+d/2}{s})^2})} \phi'(\lambda) d\lambda \right|. \end{aligned}$$

Since $\phi'(\lambda) \leq 0$ ($\forall \lambda \geq 0$), we get

$$\begin{aligned} |\nabla W_s f_{l_d}(x_\rho, y_\rho)| &= -2|c_f| \int_0^{\infty} dt \int_{(\sqrt{t^2 + (\frac{\rho-d/2}{s})^2})}^{(\sqrt{t^2 + (\frac{\rho+d/2}{s})^2})} \phi'(\lambda) d\lambda \\ &= -2|c_f| \int_0^{\infty} [\phi(\sqrt{t^2 + (\frac{\rho+d/2}{s})^2}) \\ &\quad - \phi(\sqrt{t^2 + (\frac{\rho-d/2}{s})^2})] dt. \end{aligned}$$

To find the parameter ρ such that $|\nabla W_s f_{l_d}(x_\rho, y_\rho)|$ reaches the local maximum, we consider the derivative on ρ :

$$\begin{aligned}
& \frac{d}{d\rho} |\nabla W_s f_{l_d}(x_\rho, y_\rho)| = \\
& -2|c_f| \int_0^\infty [\phi' \left(\sqrt{t^2 + (\frac{\rho+d/2}{s})^2} \right) \\
& \frac{\left(\frac{\rho+d/2}{s} \right) \frac{1}{s}}{\sqrt{t^2 + \left(\frac{\rho+d/2}{s} \right)^2}} - \\
& \phi' \left(\sqrt{t^2 + (\frac{\rho-d/2}{s})^2} \right) \frac{\left(\frac{\rho-d/2}{s} \right) \frac{1}{s}}{\sqrt{t^2 + \left(\frac{\rho-d/2}{s} \right)^2}}] dt.
\end{aligned}$$

Let $\frac{d}{d\rho} |\nabla W_s f_{l_d}(x_\rho, y_\rho)| = 0$, we get

$$\begin{aligned}
& \left(\rho + \frac{d}{2} \right) \int_0^\infty \frac{\phi' \left(\sqrt{t^2 + (\frac{\rho+d/2}{s})^2} \right)}{\sqrt{t^2 + \left(\frac{\rho+d/2}{s} \right)^2}} dt = \\
& \left(\rho - \frac{d}{2} \right) \int_0^\infty \frac{\phi' \left(\sqrt{t^2 + (\frac{\rho-d/2}{s})^2} \right)}{\sqrt{t^2 + \left(\frac{\rho-d/2}{s} \right)^2}} dt. \tag{6.15}
\end{aligned}$$

To facilitate solving parameter ρ , Eq. (6.15) can be rewritten by

$$G \left(\frac{\rho + \frac{d}{2}}{s} \right) = G \left(\frac{\rho - \frac{d}{2}}{s} \right) \tag{6.16}$$

according to

$$G(\xi) := \xi \int_0^\infty \frac{\phi'(\sqrt{t^2 + \xi^2})}{\sqrt{t^2 + \xi^2}} dt. \tag{6.17}$$

Now, our question turns to find ρ , such that Eq. (6.16) will be held. In general, it is difficult and even impossible to solve ρ from Eq. (6.16) directly, for an arbitrary wavelet ϕ' , because of the complexity of (6.17). To simplify the estimation of $G(\xi)$ without losing its characterization for the Dirac-structure edges, we utilize the quadratic spline function (6.7). First of all, we calculate $G(\xi)$, and the result is presented below:

(1). If $\xi \geq 1$, it is easy to see that

$$G(\xi) = 0. \quad (6.18)$$

(2). If $\frac{1}{2} \leq \xi < 1$, $G(\xi)$ is given by

$$\begin{aligned} G(\xi) &= \xi \int_0^{\sqrt{1-\xi^2}} \frac{\phi'(\sqrt{t^2 + \xi^2})}{\sqrt{t^2 + \xi^2}} dt \\ &= -8\xi \int_0^{\sqrt{1-\xi^2}} \frac{(\sqrt{t^2 + \xi^2} - 1)^2}{\sqrt{t^2 + \xi^2}} dt \\ &= -8\xi \left[-\frac{3}{2}\sqrt{1-\xi^2} - \frac{2+\xi^2}{2} \log \xi + \frac{2+\xi^2}{2} \log(1 + \sqrt{1-\xi^2}) \right] \\ &= -8\xi \left[-\frac{3}{2}\sqrt{1-\xi^2} + \frac{2+\xi^2}{2} \log \frac{1 + \sqrt{1-\xi^2}}{\xi} \right]. \end{aligned} \quad (6.19)$$

(3). If $0 \leq \xi < \frac{1}{2}$, we obtain

$$\begin{aligned} G(\xi) &= \xi \int_0^{\sqrt{\frac{1}{4}-\xi^2}} + \int_{\sqrt{\frac{1}{4}-\xi^2}}^{\sqrt{1-\xi^2}} \\ &= 8\xi \int_0^{\sqrt{\frac{1}{4}-\xi^2}} \frac{3(t^2 + \xi^2) - 2\sqrt{t^2 + \xi^2}}{\sqrt{t^2 + \xi^2}} dt \\ &\quad - 8\xi \int_{\sqrt{\frac{1}{4}-\xi^2}}^{\sqrt{1-\xi^2}} \frac{(\sqrt{t^2 + \xi^2} - 1)^2}{\sqrt{t^2 + \xi^2}} dt \\ &= 4\xi \left[-3\sqrt{1-4\xi^2} + 3\sqrt{1-\xi^2} - 3\xi^2 \log \xi \right. \\ &\quad \left. + 2 \log \frac{1 + \sqrt{1-4\xi^2}}{2} \right. \\ &\quad \left. + 4\xi^2 \log \frac{1 + \sqrt{1-4\xi^2}}{2} - 2 \log(1 + \sqrt{1-\xi^2}) \right. \\ &\quad \left. - \xi^2 \log(1 + \sqrt{1-\xi^2}) \right]. \end{aligned} \quad (6.20)$$

The above mathematical formula are too complicated, so that it is difficult to solve them. As an alternate, in this chapter, a graphic-solution method will be developed.

The graphical description of $G(\xi)$ on $[0, \infty)$ is illustrated in Fig. 6.7, based on Eqs. (6.18), (6.19) and (6.20). The value of the parameter ρ can be solved from this figure approximately.

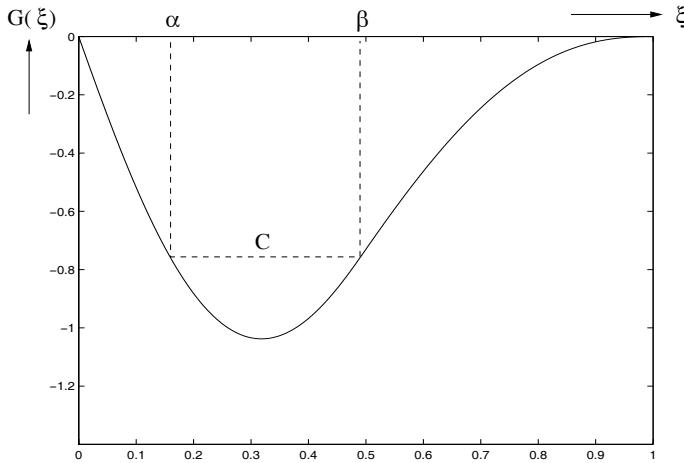


Fig. 6.7 Using the graphical description of $G(\xi)$ to find the value of the parameter ρ approximately.

By Fig. 6.7, it is easy to see that, for a constant C , $0 < C \leq 1$, there exists a pair of numbers (α, β) satisfying $0 \leq \alpha < \beta \leq 1$, such that

$$\begin{cases} \beta - \alpha = C \\ G(\alpha) = G(\beta). \end{cases} \quad (6.21)$$

It is obvious that both α and β depend on only the constant C , that means they are functions on C . Therefore, α and β can be denoted as

$$\alpha = \alpha(C), \quad \beta = \beta(C).$$

In the following, we will solve ρ satisfying Eq. (6.16). Comparing Eq. (6.16) with Eq. (6.21) yields

$$\frac{\rho + d/2}{s} = \beta \quad (6.22)$$

$$\frac{\rho - d/2}{s} = \alpha. \quad (6.23)$$

Adding Eq. (6.22) and Eq. (6.23), we obtain

$$\rho = \frac{s}{2}(\alpha + \beta). \quad (6.24)$$

By subtracting Eq. (6.23) from Eq. (6.22), we have

$$\beta - \alpha = \frac{d}{s}. \quad (6.25)$$

Substituting Eq. (6.25) into Eq. (6.24) produces

$$\rho = \frac{s}{2} \left[\alpha\left(\frac{d}{s}\right) + \beta\left(\frac{d}{s}\right) \right] = \rho(s, d). \quad (6.26)$$

Thus, ρ is dependent on s and d . That means ρ is a function on s and d . The relationship among ρ , s and d can be displayed graphically in Fig. 6.8.

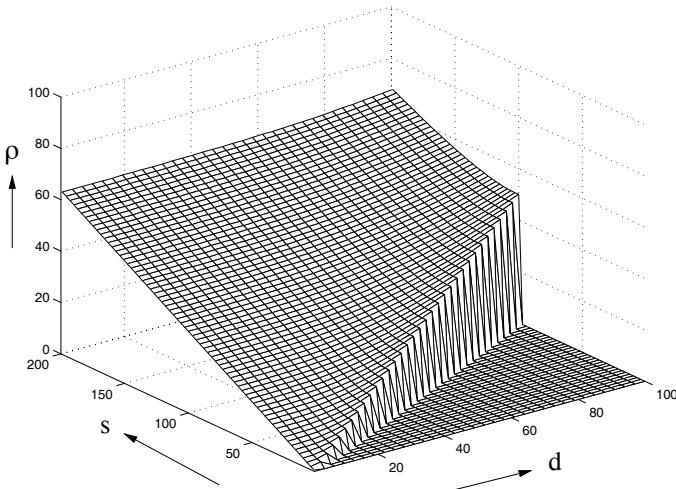


Fig. 6.8 The graphical description of $\rho = \rho(s, d)$.

From Fig. 6.8, it is easy to understand that the dependence of ρ on d is very light under the condition of $0 < \frac{d}{s} \leq 1$, i.e. $s \geq d > 0$. The larger the ratio of $\frac{d}{s}$, the smaller the dependence of ρ on d is. When the scale s of wavelet transform is large enough, the parameter ρ is independent on the width d of the curve image.

Fig. 6.9 displays the relationship between ρ and d exactly, when some specific transform scales are given, for instance $s = 4, 8, 12, 16$.

Another way to view Eq. (6.26) and Fig. 6.8 can also be explained in Fig. 6.10. From this figure, the following facts can be fund: For fixed d 's,

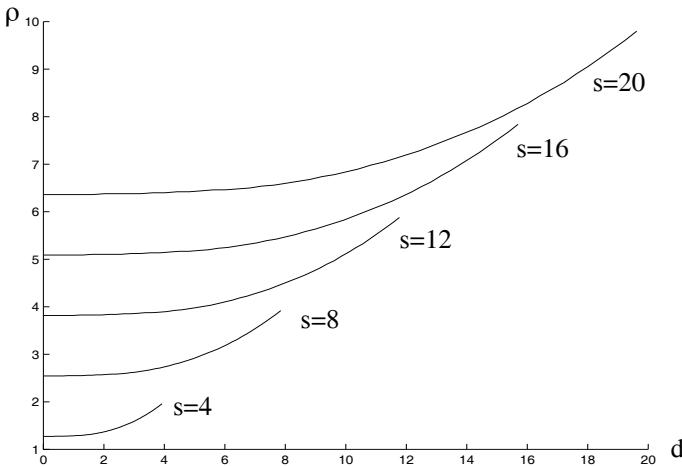


Fig. 6.9 The relationship between ρ and d , when some specific scales are given ($s = 4, 8, 12, 16$).

the dependence of ρ on s is closed to the linear function, when the scale s of wavelet transform is large enough. That means the larger the scale s , the closer to the straight line the parameter ρ is. Therefor, the relationship between ρ and d can be represented by a straight line, when the scale s of the wavelet transform has sufficient value. The equation of this straight line can approximately be fund in Fig. 6.7, i.e. the minimum of $G(\xi)$. That is $\rho \approx 0.318s$.

Some graphical examples are illustrated in Fig. 6.11. The left side of Fig. 6.11 contains two original images, which are circles. The width of them differ from each other. The same wavelet transforms are applied to these circles. The resulting maximum moduli of the wavelet transforms with respect to them are graphically displayed in the right side of Fig. 6.11. It is obvious that the same images, i.e. the double-line circles, are obtained.

6.3 Experiments

Based on the theoretical analyses in the previous sections, an algorithm to detect the Dirac-structure edges such as curves from a multi-structure-edge image is designed as follows.

Algorithm 6.1 For a multi-structure-edge image containing the Dirac-

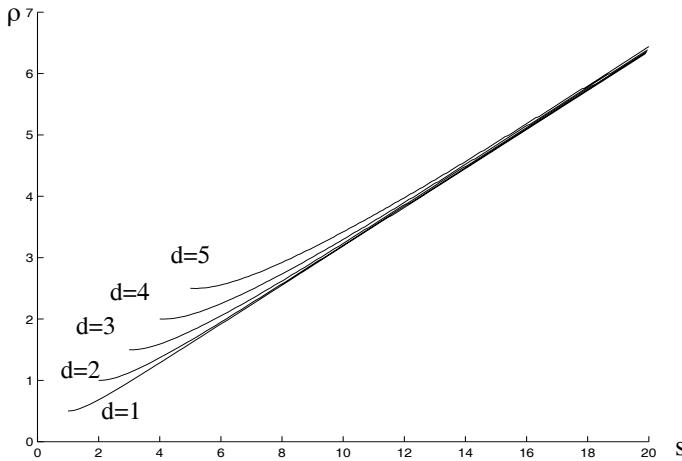


Fig. 6.10 The relationship between ρ and s , when some specific widths are given ($d = 1 - 5$).

structure edges, such as curves, and the wavelet transform scale $s > 0$, the edges can be detected by the following steps:

- step 1** Calculating all the wavelet transforms $\{W_s^1 f(x, y), W_s^2 f(x, y)\}$ under the quadratic spline wavelet .
- step 2** Calculating the local maxima f_{locmax} of $|\nabla W_s f(x, y)|$ and the gradient direction $f_{gradient}$.
- step 3** For each point (x, y) with local maximum, searching the point whose distance from (x, y) to it is $0.6424s$ along the gradient direction. If it is still a point with local maximum, the center point is detected.
- step 4** The curves are formed by all the points detected by the above steps.

In Fig. 6.12, four circles with various gray-levels and widths are tested using this method. The original images are illustrated on the left column of Fig. 6.12. After applying the steps 1 and 2 of the proposed wavelet transform algorithm to these circles, the local maximum modulus of the wavelet transform with respect to them can be computed and the results are given on the middle column in Fig. 6.12. Finally, the central lines of these circles are extracted using the steps 3 and 4 of the above algorithm, and presented on the right column in Fig. 6.12.

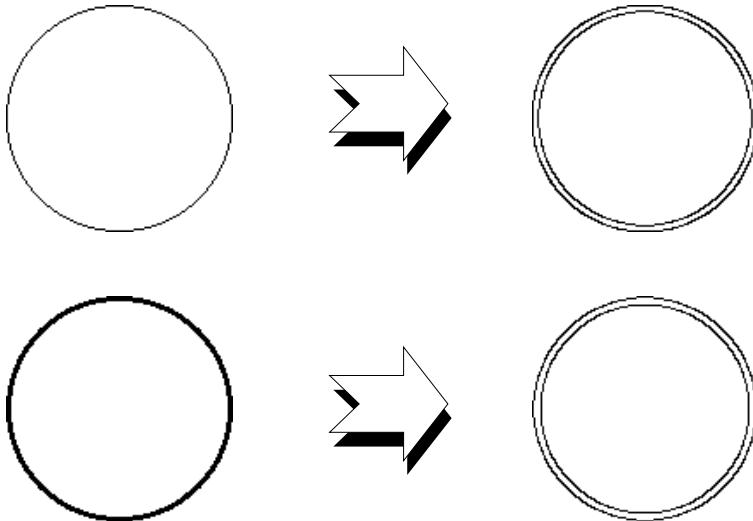


Fig. 6.11 Graphical examples of the property of the width light-invariant.

Let us look at Fig. 6.13. The particular task is that we are required to identify different structures of edges, thereafter, detect the Dirac-structure edges such as drawing lines and texts in Fig. 6.13(a) and eliminate the step-structure edges such as the contour of the aircraft in Fig. 6.13(a). Unfortunately, the algorithm based on the method proposed in [Mallat and Hwang, 1992; Mallat and Zhong, 1992] has detected all edges without identifying different structures of edges. Thus, the resulting image contains the step-structure edges such as the contours which are required to be deleted from the image. The method developed in this chapter possesses three significant characteristics, namely: (1) slope invariant: the local maximum modulus of the wavelet transform of a Dirac-structure edge is independent on the slope of the edge; (2) grey-level invariant: the local maximum modulus of the wavelet transform with respect to a Dirac-structure edge takes place at the same points when the images with different grey-levels will be processed; (3) width light-dependent: for various widths of the Dirac-structure edge images, the location of maximum modulus of the wavelet transform varies lightly under the certain circumscription. According to these characteristics and our early work [Tang et al., 1998c], we can recognize three basic structures of edges, and further, extract the Dirac-structure ones. After applying the above algorithm, new result can be found in Fig. 6.13(b). It

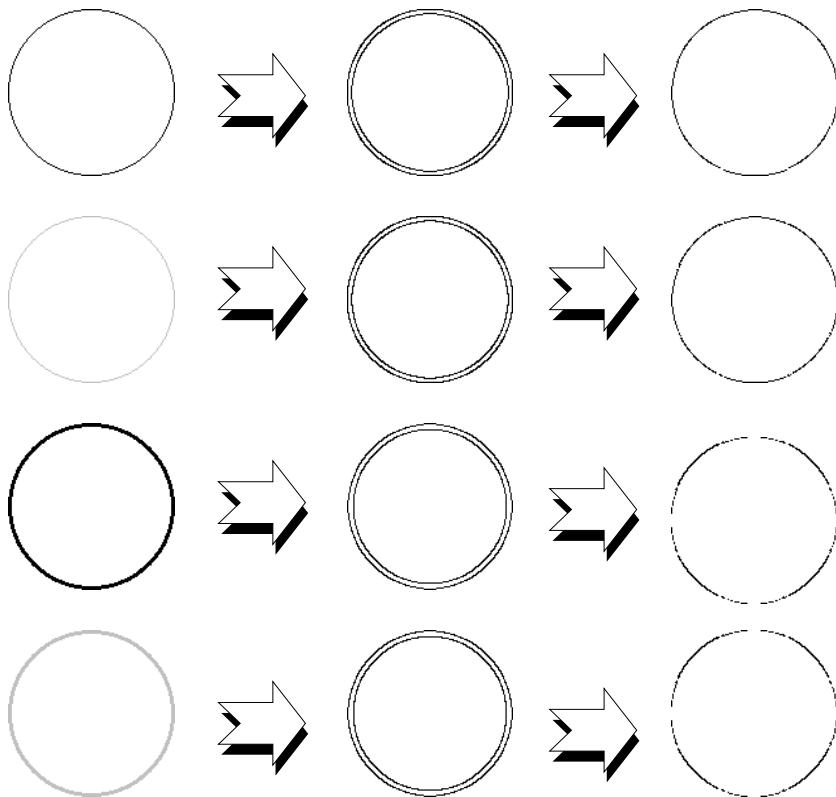


Fig. 6.12 Detection of the central lines with various gray-levels and widths by proposed wavelet transform algorithm, (the scale of the wavelet transform: $s=6$). Left column: original images, i.e. four circles with various gray-levels and widths. Middle column: the local maximum modulus of the wavelet transform. Right column: the central lines of the circles are extracted.

extracts all lines and texts, and removes the contour of the aircraft.

Now, turn to another example shown in Fig. 6.14. The original image is illustrated in Fig. 6.14(a) which has two classes of structures of edges: (1) Dirac-structure edges, i.e. three circles with different gray-levels and widths; (2) step-structure edge, that is a planner object, square, which has a contour. Upon utilizing the method proposed in this chapter, we can produce a new image as shown in Fig. 6.14(b). It contains only three circles which belong to the Dirac-structure edges, and the step-structure edge, contour of the square, is removed.

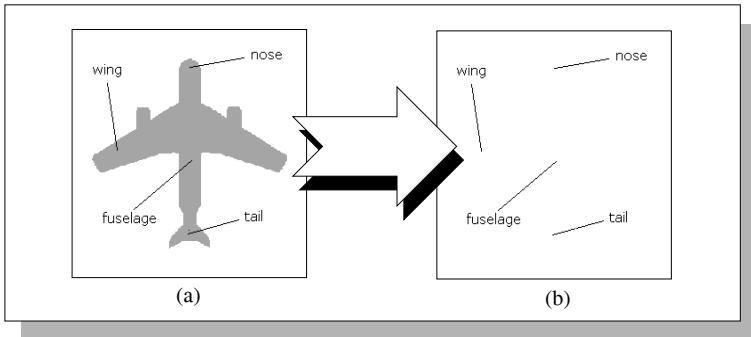


Fig. 6.13 Result using the method proposed in this chapter. (a) Original image: two classes of edges are embedded in this image, i.e. a contour of the aircraft which belongs to the step-structure edge, and some lines and texts which belong to the Dirac-structure edges. (b) The result which is obtained from the new algorithm, all lines and texts are extracted, and the contour of the aircraft is removed.

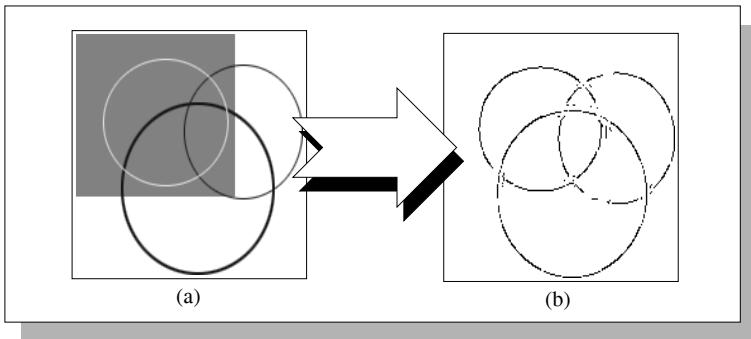


Fig. 6.14 Another result using the method proposed in this chapter (a) Original image which has two classes of structures of edges, i.e. Dirac-structure edges (three circles with different gray-levels and widths), and step-structure edge (a planer object, square which has a contour). (b) Resulting image which has only three circles, and the contour of the square is removed.

The final experiment is presented in Fig. 6.15. Original image is the image Lena plus several characters as illustrated in Fig. 6.15(a). It is a multi-structure-edge one, where several characters are embedded in it. Upon performing the steps 1 and 2 of the proposed wavelet transform algorithm to it, its local maximum modulus of the wavelet transform can be computed and the result is given in Fig. 6.15(b). Next, the charac-

ters which are embedded in the image are extracted by the steps 3 and 4 of the above algorithm, and presented in Fig. 6.15(c). Finally, the noise is removed and a fine image which contains only the characters “lena” is obtained and displays in Fig. 6.15(d).

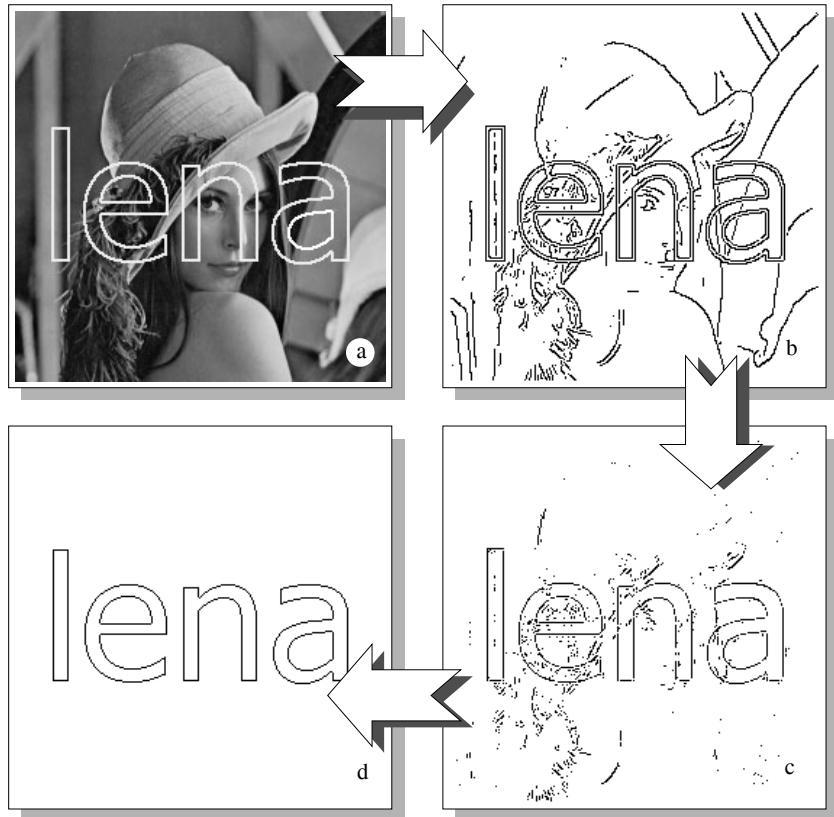


Fig. 6.15 Lines extracted from image of Lena by proposed wavelet transform algorithm (the scale of the wavelet transform: $s=6$). (a) Original image which is the image Lena with several characters. (b) Local maximum modulus of the wavelet transform of image of Lena. (c) The characters which are embedded in Lena image are extracted.

This page intentionally left blank

Chapter 7

Construction of New Wavelet Function and Application to Curve Analysis

In Chapter 6, a very important characterization of the Dirac-structure edges by wavelet transform was provided. Three significant characteristics of the local maximum modulus of the wavelet transform with respect to the Dirac-structure edges were presented, namely: (1) slope invariant: the local maximum modulus of the wavelet transform of a Dirac-structure edge is independent on the slope of the edge. (2) grey-level invariant: the local maximum modulus of the wavelet transform with respect to a Dirac-structure edge takes place at the same points when the images with different grey-levels are to be processed. (3) width light-dependent: for various widths of the Dirac-structure edge images, the location of maximum modulus of the wavelet transform varies slightly when the scale s of the wavelet transform is larger than the width d of the Dirac-structure edge images. Based on the characteristics, an algorithm to detect the Dirac-structure edges from an image has been developed. An example of applying this algorithm to detect the Dirac-structure edge can be found in Fig. 7.1. The original image shown in Fig. 7.1(a) is a multi-structure-edge one, which contains an image of the Chinese traditional word ‘book’ where several drawing curves are embedded in it. Upon performing the algorithm proposed in Chapter 6, its local maximum modulus of the wavelet transform has been computed, and is shown in Fig. 7.1(b), and the result is presented in Fig. 7.1(c). Finally, the noise is removed and a fine image which contains only the drawing curves is obtained and displayed in Fig. 7.1(d).

However, note the third property in Chapter 6, it says “width light-dependent”, does not say “width invariant”. This means that for various widths of the Dirac-structure edge images, the location of maximum mod-

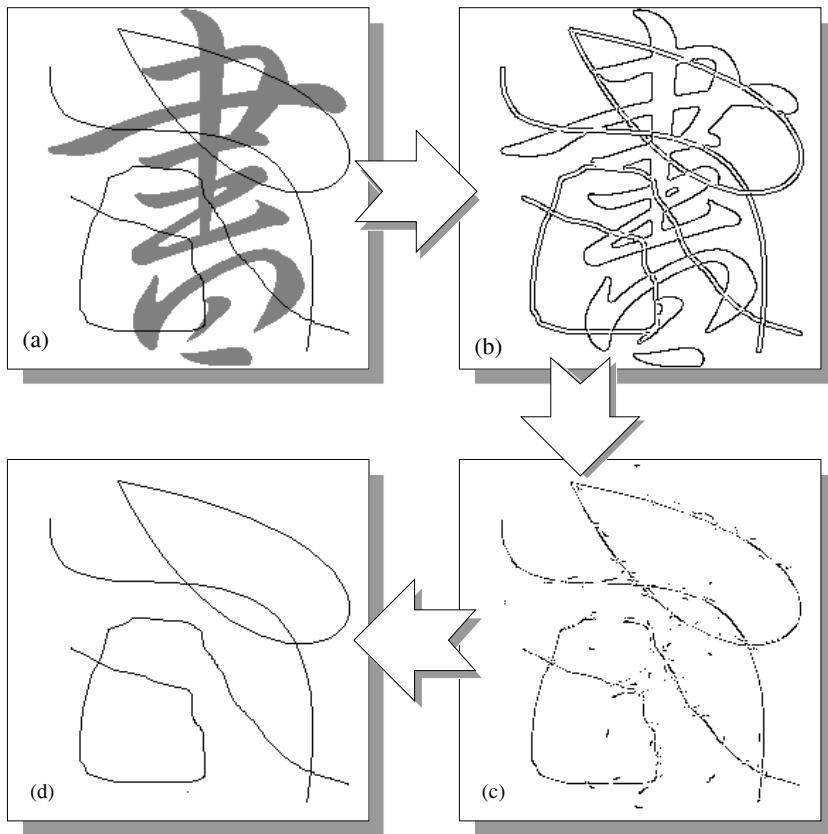


Fig. 7.1 Lines extracted from an image by proposed algorithm in Chapter 6: (a) Original image; (b) Local maximum modulus of the wavelet transform of the original image; (c) The drawing Lines extracted by the algorithm; (d) The resulting image.

ulus of the wavelet transform may change (these changes are small). What we want is that the location of maximum modulus does not change, i.e. the location of maximum modulus has the property of width invariant. Let us look at Fig. 7.2. The first row of Fig. 7.2 has three circles. The left image is the original one which contains a circle with various width. The middle one is the location of maximum modulus of the wavelet transform with scale $s = 6$, which depends on the width of the circle in some way. Finally, by utilizing the algorithm proposed in Chapter 6, the skeleton of the circle is extracted and displayed on the right of Fig. 7.2. We can find

that the skeleton of the circle is broken. The second row of Fig. 7.2 has trees, where the sizes of the branches vary, some are thick and some are thin. The left image is the original one. The middle one illustrates the location of maximum modulus of the wavelet transform with scale $s = 6$. It is clear that the location of maximum modulus of the different branches has slight changes. The right of Fig. 7.2 is the skeleton extracted utilizing the algorithm proposed in Chapter 6. It is easy to see that some branches of the tree are lost.

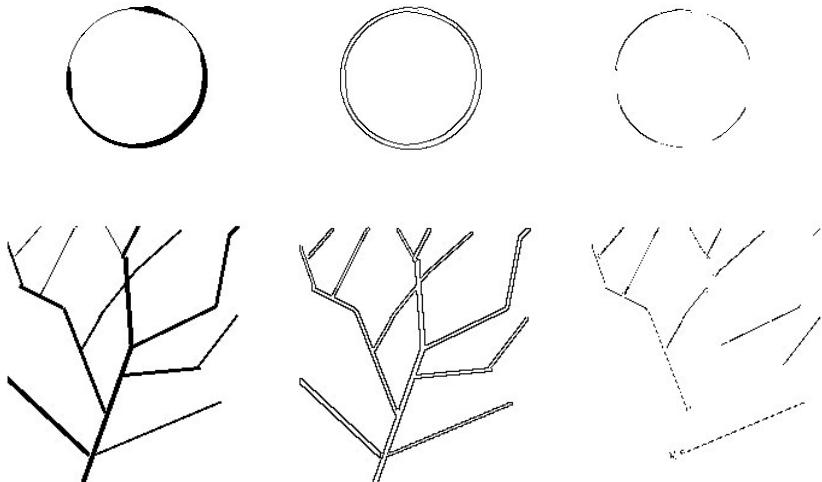


Fig. 7.2 Left: the original images; Middle: the location of maximum modulus of the wavelet transform with $s = 6$; Right: the skeleton images extracted by the algorithm of Chapter 6.

To overcome such a defect, a natural question is whether there exists a more suitable wavelet function such that the location of maximum modulus of the wavelet transform of a curve is width-invariant as well as slope invariant and grey-invariant.

In this chapter, we will present a development [Yang et al., 2003a], where a novel wavelet is constructed, so that the above “width light-dependent” properties can be improved to “width invariant” without losing the “slope invariant” and “grey-level invariant”. Due to this improvement, the detection of Dirac-structure edge is more accurate.

7.1 Construction of New Wavelet Function — Tang-Yang Wavelet

The scale wavelet transform described in Eq. (6.2) is a filter, and since $\psi \in L^2(R^2)$, its Fourier transform can be defined by

$$\hat{\psi}(\xi, \eta) := \int_R \int_R \psi(x, y) e^{-i(\xi x + \eta y)} dx dy$$

which satisfies the condition of $\hat{\psi} \in L^2(R^2)$. Thus, both functions ψ and $\hat{\psi}$ decrease at infinity. Usually, we choose $\psi \in L^2(R^2)$ such that both ψ and $\hat{\psi}$ decrease at infinity with exponents at least. We know that $\hat{\psi}(0, 0) = 0$ from Eq. (6.1), which implies that the scale wavelet transform $W_s f(x, y)$ corresponds to a band-pass filter essentially. Moreover, since ψ and $\hat{\psi}$ decrease at infinity, $W_s f(x, y)$ characterizes the local properties of f both on the time-domain and frequency domain. In fact, two points are considered, namely:

- On one hand, due to the fast decrease of ψ at infinity, formula (6.2) characterizes the local property of f around point (x, y) . To understand this, one can consider that ψ vanishes outside $[-1, 1]$. In this case, we have

$$\psi\left(\frac{x-u}{s}, \frac{y-v}{s}\right) = 0$$

for every (u, v) satisfying:

$$-1 \leq \frac{x-u}{s} \leq 1 \quad \text{or} \quad -1 \leq \frac{y-v}{s} \leq 1,$$

i.e.

$$x-s \leq u \leq x+s \quad \text{or} \quad y-s \leq v \leq y+s.$$

This concludes that

$$\begin{aligned} W_s f(x, y) &= \int_R \int_R f(u, v) \frac{1}{s^2} \psi\left(\frac{x-u}{s}, \frac{y-v}{s}\right) du dv \\ &= \int_{x-s}^{x+s} \int_{y-s}^{y+s} f(u, v) \frac{1}{s^2} \psi\left(\frac{x-u}{s}, \frac{y-v}{s}\right) du dv, \end{aligned}$$

which tells us that $W_s f(x, y)$ characterizes the local properties of f on the interval $[x-s, x+s] \times [y-s, y+s]$ centering at (x, y) . The smaller s is, the narrower the interval, i.e. the time-window, is.

- On the other hand, based on the basic properties of Fourier transform, we can obtain

$$\begin{aligned}
 (W_s f)(x, y) &= \left(\frac{1}{2\pi}\right)^2 \int_R \int_R (W_s f)(\xi, \eta) e^{i(\xi x + \eta y)} d\xi d\eta \\
 &= \left(\frac{1}{2\pi}\right)^2 \int_R \int_R \hat{f}(\xi, \eta) (\psi_s)(\xi, \eta) e^{i(\xi x + \eta y)} d\xi d\eta \\
 &= \left(\frac{1}{2\pi}\right)^2 \int_R \int_R \hat{f}(\xi, \eta) \hat{\psi}(s\xi, s\eta) e^{i(\xi x + \eta y)} d\xi d\eta.
 \end{aligned}$$

Since $\hat{\psi}(s\xi, s\eta) e^{i(\xi x + \eta y)}$ decreases as $\hat{\psi}(s\xi, s\eta)$ does at infinity, thus $(W_s f)(x, y)$ characterizes the local property of \hat{f} . The position of the locality depends on scale s and function $\hat{\psi}$. The larger s is, the narrower the frequency-window is. A detail analysis can be done similarly to the above.

Theoretically, Eq. (6.1), i.e. $\hat{\psi}(0, 0) = 0$, implies that $\psi(x, y)$ is a band-pass filter, but a high-pass one because of the decrease of its Fourier transform at infinity. It is easy to see that the partial derivatives of a low-pass function can become the candidates of the wavelet functions. Here, we consider such kind of wavelets, i.e.,

$$\psi^1(x, y) := \frac{\partial}{\partial x} \theta(x, y), \quad \psi^2(x, y) := \frac{\partial}{\partial y} \theta(x, y)$$

where $\theta(x, y)$ denotes a real function satisfying:

- $\theta(x, y)$ fast decreases at infinity;
- $\theta(x, y)$ is an even function on both x and y .
- $\hat{\theta}(0, 0) = 1$.

For wavelet $\psi^1(x, y)$ defined above, its scale wavelet transform is

$$\begin{aligned}
 W_s^1 f(x, y) &= (f * \psi_s^1)(x, y) \\
 &= (f * s \frac{\partial}{\partial x} \theta_s)(x, y) \\
 &= s \frac{\partial}{\partial x} (f * \theta_s)(x, y)
 \end{aligned}$$

where $\theta_s(x, y) := \frac{1}{s^2} \theta(\frac{x}{s}, \frac{y}{s})$.

This formula tells us that $W_s^1 f(x, y)$ is essentially the derivative of the smoothness function along the horizontal axis and then the local maxima

of the derivative function correspond to the points of the smoothness image with sharp variation along the horizontal axis. It is equivalent to the classical multi-scale edge detection [Canny, 1986; Marr and Hildreth, 1980], if $\theta(x, y)$ is set to be a Gaussian, which is defined by

$$G_\sigma(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right). \quad (7.1)$$

That is why $\theta(x, y)$ is assumed to satisfy $\hat{\theta}(0, 0) = 1$.

A similar explanation for wavelet $\psi^2(x, y)$ defined above can be made. However, the partial derivative is along the vertical direction instead of the horizontal one.

Gaussian function has been employing in image processing. It possesses some excellent properties, such as, the locality in both the time domain and frequency domain, the same widths in both the time-window and frequency-window, and so on. All these properties make it applied extensively and deeply in the area of the filtering, and it already almost becomes the best candidate of low-pass filter in practice. Unfortunately, Gaussian function is not always the best one for all applications. In fact, we have shown that it is not the best candidate for characterizing a Dirac-structure edge [Tang et al., 2000]. Even the quadratic spline wavelet is better than it, although, the quadratic spline wavelet is still not a perfect one for such applications. In [Tang et al., 2000] it has been proved that the location of maximum modulus of the wavelet transform with respect to a Dirac-structure edge is not width invariant. It still depends on the width of the edge even though it depends lightly. To avoid such dissatisfaction, a new wavelet called Tang-Yang wavelet is constructed, and its definition is described below [Yang et al., 2003a].

Let

$$\begin{cases} \psi_1(x) = -\frac{2}{\pi}(-8x \ln \frac{1+\sqrt{1-16x^2}}{4x} + \frac{1}{2x}\sqrt{1-16x^2}) \\ \psi_2(x) = -\frac{2}{\pi}(8x \ln \frac{3+\sqrt{9-16x^2}}{4x} - \frac{3}{2x}\sqrt{9-16x^2}) \\ \psi_3(x) = -\frac{2}{\pi}(-4x \ln \frac{1+\sqrt{1-x^2}}{x} + \frac{4}{x}\sqrt{1-x^2}) \end{cases}$$

Then, the 1-D wavelet $\psi(x)$ is an odd function defined on $(0, \infty)$ by

$$\psi(x) := \begin{cases} \psi_1(x) + \psi_2(x) + \psi_3(x) & x \in (0, \frac{1}{4}) \\ \psi_2(x) + \psi_3(x) & x \in [\frac{1}{4}, \frac{3}{4}) \\ \psi_3(x) & x \in [\frac{3}{4}, 1) \\ 0 & x \in [1, \infty) \end{cases} \quad (7.2)$$

Let $\phi(x) := \int_0^x \psi(t)dt$. Then $\phi(x)$ is an even function, compactly supported on $[-1, 1]$, and $\phi'(x) = \psi(x)$. Fig. 7.3 displays functions $\psi(x)$ and $\phi(x)$ graphically.

The smoothness function $\theta(x, y)$ is then defined by

$$\theta(x, y) := \phi(\sqrt{x^2 + y^2}),$$

which is graphically shown in Fig. 7.4, and the 2-D wavelets are defined by

$$\begin{cases} \psi^1(x, y) := \frac{\partial}{\partial x} \theta(x, y) = \phi'(\sqrt{x^2 + y^2}) \frac{x}{\sqrt{x^2 + y^2}} \\ \psi^2(x, y) := \frac{\partial}{\partial y} \theta(x, y) = \phi'(\sqrt{x^2 + y^2}) \frac{y}{\sqrt{x^2 + y^2}} \end{cases} \quad (7.3)$$

and are illustrated in Fig. 7.5.

The gradient direction and the amplitude of the wavelet transform are denoted respectively by

$$\nabla W_s f(x, y) := \begin{pmatrix} W_s^1 f(x, y) \\ W_s^2 f(x, y) \end{pmatrix}, \quad (7.4)$$

and

$$|\nabla W_s f(x, y)| := \sqrt{|W_s^1 f(x, y)|^2 + |W_s^2 f(x, y)|^2}. \quad (7.5)$$

By locating the local maxima of $|\nabla W_s f(x, y)|$ we can detect the edges of the images.

7.2 Characterization of Curves through New Wavelet Transform

Three significant characteristics of the local maximum modulus of the wavelet transform with respect to the Dirac-structure edges in images will be held, namely:

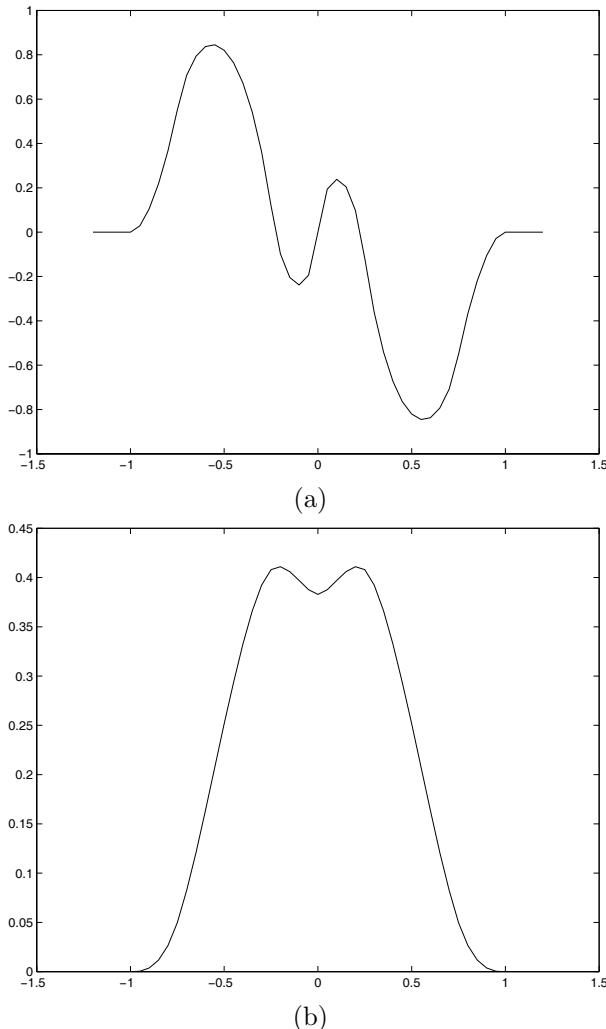


Fig. 7.3 The graphical descriptions of $\psi(x)$ and $\phi(x)$: (a) the graphical descriptions of function $\psi(x)$; (b) its primitive function $\phi(x)$.

- Grey-level invariant: the local maximum modulus of the wavelet transform with respect to a Dirac-structure edge takes place at the same points when the images with different grey-levels are to be processed.

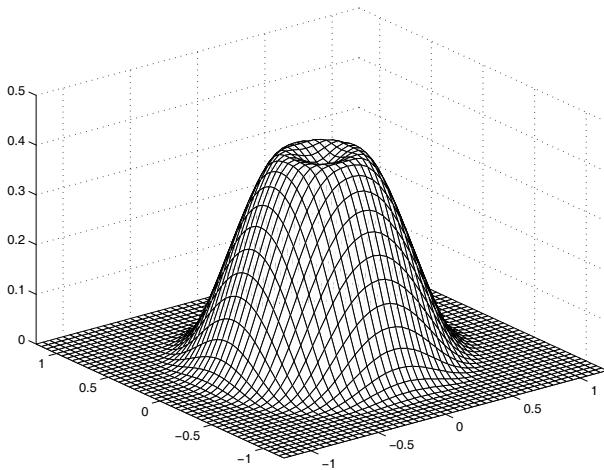


Fig. 7.4 The graphical descriptions of function $\theta(u, v)$.

- Slope invariant: the local maximum modulus of the wavelet transform of a Dirac-structure edge is independent on the slope of the edge.
- Width invariant: for various widths of the Dirac-structure edges in an image, the location of maximum modulus of the wavelet transform does not vary under certain circumstance.

The discussion of the Grey-level invariant and slope invariant is the same as that in Chapter 6. In the following we present the characteristic of the width invariant.

In this section. It will be proved that for various widths of curves, d 's, the location of maximum modulus of the wavelet transform does not vary. It means that parameter ρ retains the same value when $\rho \geq d$ and the distance from the central line is just $\rho = \frac{s}{2}$.

Now that $\phi(r)$ is compactly supported and the support is very short, for the point t_0 far away from a and b , we have

$$\begin{aligned} w_1 &= \int_{(t_0-b)/s}^{(t_0-a)/s} dt \int_{(\rho-\frac{d}{2})/s}^{(\rho+\frac{d}{2})/s} \phi'(\sqrt{t^2 + \lambda^2}) \frac{t}{\sqrt{t^2 + \lambda^2}} d\lambda \\ &= \int_{-\infty}^{\infty} dt \int_{(\rho-\frac{d}{2})/s}^{(\rho+\frac{d}{2})/s} \phi'(\sqrt{t^2 + \lambda^2}) \frac{t}{\sqrt{t^2 + \lambda^2}} d\lambda = 0, \end{aligned}$$

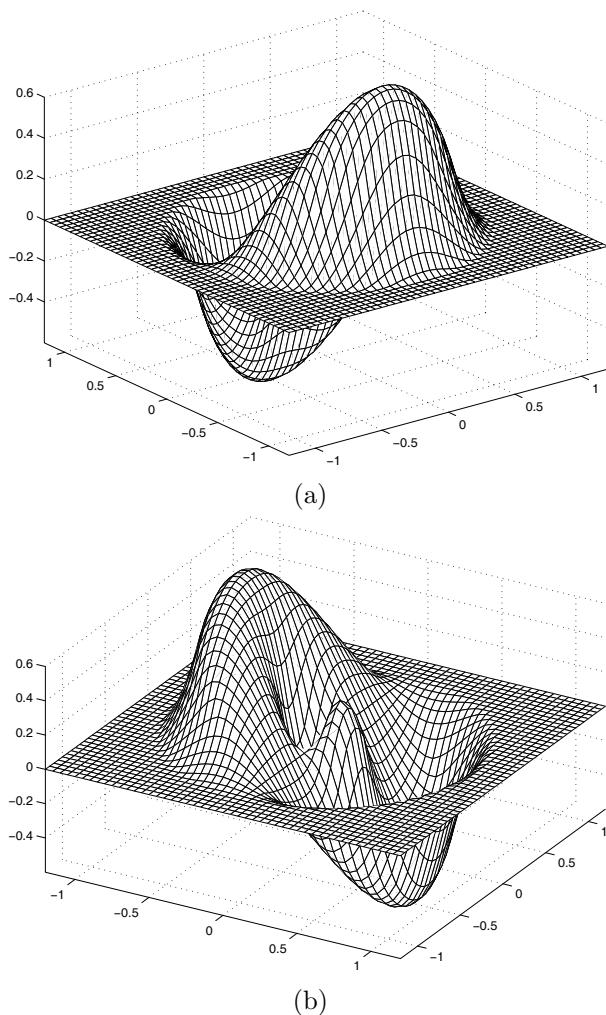


Fig. 7.5 The graphical descriptions of 2-D wavelet functions: (a) Function $\psi^1(x, y)$; (b) Function $\psi^2(x, y)$.

which concludes that

$$\begin{aligned} & |\nabla W_s f_{l_d}(x_\rho, y_\rho)| \\ &= \sqrt{c_f^2 w_2^2} = |c_f w_2| \end{aligned}$$

$$\begin{aligned}
&= |c_f| \left| \int_{-\infty}^{\infty} dt \int_{(\rho - \frac{d}{2})/s}^{(\rho + \frac{d}{2})/s} \phi'(\sqrt{t^2 + \lambda^2}) \frac{\lambda}{\sqrt{t^2 + \lambda^2}} d\lambda \right| \\
&= 2|c_f| \left| \int_0^{\infty} dt \int_{(\sqrt{t^2 + (\frac{\rho+d/2}{s})^2})}^{(\sqrt{t^2 + (\frac{\rho-d/2}{s})^2})} \phi'(\lambda) d\lambda \right| \\
&= 2 \left| c_f \int_0^{\infty} [\phi(\sqrt{t^2 + (\frac{\rho+d/2}{s})^2}) - \phi(\sqrt{t^2 + (\frac{\rho-d/2}{s})^2})] dt \right|.
\end{aligned}$$

To find ρ such that $|\nabla W_s f_{l_d}(x_\rho, y_\rho)|$ reaches the local maximum, we consider its derivative on ρ . We denote

$$G(x) := x \int_0^{\infty} \frac{\phi'(\sqrt{t^2 + x^2})}{\sqrt{t^2 + x^2}} dt = x \int_0^{\infty} \frac{\psi(\sqrt{t^2 + x^2})}{\sqrt{t^2 + x^2}} dt, \quad (7.6)$$

then have

$$\frac{\partial}{\partial x} \int_0^{\infty} \phi(\sqrt{t^2 + x^2}) dt = G(x).$$

If $G(x) \leq 0$, we have

$$\begin{aligned}
|\nabla W_s f_{l_d}(x_\rho, y_\rho)| &= -2|c_f| \int_0^{\infty} [\phi(\sqrt{t^2 + (\frac{\rho+d/2}{s})^2}) \\
&\quad - \phi(\sqrt{t^2 + (\frac{\rho-d/2}{s})^2})] dt. \quad (7.7)
\end{aligned}$$

Therefore

$$\begin{aligned}
&\frac{d}{d\rho} |\nabla W_s f_{l_d}(x_\rho, y_\rho)| \\
&= -2|c_f| \int_0^{\infty} [\phi' \left(\sqrt{t^2 + (\frac{\rho+d/2}{s})^2} \right) \frac{\left(\frac{\rho+d/2}{s} \right) \frac{1}{s}}{\sqrt{t^2 + \left(\frac{\rho+d/2}{s} \right)^2}} \\
&\quad - \phi' \left(\sqrt{t^2 + (\frac{\rho-d/2}{s})^2} \right) \frac{\left(\frac{\rho-d/2}{s} \right) \frac{1}{s}}{\sqrt{t^2 + \left(\frac{\rho-d/2}{s} \right)^2}}] dt \\
&= -2|c_f| \frac{1}{s} \left[G \left(\frac{\rho}{s} + \frac{d}{2s} \right) - G \left(\frac{\rho}{s} - \frac{d}{2s} \right) \right].
\end{aligned}$$

To guarantee that $|\nabla W_s f_{l_d}(x_\rho, y_\rho)|$ reaches the local maximum, we let $\frac{d}{d\rho} |\nabla W_s f_{l_d}(x_\rho, y_\rho)| = 0$, and we can get

$$G\left(\frac{\rho}{s} + \frac{d}{2s}\right) = G\left(\frac{\rho}{s} - \frac{d}{2s}\right). \quad (7.8)$$

By Eq. (7.7), it is easily seen that ρ must satisfy the following condition:

$$\left| \frac{\rho}{s} - \frac{d}{2s} \right| \leq 1, \quad \left| \frac{\rho}{s} + \frac{d}{2s} \right| \leq 1, \quad (7.9)$$

so that $|\nabla W_s f_{l_d}(x_\rho, y_\rho)|$ reaches the local maximum.

Hence, our question turns to solve ρ satisfying Eqs. (7.8) and (7.9). To do this, we can evaluate $G(x)$ first. For ψ defined by Eq. (8.1), it is clear that

(1). For $x \in (0, \frac{1}{4})$,

$$\begin{aligned} G(x) &= x \int_x^{\frac{1}{4}} \frac{\psi_1(t) + \psi_2(t) + \psi_3(t)}{\sqrt{t^2 - x^2}} dt \\ &\quad + x \int_{\frac{1}{4}}^{\frac{3}{4}} \frac{\psi_2(t) + \psi_3(t)}{\sqrt{t^2 - x^2}} dt + x \int_{\frac{3}{4}}^1 \frac{\psi_3(t)}{\sqrt{t^2 - x^2}} dt \\ &= x \int_x^{\frac{1}{4}} \frac{\psi_1(t)}{\sqrt{t^2 - x^2}} dt + x \int_x^{\frac{3}{4}} \frac{\psi_2(t)}{\sqrt{t^2 - x^2}} dt + x \int_x^1 \frac{\psi_3(t)}{\sqrt{t^2 - x^2}} dt; \end{aligned}$$

(2). For $x \in [\frac{1}{4}, \frac{3}{4})$,

$$\begin{aligned} G(x) &= x \int_x^{\frac{3}{4}} \frac{\psi_2(t) + \psi_3(t)}{\sqrt{t^2 - x^2}} dt + x \int_{\frac{3}{4}}^1 \frac{\psi_3(t)}{\sqrt{t^2 - x^2}} dt \\ &= x \int_x^{\frac{3}{4}} \frac{\psi_2(t)}{\sqrt{t^2 - x^2}} dt + x \int_x^1 \frac{\psi_3(t)}{\sqrt{t^2 - x^2}} dt; \end{aligned}$$

(3). For $x \in [\frac{3}{4}, 1)$,

$$G(x) = x \int_x^1 \frac{\psi_3(t)}{\sqrt{t^2 - x^2}} dt.$$

Thus, our problem now turns to evaluate:

$$\begin{cases} J_1(x) := x \int_x^{\frac{1}{4}} \frac{\psi_1(t)}{\sqrt{t^2 - x^2}} dt, \\ J_2(x) := x \int_x^{\frac{3}{4}} \frac{\psi_2(t)}{\sqrt{t^2 - x^2}} dt, \\ J_3(x) := x \int_x^1 \frac{\psi_3(t)}{\sqrt{t^2 - x^2}} dt. \end{cases}$$

According to the definition of ψ described by Eq. (8.1), we have

$$\begin{aligned} -\frac{\pi}{2} J_1(x) &= x \int_x^{\frac{1}{4}} \frac{1}{\sqrt{t^2 - x^2}} \left(-8t \ln \frac{1 + \sqrt{1 - 16t^2}}{4t} + \frac{\sqrt{1 - 16t^2}}{2t} \right) dt \\ &= -4x \int_x^{\frac{1}{4}} \frac{1}{\sqrt{t^2 - x^2}} \ln \frac{1 + \sqrt{1 - 16t^2}}{4t} dt^2 \\ &\quad + x \int_x^{\frac{1}{4}} \frac{\sqrt{1 - 16t^2}}{4t^2 \sqrt{t^2 - x^2}} dt^2 \\ &= -8x \int_x^{\frac{1}{4}} \ln \frac{1 + \sqrt{1 - 16t^2}}{4t} d\sqrt{t^2 - x^2} + \frac{x}{4} \int_{x^2}^{\frac{1}{16}} \frac{\sqrt{1 - 16t}}{t \sqrt{t - x^2}} dt \\ &= -8x \left[\sqrt{t^2 - x^2} \ln \frac{1 + \sqrt{1 - 16t^2}}{4t} \Big|_{t=x}^{\frac{1}{4}} + \int_x^{\frac{1}{4}} \frac{\sqrt{t^2 - x^2}}{t \sqrt{1 - 16t^2}} dt \right] \\ &\quad + \frac{x}{4} \int_{x^2}^{\frac{1}{16}} \frac{\sqrt{1 - 16t}}{t \sqrt{t - x^2}} dt \\ &= -4x \int_x^{\frac{1}{4}} \frac{\sqrt{t^2 - x^2}}{t^2 \sqrt{1 - 16t^2}} dt^2 + \frac{x}{4} \int_{x^2}^{\frac{1}{16}} \frac{\sqrt{1 - 16t}}{t \sqrt{t - x^2}} dt \\ &= -4x \int_{x^2}^{\frac{1}{16}} \frac{\sqrt{t - x^2}}{t \sqrt{1 - 16t}} dt + \frac{x}{4} \int_{x^2}^{\frac{1}{16}} \frac{\sqrt{1 - 16t}}{t \sqrt{t - x^2}} dt \\ &= -\frac{x}{4} \int_{x^2}^{\frac{1}{16}} \frac{1}{t} \left(\frac{16\sqrt{t - x^2}}{\sqrt{1 - 16t}} - \frac{\sqrt{1 - 16t}}{\sqrt{t - x^2}} \right) dt. \end{aligned}$$

Hence

$$J_1(x) = \frac{x}{2\pi} \int_{x^2}^{\frac{1}{16}} \frac{1}{t} \left(\frac{16\sqrt{t - x^2}}{\sqrt{1 - 16t}} - \frac{\sqrt{1 - 16t}}{\sqrt{t - x^2}} \right) dt.$$

$J_2(x)$ and $J_3(x)$ can be calculated similarly. We omit the details here, and give the results as follows:

$$J_2(x) = -\frac{3x}{2\pi} \int_{x^2}^{\frac{9}{16}} \frac{1}{t} \left(\frac{16\sqrt{t-x^2}}{\sqrt{9-16t}} - \frac{\sqrt{9-16t}}{\sqrt{t-x^2}} \right) dt,$$

$$J_3(x) = \frac{x}{\pi} \int_{x^2}^1 \frac{1}{t} \left(\frac{16\sqrt{t-x^2}}{\sqrt{16-16t}} - \frac{\sqrt{16-16t}}{\sqrt{t-x^2}} \right) dt.$$

For $k = 1, 3, 4$, we have

$$\begin{aligned} & \int_{x^2}^{\frac{k^2}{16}} \frac{1}{t} \left(\frac{16\sqrt{t-x^2}}{\sqrt{k^2-16t}} - \frac{\sqrt{k^2-16t}}{\sqrt{t-x^2}} \right) dt \\ &= \int_{x^2}^{\frac{k^2}{16}} \frac{32t - 16x^2 - k^2}{t\sqrt{-16t^2 + (16x^2 + k^2)t - k^2x^2}} dt \\ &= 32 \int_{x^2}^{\frac{k^2}{16}} \frac{dt}{\sqrt{-16t^2 + (16x^2 + k^2)t - k^2x^2}} \\ &\quad - (16x^2 + k^2) \int_{x^2}^{\frac{k^2}{16}} \frac{dt}{t\sqrt{-16t^2 + (16x^2 + k^2)t - k^2x^2}} \\ &= 8 \arcsin \frac{32t - (16x^2 + k^2)}{k^2 - 16x^2} \Big|_{t=x^2}^{\frac{k^2}{16}} \\ &\quad - \frac{16x^2 + k^2}{kx} \arcsin \frac{(16x^2 + k^2)t - 2k^2x^2}{t(k^2 - 16x^2)} \Big|_{t=x^2}^{\frac{k^2}{16}} \\ &= \pi \left(8 - \frac{16x^2 + k^2}{kx} \right). \end{aligned}$$

Hence

$$\begin{aligned} J_1(x) &= \frac{x}{2\pi} \pi \left(8 - \frac{16x^2 + 1}{x} \right) = -8x^2 + 4x - \frac{1}{2}; \\ J_2(x) &= -\frac{3x}{2\pi} \pi \left(8 - \frac{16x^2 + 9}{3x} \right) = 8x^2 - 12x + \frac{9}{2}; \\ J_3(x) &= \frac{x}{\pi} \pi \left(8 - \frac{16x^2 + 16}{4x} \right) = -4(1-x)^2. \end{aligned}$$

Consequently,

(1). for $x \in (0, \frac{1}{4})$,

$$G(x) = J_1(x) + J_2(x) + J_3(x) = -4x^2;$$

(2). for $x \in [\frac{1}{4}, \frac{3}{4})$,

$$G(x) = J_2(x) + J_3(x) = 4x^2 - 4x + \frac{1}{2};$$

(3). for $x \in [\frac{3}{4}, 1)$,

$$G(x) = J_3(x) = -4(x-1)^2.$$

i.e.,

$$G(x) = \begin{cases} -4x^2, & x \in (0, \frac{1}{4}), \\ 4x^2 - 4x + \frac{1}{2}, & x \in [\frac{1}{4}, \frac{3}{4}), \\ -4(x-1)^2, & x \in [\frac{3}{4}, 1), \\ 0, & x = 0 \text{ or } x \geq 1, \\ -G(-x), & x < 0. \end{cases}$$

It is clear that $G(x) \leq 0$ and is symmetric about $x = \frac{1}{2}$, therefore, Eqs. (7.8) and (7.9) are equivalent to the followings:

$$\begin{cases} 0 \leq \frac{\rho}{s} - \frac{d}{2s} \leq \frac{\rho}{s} + \frac{d}{2s} \leq 1 \\ \frac{1}{2} - \left(\frac{\rho}{s} - \frac{d}{2s}\right) = \left(\frac{\rho}{s} + \frac{d}{2s}\right) - \frac{1}{2} \end{cases} \quad \text{or} \quad \begin{cases} -1 \leq \frac{\rho}{s} - \frac{d}{2s} \leq \frac{\rho}{s} + \frac{d}{2s} \leq 0 \\ -\frac{1}{2} - \left(\frac{\rho}{s} - \frac{d}{2s}\right) = \left(\frac{\rho}{s} + \frac{d}{2s}\right) + \frac{1}{2} \end{cases},$$

i.e.

$$\begin{cases} 0 \leq \rho - \frac{d}{2} \leq \rho + \frac{d}{2} \leq s \\ \rho = \frac{s}{2} \end{cases} \quad \text{or} \quad \begin{cases} -s \leq \rho - \frac{d}{2} \leq \rho + \frac{d}{2} \leq 0 \\ \rho = -\frac{s}{2} \end{cases},$$

i.e.

$$\begin{cases} s \geq d \\ \rho = \frac{s}{2} \end{cases} \quad \text{or} \quad \begin{cases} s \geq d \\ \rho = -\frac{s}{2} \end{cases}.$$

Hence, Eqs. (7.8) and (7.9) are solvable if and only if $s \geq d$. There are two solutions:

$$\rho_1 = -\frac{s}{2}, \quad \rho_2 = \frac{s}{2}.$$

These two solutions realize that the local maxima of $|\nabla W_s f_{l_d}(x_\rho, y_\rho)|$ arrive at both sides of the central line l of l_d and the distance from l is $\frac{s}{2}$, which is independent on the width d .

In summary, The above three invariance properties can be rewritten as the following theorem:

Theorem 7.1 *Let l_d be a Dirac-structure edge with width d and l be its central line. The local maxima modulus of the wavelet transform corresponding to the wavelets of Eq. (8.2) forms two new lines which are located symmetrically on both sides of the central line, and have the same direction with it. If scale $s \geq d$, then the distance between the two new ones equals to s .*

This theorem describes the width invariant property, which is important. It improves our former results in Chapter 6 and [Tang et al., 2000]. A couple of graphical examples are shown in Fig. 7.6.

Such a invariance property can be illustrated in Figs. 7.7 and 7.8 clearly. The graph in Fig. 7.7(a) displays function $G(\alpha)$ defined by Eq. (7.6). The graph in Fig. 7.7(b) indicates the positions of local maximum modulus by the wavelets defined by Eq. (8.1), where, the three horizontal straight segments correspond to the widths of three curves. Point 0 at the horizontal axis corresponds to the center of the curves. The vertical axis corresponds to the scale s of the wavelet transform. The slant straight lines correspond to the positions of the local maximum moduli of the wavelet transforms. For each scale s , the local maximum moduli of the wavelet transforms with respect to the curves of different widths are located at the same positions.

To be more clear, we look at Fig. 7.8. There are three curves with the different widths of $d = 1, d = 2$ and $d = 3$, as shown in Figs. 7.8(a), (b) and (c). The three horizontal straight segments in Fig. 7.8(d) correspond to these three curves. This width invariance can guarantee that for each scale s , the local maximum moduli of the wavelet transforms with respect to the curves of different widths are located at the same positions. For example, when we apply the wavelet transform with scale $s = 8$ to these three curves, the locations of local maximum moduli of the wavelet transform for three curves, which have different widths of $d = 1, 2$ and 3 , are the same. They are located on the points -4 and 4 , and denoted by p_1 and p_2 respectively in Fig. 7.8(d).

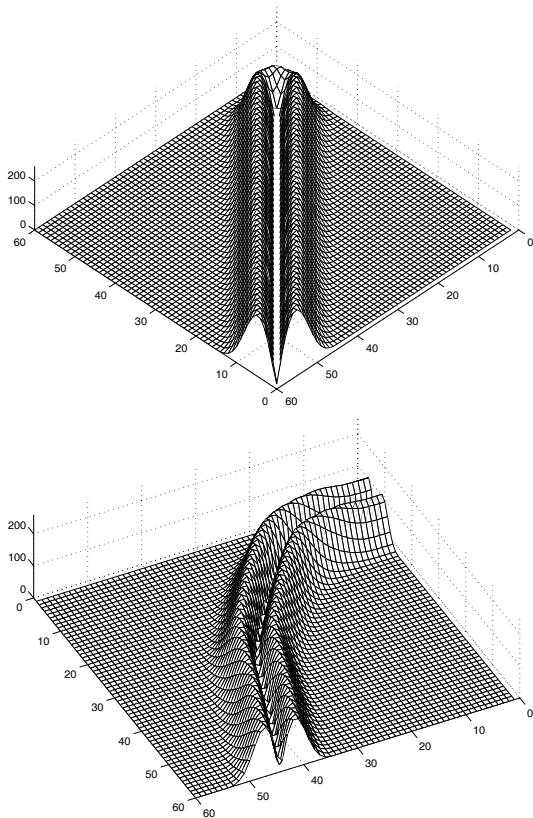


Fig. 7.6 Modulus of wavelet transforms corresponding a segment of straight line and a curve.

7.3 Comparison with Other Wavelets

In this section, we will compare our new wavelet with two other kinds of wavelets, namely:

- Gaussian wavelets: they are derived from the smoothness functions $\theta(x)$, which are Gaussian function as Eq. (7.1);
- the quadratic spline wavelets [Tang et al., 2000].

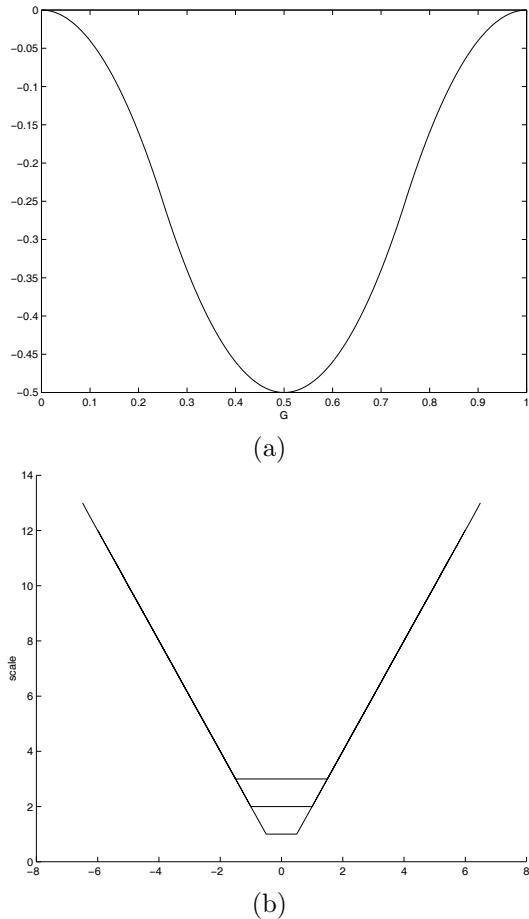


Fig. 7.7 (a) function $G(\alpha)$ defined by Eq. (7.6); (b) positions of local maximum modulus of the wavelets defined by Eq. (8.1).

7.3.1 Comparison with Gaussian Wavelets

For the first case, $\theta(x)$ is a Gaussian function, we let $\sigma := 1/4$ for simplicity. Obviously, its derivative is a wavelet which is called a Gaussian wavelet as shown in Fig. 7.9(a). A comparison between the new wavelet defined by Eq. (8.1) and Gaussian wavelet will be described below.

Let $\phi(x)$ in Eq. (7.6) be the Gaussian function of $\sigma := 1/4$ and $\psi(x)$ be

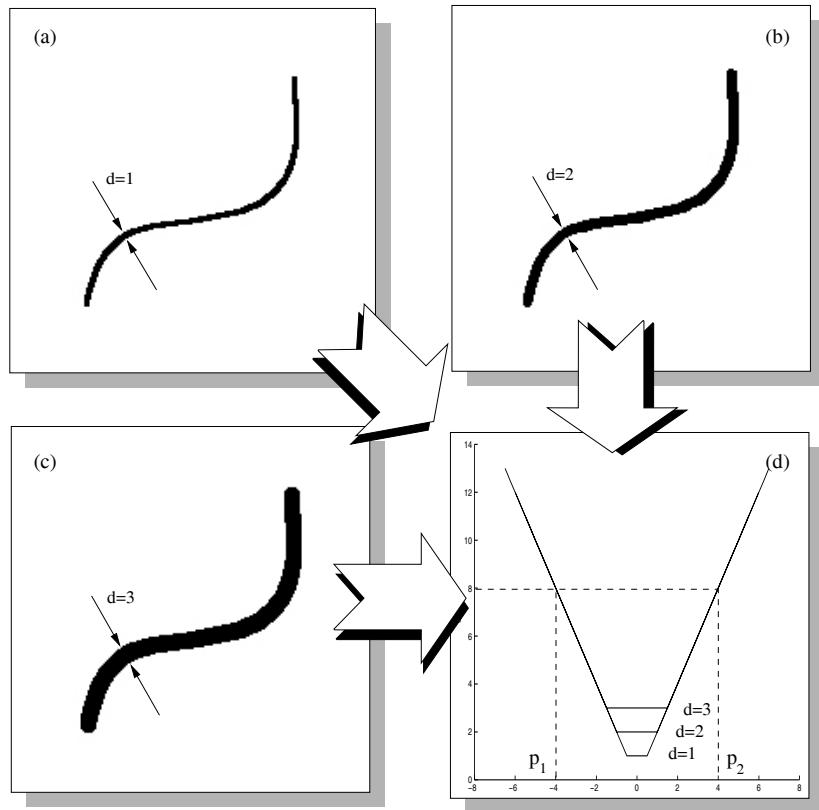


Fig. 7.8 Three curves with different widths of (a) $d = 1$, (b) $d = 2$, (c) $d = 3$, and (d) the positions of local maximum modulus by the wavelets defined by Eq. (8.1).

its derivative, i.e.,

$$\phi(x) = \frac{4}{\sqrt{2\pi}} e^{-8x^2}, \quad \psi(x) = -\frac{64}{\sqrt{2\pi}} xe^{-8x^2}.$$

Then the function $G(x)$ defined by (7.6) can be calculated as follows:

$$\begin{aligned} G(x) &:= x \int_0^\infty \frac{\psi(\sqrt{t^2 + x^2})}{\sqrt{t^2 + x^2}} dt \\ &= -\frac{64}{\sqrt{2\pi}} x \int_0^\infty e^{-8(x^2 + t^2)} dt \end{aligned}$$

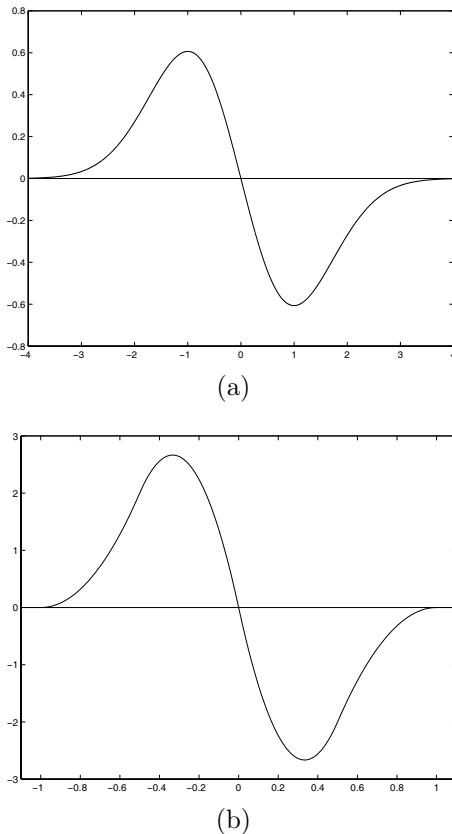


Fig. 7.9 (a) Gaussian wavelet; (b) quadratic spline wavelet.

$$= \left(\int_0^\infty e^{-8t^2} dt \right) \psi(x).$$

Hence, Eq. (7.8) is equivalent to

$$\psi\left(\frac{\rho}{s} + \frac{d}{2s}\right) = \psi\left(\frac{\rho}{s} - \frac{d}{2s}\right), \quad (7.10)$$

i.e.,

$$\left(\frac{\rho}{s} + \frac{d}{2s}\right) \exp\left[-8\left(\frac{\rho}{s} + \frac{d}{2s}\right)^2\right] = \left(\frac{\rho}{s} - \frac{d}{2s}\right) \exp\left[-8\left(\frac{\rho}{s} - \frac{d}{2s}\right)^2\right].$$

Although this equation can not be solved analytically, a numerical solution can be obtained. According to Fig. 7.9 (left) which describes function $\psi(x)$, for a constant C : $0 < C \leq 1$, there exists a pair of numbers (α, β) satisfying $0 \leq \alpha < \beta \leq 1$, such that

$$\begin{cases} \beta - \alpha = C \\ \psi(\alpha) = \psi(\beta) \end{cases}. \quad (7.11)$$

It is obvious that both α and β depend on only the constant C , which means that they are functions on C . Therefore, α and β can be denoted by

$$\alpha = \alpha(C), \quad \beta = \beta(C). \quad (7.12)$$

In the following, we will solve ρ to satisfy Eq. (7.10). Comparing Eq. (7.10) and Eq. (7.11) yields

$$\frac{\rho}{s} + \frac{d}{2s} = \beta, \quad \frac{\rho}{s} - \frac{d}{2s} = \alpha. \quad (7.13)$$

Therefore, we obtain

$$\rho = \frac{s}{2}(\alpha + \beta), \quad C = \beta - \alpha = \frac{d}{s}.$$

By Eq. (7.12), we can deduce that

$$\rho = \frac{s}{2}(\alpha + \beta) = \frac{s}{2} \left(\alpha \left(\frac{d}{s} \right) + \beta \left(\frac{d}{s} \right) \right).$$

Then, for a fixed d , parameter ρ is a function of s . Fig. 7.10(a) shows function $G(\alpha)$ with respect to the Gaussian wavelet, and Fig. 7.10(b) illustrates the positions of local maximum modulus by Gaussian wavelet for $d = 1, 2, 3$ respectively. Fig. 7.10(e) shows function $G(\alpha)$ with respect to the new wavelet, and Fig. 7.10(f) illustrates the positions of local maximum modulus by the new wavelet for $d = 1, 2, 3$ respectively. The readers would be suggested to compare Figs. 7.10(b) and 7.10(f). It is obvious that the locations of the local maximum of the wavelet transform with respect to Gaussian wavelet depend nonlinearly on the curve width d heavily, while the locations of the local maximum of the wavelet transform using the new wavelet are fixed.

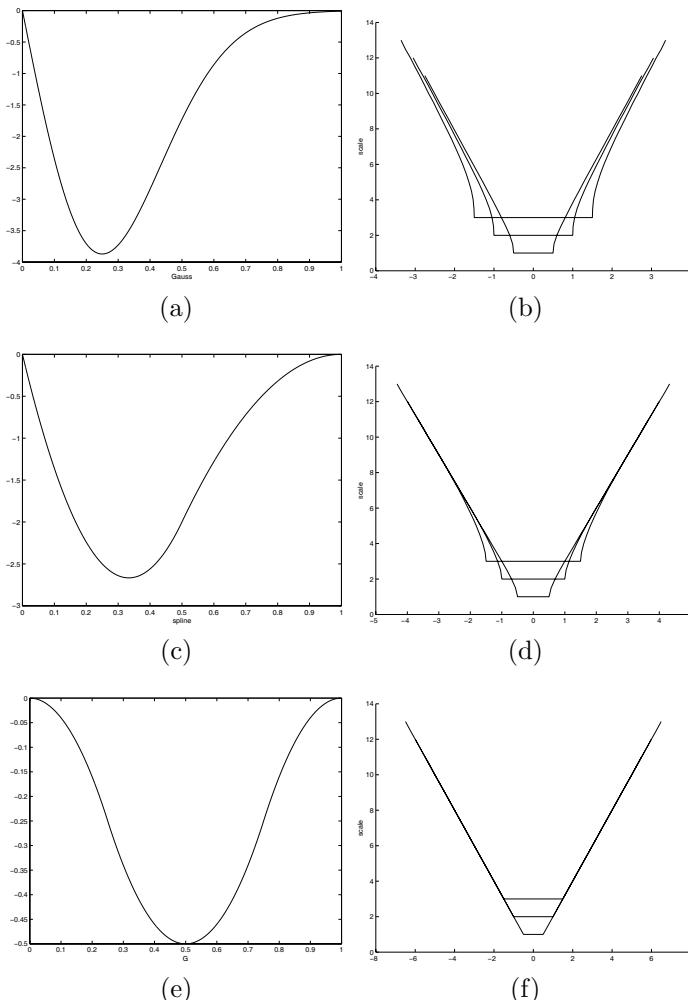


Fig. 7.10 Left, i.e. (a), (c) and (e): The functions $G(\alpha)$ with respect to the Gaussian wavelet, quadratic spline wavelet and the new wavelet defined by Eq. (8.1) respectively. Right, i.e. (b), (d) and (f): The positions of local maximum moduli by Gaussian wavelet, quadratic spline wavelet and the new wavelet defined by Eq. (8.1).

7.3.2 Comparison with Quadratic Spline Wavelets

$\theta(x)$ is set to be the quadratic spline as discussed in Chapter 6 and [Tang et al., 2000], and a graphical display is presented in Fig. 7.9(b). Fig. 7.10(c)

presents function $G(\alpha)$ with respect to the quadratic spline wavelet. Fig. 7.10(d) illustrates the corresponding locations of the maximum modula for $d = 2, 4, 8$ respectively. Similarly, it can be shown that ρ depends nonlinearly on d lightly.

7.4 Algorithm and Experiments

In this section, the algorithm for extracting the Dirac-structure edges will be presented, including the calculation of the nonzero coefficients $\{\phi_{k,l}^s\}$ for different scales. Several experiments will also be conducted.

7.4.1 Algorithm

In practice, the wavelet transform should be calculated discretely. We have the following formula:

$$\begin{aligned} W_s^i f(n, m) &= \int \int f(u, v) \psi_s^i(n - u, m - v) du dv \\ &= \sum_{k,l} f(k, l) \int_k^{k+1} \int_l^{l+1} \psi_s^i(n - u, m - v) du dv \\ &= \sum_{k,l} f(k, l) \int_{n-k-1}^{n-k} \int_{m-l-1}^{m-l} \psi_s^i(u, v) du dv \\ &= \sum_{k,l} f(n - k - 1, m - l - 1) \psi_{k,l}^{s,i}, \quad (i = 1, 2), \end{aligned}$$

where

$$\begin{aligned} \psi_{k,l}^{s,i} &= \int_k^{k+1} \int_l^{l+1} \psi_s^i(u, v) du dv \\ &= \int_{k/s}^{(k+1)/s} \int_{l/s}^{(l+1)/s} \psi^i(u, v) du dv, \quad (i = 1, 2). \end{aligned}$$

In fact, the computation of the above formula is a discrete convolution. For a small scale s , it can be performed directly. But for a large scale s , considering the amount of computation, FNT (fast numerous transform) or other fast algorithms can be utilized instead. The details can be found in [Mallat and Hwang, 1992; Tang et al., 1998b]. In the following, we discuss the calculation of the coefficients $\{\psi_{k,l}^{s,i}\}$.

For the wavelets defined by Eq.(8.2), since

$$\psi^1(u, v) = \phi'(\sqrt{u^2 + v^2}) \frac{u}{\sqrt{u^2 + v^2}} = \psi^2(v, u),$$

we deduce that

$$\begin{aligned} \psi_{k,l}^{s,1} &= \int_{k/s}^{(k+1)/s} du \int_{l/s}^{(l+1)/s} \psi^1(u, v) dv \\ &= \int_{k/s}^{(k+1)/s} du \int_{l/s}^{(l+1)/s} \psi^2(v, u) dv \\ &= \int_{k/s}^{(k+1)/s} dv \int_{l/s}^{(l+1)/s} \psi^2(u, v) du \\ &= \psi_{k,l}^{s,2}. \end{aligned}$$

Therefore, we need to calculate only $\psi_{k,l}^{s,1}$ for all $k, l \in Z$. Note that $\psi^1(u, v)$ is odd on u and even on v , thus we have

$$\begin{aligned} \psi_{-k,l}^{s,1} &= \int_{-k/s}^{(-k+1)/s} du \int_{l/s}^{(l+1)/s} \psi^1(u, v) dudv \\ &= - \int_{(k-1)/s}^{k/s} du \int_{l/s}^{(l+1)/s} \psi^1(u, v) dudv \\ &= -\psi_{k-1,l}^{s,1}. \end{aligned}$$

Similarly,

$$\psi_{k,-l}^{s,1} = \psi_{k,l-1}^{s,1}, \quad \psi_{-k,-l}^{s,1} = -\psi_{k-1,l-1}^{s,1}.$$

Consequently, we further need to calculate only $\psi_{k,l}^{s,1}$ for all positive integers k, l . By $\psi^1(u, v) = \frac{\partial}{\partial u}[\phi(\sqrt{u^2 + v^2})]$, we have

$$\begin{aligned} \psi_{k,l}^{s,1} &= \int_{k/s}^{(k+1)/s} \int_{l/s}^{(l+1)/s} \frac{\partial}{\partial u}[\phi(\sqrt{u^2 + v^2})] dudv \\ &= \int_{l/s}^{(l+1)/s} \left[\phi\left(\sqrt{v^2 + \left(\frac{k+1}{s}\right)^2}\right) - \phi\left(\sqrt{v^2 + \left(\frac{k}{s}\right)^2}\right) \right] dv \\ &= \phi_{l,k+1}^s + \phi_{l+1,k}^s - \phi_{l+1,k+1}^s - \phi_{l,k}^s, \end{aligned}$$

where

$$\phi_{k,l}^s = \int_{k/s}^{\infty} \phi \left(\sqrt{v^2 + \left(\frac{l}{s}\right)^2} \right) dv.$$

Hence, our question turns to calculate $\phi_{k,l}^s$ for all non-negative integers k and l . Since

$$\psi^1(u, v) = \phi'(\sqrt{u^2 + v^2}) \frac{u}{\sqrt{u^2 + v^2}} = \psi(\sqrt{u^2 + v^2}) \frac{u}{\sqrt{u^2 + v^2}},$$

where ψ is defined by Eq. (8.1), we have for any non-negative integers k, l satisfying $k^2 + l^2 < s^2$,

$$\begin{aligned} \phi_{k,l}^s &= \int_{k/s}^{\infty} \phi(\sqrt{v^2 + (l/s)^2}) dv \\ &= \int_{\frac{\sqrt{k^2+l^2}}{s}}^{\infty} \phi(v) d\sqrt{v^2 - (l/s)^2} \\ &= \left. \sqrt{v^2 - (l/s)^2} \phi(v) \right|_{v=\frac{\sqrt{k^2+l^2}}{s}}^{\infty} - \int_{\frac{\sqrt{k^2+l^2}}{s}}^{\infty} \phi'(v) \sqrt{v^2 - (l/s)^2} dv \\ &= -\frac{l}{s} \phi\left(\frac{\sqrt{k^2+l^2}}{s}\right) - \int_{\frac{\sqrt{k^2+l^2}}{s}}^1 \psi(v) \sqrt{v^2 - (l/s)^2} dv \\ &= -\frac{l}{s} \int_{-\infty}^{\frac{\sqrt{k^2+l^2}}{s}} \psi(v) dv - \int_{\frac{\sqrt{k^2+l^2}}{s}}^1 \psi(v) \sqrt{v^2 - (l/s)^2} dv \\ &= -\frac{l}{s} \int_0^{\frac{1}{s}\sqrt{k^2+l^2}} \psi(v) dv - \frac{l}{s} \int_{-1}^0 \psi(v) dv \\ &\quad - \int_{\frac{1}{s}\sqrt{k^2+l^2}}^1 \psi(v) \sqrt{v^2 - (l/s)^2} dv \\ &= -\frac{l}{s} \int_0^{\frac{1}{s}\sqrt{k^2+l^2}} \psi(v) dv + \frac{l}{s} \int_0^1 \psi(v) dv \\ &\quad - \int_{\frac{1}{s}\sqrt{k^2+l^2}}^1 \psi(v) \sqrt{v^2 - (l/s)^2} dv \\ &= \frac{l}{s} \int_{\frac{1}{s}\sqrt{k^2+l^2}}^1 \psi(v) dv - \int_{\frac{1}{s}\sqrt{k^2+l^2}}^1 \psi(v) \sqrt{v^2 - (l/s)^2} dv \\ &= \int_{\frac{1}{s}\sqrt{k^2+l^2}}^1 \left[\frac{l}{s} - \sqrt{v^2 - (l/s)^2} \right] \psi(v) dv. \end{aligned}$$

On the other hand, it is easy to see that $\phi_{k,l}^s = 0$ for all integers k, l satisfying $k^2 + l^2 \geq s^2$ due to the compact support $[-1, 1]$ of $\phi(x)$. According to the composite trapezoidal formula of numerical quadrature [Chao-hao, 1992]

$$\int_a^b f(x)dx \approx \frac{b-a}{2n} \left[f(a) + f(b) + 2 \sum_{i=1}^{n-1} f(a + i\frac{b-a}{n}) \right],$$

we can calculate all the coefficients $\phi_{k,l}^s$ numerically for non-negative integers k, l . The positive integer n in the formula can be set to be so large that the error is smaller than any prior number. The possible nonzero items of $\phi_{k,l}^s$ for $s = 2, 4, 8$ are listed in the following four tables.

$l \setminus k$	$k = 0$	$k = 1$
$l = 0$	0.2500	0.1617
$l = 1$	0.1241	0.0656

Table 7.1 The nonzero coefficients $\{\phi_{k,l}^s\}$ for $s = 2$.

$l \setminus k$	$k = 0$	$k = 1$	$k = 2$	$k = 3$
$l = 0$	0.2500	0.2138	0.1170	0.0435
$l = 1$	0.1520	0.1171	0.0612	0.0236
$l = 2$	0.0602	0.0477	0.0260	0.0084
$l = 3$	0.0207	0.0169	0.0078	0

Table 7.2 The nonzero coefficients $\{\phi_{k,l}^s\}$ for $s = 4$.

$l \setminus k$	$k = 0$	$k = 1$	$k = 2$	$k = 3$
$l = 0$	0.2500	0.2254	0.1871	0.1343
$l = 1$	0.1844	0.1698	0.1374	0.0963
$l = 2$	0.1202	0.1110	0.0882	0.0603
$l = 3$	0.0683	0.0630	0.0495	0.0330
$l = 4$	0.0335	0.0308	0.0239	0.0158
$l = 5$	0.0144	0.0133	0.0104	0.0071
$l = 6$	0.0060	0.0056	0.0046	0.0033
$l = 7$	0.0025	0.0023	0.0018	0.0009

Table 7.3 The nonzero coefficients $\{\phi_{k,l}^s\}$ for $s = 8$, $k = 0 - 3$.

$l \setminus k$	$k = 4$	$k = 5$	$k = 6$	$k = 7$
$l = 0$	0.0821	0.0423	0.0190	0.0079
$l = 1$	0.0577	0.0292	0.0133	0.0056
$l = 2$	0.0350	0.0173	0.0081	0.0032
$l = 3$	0.0187	0.0094	0.0046	0.0013
$l = 4$	0.0091	0.0049	0.0023	0
$l = 5$	0.0045	0.0024	0.0005	0
$l = 6$	0.0020	0.0005	0	0
$l = 7$	0	0	0	0

Table 7.4 The nonzero coefficients $\{\phi_{k,l}^s\}$ for $s = 8$, $k = 4 - 7$.

Based on the characterization of a straight line in an image developed in Section 7.2, an algorithm to detect straight lines in an image can be designed. The result is also valid for general curves since a short segment of a curve can be regarded as a straight line approximately. In fact, wavelet transforms are essentially local analysis. Therefore the result of Theorem 8.1 can be applied to the general curves in an image. Our algorithm to detect curves in an image is designed as follows.

Algorithm 7.1 Let $f(x, y)$ be an image containing curves. For a scale $s > 0$,

- Step 1** Calculate all the wavelet transforms $\{W_s^1 f(x, y), W_s^2 f(x, y)\}$ with respect to the wavelets defined by Eq.(8.2).
- Step 2** Calculate the local maxima f_{locmax} of $|\nabla W_s f(x, y)|$ and the gradient direction $f_{gradient}$.
- Step 3** For each point (x, y) with local maximum, search the point whose distance along the gradient direction from (x, y) is s . If it is a point of local maxima, the center point is detected.
- Step 4** The curves formed by all the points detected in Step 3 are what we need.

7.4.2 Experiments

Four circles with various gray-levels and widths as shown in Fig. 7.11 are tested using the new method. The original images are illustrated on the left column of Fig. 7.11. After applying Steps 1 and 2 of the proposed wavelet transform algorithm to these circles, the local maximum modulus of the

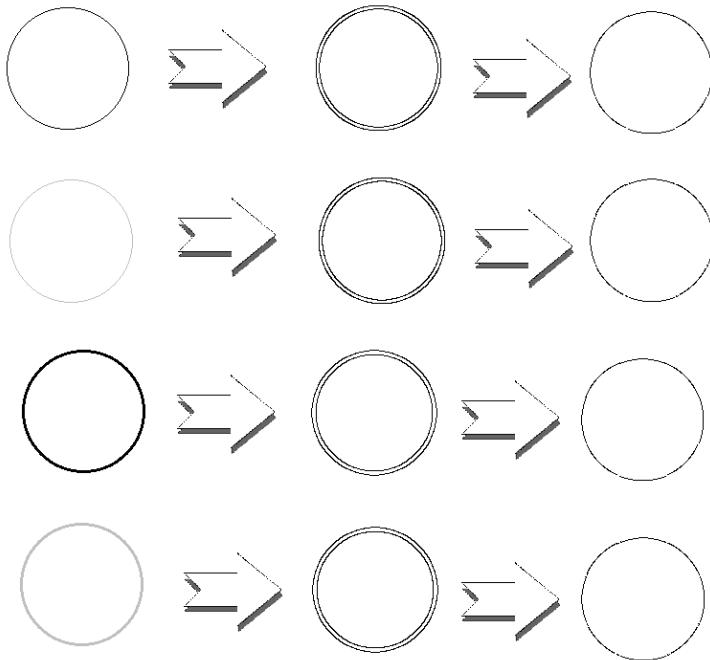


Fig. 7.11 Detection of lines by new wavelet transform, $s=6$.

wavelet transform with respect to them can be computed and the results are given on the middle column in Fig. 7.11. Finally, the central lines of these circles are extracted using Steps 3 and 4 of the above algorithm, and presented on the right column in Fig. 7.11.

Next, let us turn back to the beginning of this chapter, and look at Fig. 7.12. The particular task is that we are required to extract the skeleton of the circle with various widths. Unfortunately, as we have shown in Fig. 7.12, the algorithm based on the spline wavelet in Chapter 6 can not work well due to the width dependence of the detection. Fortunately, as described in detail in Section 7.2, the method developed in this chapter possesses the width invariant, grey-level invariant as well as slope invariant. According to these properties, the skeleton of the circle and tree in Fig. 7.12 can be extracted. After applying Steps 1 and 2 of the above algorithm to

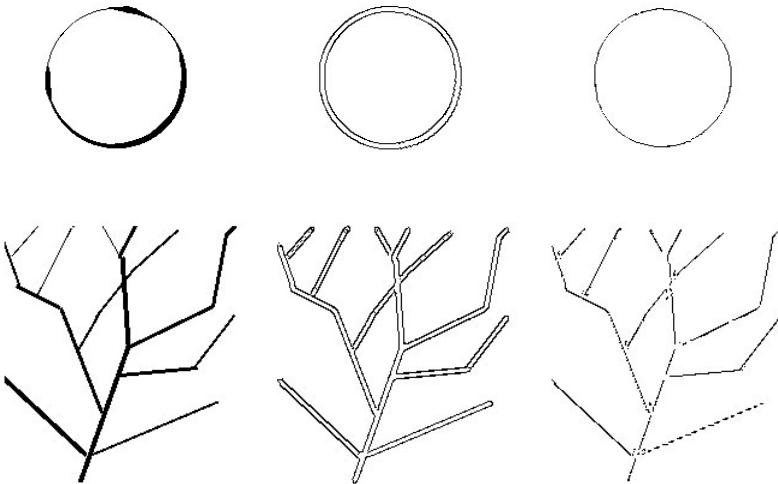


Fig. 7.12 Left: the original image; Middle: the location of maximum modulus of the wavelet transform corresponding to $s = 6$; Right: the skeleton extracted by the algorithm in this chapter.

the original image as displayed on the left column of Fig. 7.12, the local maximum modulus of the wavelet transform with respect to them can be computed and presented on the middle column in Fig. 7.12. At last, the central lines are extracted using Steps 3 and 4 of the above algorithm, and presented on the right column in Fig. 7.12.

Figs. 7.13, 7.14 and 7.15 show experimental results for four other images. Fig. 7.13(a) is the original image which consists of some basic strokes of Chinese and Japanese characters in different fonts. Its MMWTs (maximum moduli of the wavelet transform) are displayed in Fig. 7.13(b). It is easy to see that the locations of its MMWTs are independent of the width of the strokes. Fig. 7.13(c) is the original image of a noisy contour of an airplane. There are many anomalous blurs along the contour and the widths along the contour are erratically different. It is interesting to see that its MMWTs, which are shown in Fig. 7.13(d), are smooth with the same widths along the contour. Fig. 7.14(a) and Fig. 7.15(a) are two circuit diagrams. Some lines of the schematic diagrams are eroded because of aging. Some pixels on the lines of the schematic diagrams disappear and some other noises appear instead. The widths of each line are thus not the

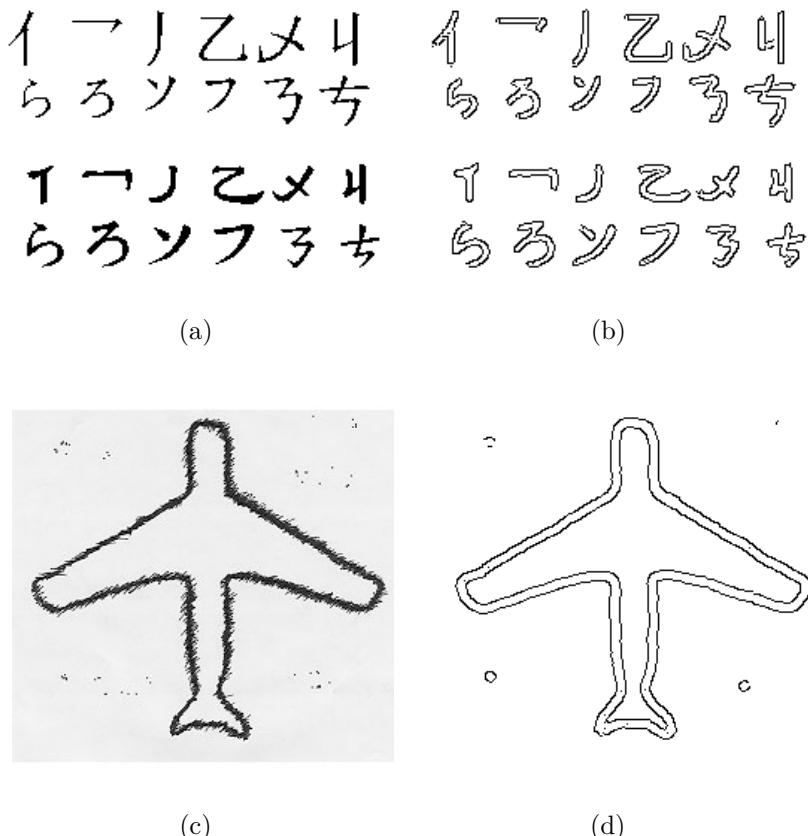


Fig. 7.13 (a) An original image containing some basic strokes of Chinese and Japanese characters in different fonts; (b) the location of the MMWTs of (a) with the new wavelet corresponding to $s = 5$; (c) an original image of a noisy contour of an airplane; (d) the location of the MMWTs of (c) with the new wavelet corresponding to $s = 5$.

same. With the novel wavelet presented in this chapter, their MMWTs, which are shown in Fig. 7.14(b) and Fig. 7.15(b) respectively, remain the same.

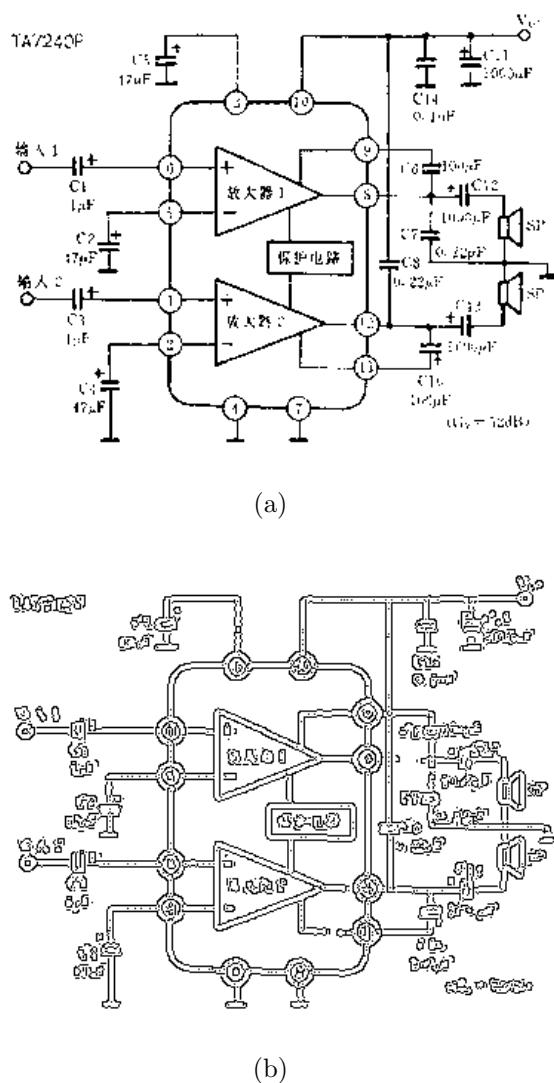


Fig. 7.14 (a) A circuit diagram; (b) The locations of the MMWTs of (a) with the new wavelet corresponding to $s = 4$.

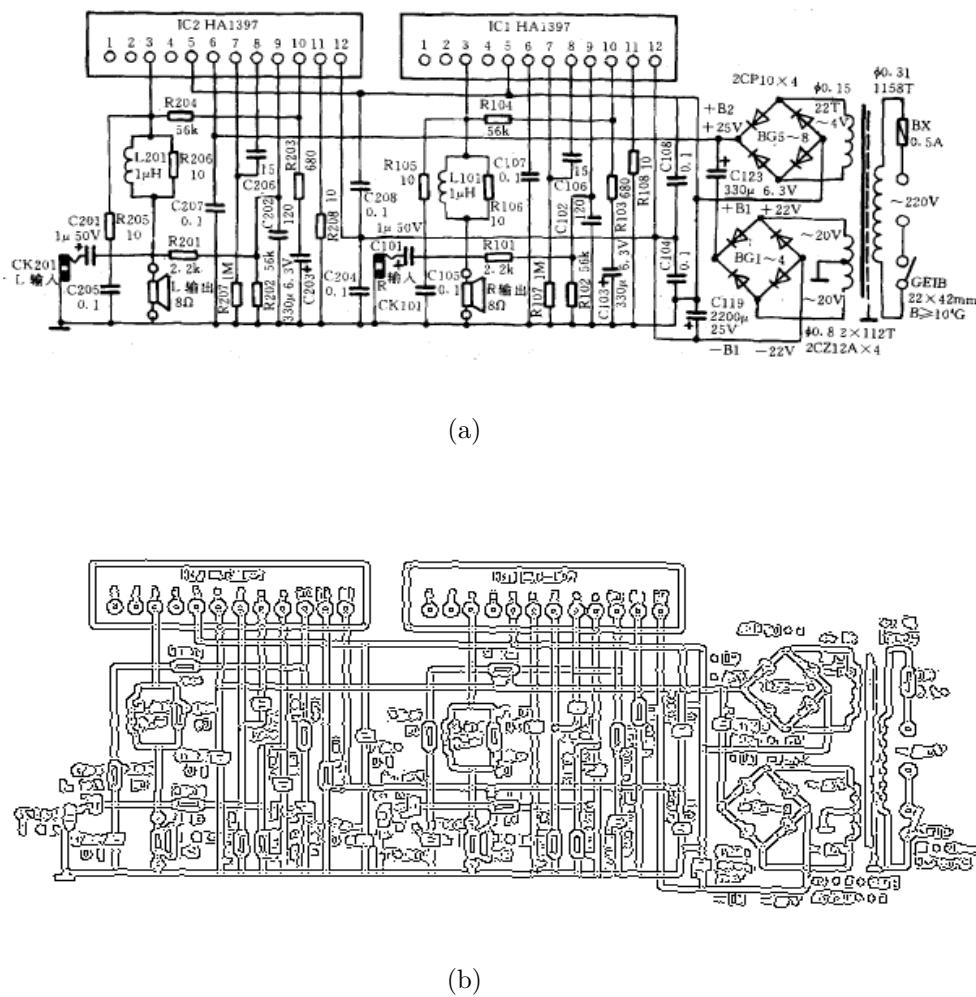


Fig. 7.15 (a) A circuit diagram; (b) The locations of the MMWTs of (a) with the new wavelet corresponding to $s = 4$.

Chapter 8

Skeletonization of Ribbon-like Shapes with New Wavelet Function

One of the most significant topics in pattern recognition is analysis of Ribbon-like shapes. It can be applied to character recognition, signature verification, understanding of paper-based graphics including maps and engineering drawings, computer-assisted cartoon, fingerprint analysis, etc. [Janssen, 1997; Tombre, 1998]. Representation of a shape using a suitable form is essential in these recognition systems. An appropriate representation of a Ribbon-like shape is its skeleton [Zou and Yan, 2001]. The objective of this study is to extract skeletons from planar Ribbon-like shapes, which have the following properties: (1) conforming to human perceptions of the original shapes, (2) being centred inside the original shapes, (3) being efficiently computable, (4) being robust against noise and geometric transformation.

Skeletons have been defined in various ways in the different literatures. Generally, the skeleton of a shape is referred to as the locus of the symmetric points or symmetric axes of the local symmetries of the shape, in other words, different local symmetry analyses may result in different symmetric points, and hence different skeletons are produced. There are three methods of the local symmetry analysis, which are well-known to the shape analysis community, namely:

- Blum's Symmetric Axis Transform (SAT) [Blum, 1967],
- Brady's Smoothed Local Symmetry (SLS) [Brady, 1983],
- Leyton's Process-Inferring Symmetry Analysis (PISA) [Leyton, 1988].

Each method contributes to skeletonization of the shapes greatly. The

difference in selection of the location of a symmetric point makes these methods distinguishable from each other. Nevertheless, the major problem of the SAT and PISA is that the symmetric points of a local symmetry, and hence a skeleton segment may lie in a perceptually distinct part of the underlying shape. Although the skeleton obtained from the SLS has a pleasing visual appearance, the major shortcoming of the SLS is that some perceptually irrelevant symmetric axes may be created. Some grouping rules can be used [Connell and Brady, 1987]. However, these rules may not be appropriate for a wide range of shapes. An alternative way is to divided the shape into several parts, and thereafter, the SLS axes are computed within each part [Rom and Medioni, 1984]. Dividing a shape into suitable parts is the crucial in this method, unfortunately, it is a difficult task. Another powerful measure of the symmetry proposed by Kovesi [Kovesi, 1995; Kovesi, 1997] is based on the analysis of phase information. The mathematics of phase congruency using wavelets has been studied by Kovesi [Kovesi, 1995], where the phase information can be used to construct a contrast invariant measure of the symmetry that does not require any prior recognition or segmentation. Although it is possible to calculate the phase congruency directly [Venkatesh and Owens, 1990], it is a rather awkward quantity to calculate. Moreover, it provides limited information about the overall shape of an object, and it is not shown further whether it is capability of skeleton extraction.

The computation of the skeleton of a shape is another challenge for the skeletonization. In other words, to determine the symmetric points of a shape by using the above symmetry analysis is a key issue. Computing the skeleton of a shape in the continuous domain is not an easy task. Suppose the boundaries of a shape are represented by two curves, generally, it is difficult to determine the symmetric points from those curves. Especially in the discrete domain, it is even more difficult to compute the skeleton of a shape using the definitions of the skeleton given above, since the shape is represented by a set of discrete pixels, not in a continuous form. The constrained Delaunary triangulation technique [Zou and Yan, 2001] is a sound solution for this problem. However, it has to suffer from the complicated computation and it costs too much computation time.

Therefore, many approximation methods have been proposed to compute skeletons in the discrete world, which can be divided into two groups: pixel-based method and non-pixel-based method.

In the pixel-based method, each foreground pixel is utilized for compu-

tation in the skeletonization process. Techniques used in the pixel-based method include thinning [Lam et al., 1992; Smith, 1987] and distance transform [Smith, 1987]. The former is applied to most of the existing skeletonization methods. The basic principle of them is to repetitively remove the pixels of the outside layer of an image until only the central pixels remain, which establish the approximated skeleton of the shape. It can be regarded as the grassfire process, that produces homotopic skeleton, does not alter the topology of the shape, and is easy to implement. The distance transform possesses the following advantages: (1) A skeleton is produced in a fixed number of passes through the image regardless of the sizes of the object. (2) It may operate faster than thinning method does, since it requires $O(n^2)$ to $O(n^3)$ time to compute the skeleton of an $n \times n$ image [Smith, 1987] rather than the later, which usually takes $O(n^3)$ to do so [Smith, 1987].

The pixel-based method often suffers from one or more of the following drawbacks:

First, the generated skeletons are generally in discrete forms, where the skeleton points are discrete pixels. However, such a skeleton is not helpful for recognizing the underlying shape unless skeleton pixels are linked by lines or curves implicitly or explicitly to form a graph.

Second, even if the skeleton pixels are linked, the resulting skeleton may not be centred inside the underlying shape due to the use of discrete data.

Third, the computational complexity is quite high, since all foreground pixels are needed for the computation in the skeletonization process.

In the non-pixel-based method, the skeleton of a shape is analytically derived from the border of the image. There are two types of non-pixel-based methods, which are based on either cross-section [Pavlidis, 1986] or Voronoi diagrams [Ogniewicz and Kubler, 1995]. These methods attempt to determine the symmetric points of a shape without the intermediate step of the grassfire propagation. The fundamental concept of these methods is that the local symmetric axes of a shape are derived from pairs of contour pixels or a contour segment representing a sequence of the contour pixels. The mid-points or centre-lines of the pairs of the contour elements are connected to generate the skeleton. An obvious advantage of these methods over the pixel-based methods is that fewer data (only the contour pixels) are used for the skeletonization. Hence, they are expected to be faster. Especially, comparing with the thinning technique, another advantage of the non-pixel-based method is that the real coordinates are available for

the skeleton points, which are not restricted by the discrete grids. Thus, the generated skeleton can be centred inside the underlying shapes.

On the other hand, the accurate identification of the local symmetries of the underlying shape is the major problem in the non-pixel-based methods. In fact, given a contour pixel of a digital shape, it may be impossible to find another pixel on the opposite contour, such that they are exactly mathematically symmetrical, due to the digitization. Thus, a fundamental problem needs to be solved is to define symmetries in the discrete domain, so that the skeleton of a digital shape can be computed correctly.

All in all, although more than 300 skeletonization algorithms have been proposed, the improvement is still required, since the existing approximation algorithms of skeletonization often suffer from one or more of the following drawbacks [Chang and Yan, 1999; Ge and Fitzpatrick, 1996; Lam et al., 1992; Smith, 1987; You and Tang, 2007]:

- (1) It may take a long time to skeletonize a high-resolution image.
- (2) Skeletons may not be centred inside the underlying shapes.
- (3) Skeletons are sensitive to noise and shape variations, such as rotation and scaling, etc.
- (4) A shape and its skeleton may have different number of connected components.
- (5) Skeleton may contain artifacts such as noisy spurs and spurious short branch between split junction points.
- (6) Skeleton branches may be seriously erode.
- (7) In addition, most methods are suitable for only the binary images rather than the gray images.

To overcome the above problems, a novel wavelet-based method is presented in [Tang and You, 2003; You and Tang, 2007]. In this way, a new wavelet function called Tang-Yang wavelet, which has been constructed by our research group [Yang et al., 2001; Yang et al., 2003a], is applied; a new symmetry analysis, which is called symmetry analysis of maximum moduli of wavelet transform, is also proposed. This approach benefits from the following desired characteristics of new wavelet function investigated.

- The position of the local maximum moduli of the wavelet transform with respect to the Ribbon-like shape is independent of the gray-levels of the image.
- The local maximum moduli of the wavelet transform of the Ribbon-like shape form two new contours, when the appropriate scale of

the wavelet transform is selected according to the width of the shape. Meantime, the new contours are located symmetrically at the both sides of the original image, and keep the same topological and geometric properties.

- The distance from one contour to opposite one, which is completely independent of the width of the shape, and equals to the scale of the wavelet transform.

Based on the above symmetry analyses, a new skeleton called wavelet skeleton can be defined, which not only remain the good properties provided by the existing methods, but also improves the technique in the following respects: (1) Generally, it is easy to determine the symmetric points from the symmetric curves proposed; and a skeleton may be centred exactly inside the underlying shape; (2) The computation of the skeleton of a shape is readily simple, moreover, it takes less time to perform the skeleton; (3) A skeleton representation is robust against noise and insensitive to linear geometric transformations, such as translation, rotation and scaling; (4) The image to be processed may be extended to any gray levels.

This chapter is organized as follows: In Section 8.1, the Tang-Yang wavelet function is introduced. In Section 8.2, a corresponding symmetry analysis will be developed based on three important properties of the maximum moduli of the wavelet transform. The characterization of the skeleton produced by the proposed method will also be compared with the existing ones. In Section 8.3, the wavelet skeleton and its implementation will further be discussed. A set of techniques for modifying the artifacts of the primary wavelet skeleton will also be given. An algorithm to perform the proposed new scheme to extract the skeleton of the Ribbon-like shapes as well as experiments will be illustrated in Section 8.4.

8.1 Tang-Yang Wavelet Function

It's well known that to settle the contour or boundary of shape is a key of extracting the skeleton. With the growth of wavelet theory, the wavelet transform has been found to be a remarkable mathematical tool to analyze the singularities including the edges, in particular, the sharp boundary of the shape, and further, to detect them effectively [Canny, 1986; Mallat, 1998; Mallat and Hwang, 1992; Mallat and Zhong, 1992; Tang et al., 2000].

Although lots of wavelet functions have been found so far, the construction of a appropriate one according to the application in practice is still a great challenge for the worldwide researchers. In order to extract the skeleton of a shape, detecting the boundary or contour of the shape is required. In this chapter, a novel wavelet function called Tang-Yang wavelet, which is constructed in our work [Yang et al., 2001; Yang et al., 2003a], is utilized.

As we have known, the edge points of a digital image are the pixels, which are as close as possible to the center of the true edge but absolute. Hence, when we extract the central point of shape via its edges, the location of edge pexels may be shifted around the centre of the true edge without loosing the human vision. In our motivation, when we detect the edges from the local maxima of the wavelet transform modulus, the desired operator not only detect the singular points of the signal but also adjust properly the location of the edge points around the centre of the true edge, which depends strongly on the scale of the wavelet transform. Therefore, a novel wavelet function $\psi(x, y)$ is a considerable candidate, which satisfies the mentioned extra requirement besides those conditions, that Gaussian function and quadratic spline do. Namely, when it is used as the wavelet function, the location of maximum moduli of the wavelet transform is relatively dependent of the scale of the wavelet transform. Obviously, it is preferable for the boundary detection of the Ribbon-like shapes. An odd function $\psi(x)$ (for the sake of simplicity, we discuss one dimension case first) is considered as the new wavelet function. Consequently, we consider an odd function $\psi(x)$, which is defined on $(0, \infty)$ by

$$\psi(x) := \begin{cases} \psi_1(x) + \psi_2(x) + \psi_3(x) & x \in (0, \frac{1}{4}) \\ \psi_2(x) + \psi_3(x) & x \in [\frac{1}{4}, \frac{3}{4}) \\ \psi_3(x) & x \in [\frac{3}{4}, 1) \\ 0 & x \in [1, \infty) \end{cases} \quad (8.1)$$

where

$$\begin{cases} \psi_1(x) = -\frac{2}{\pi}(-8x \ln \frac{1+\sqrt{1-16x^2}}{4x} + \frac{1}{2x}\sqrt{1-16x^2}) \\ \psi_2(x) = -\frac{2}{\pi}(8x \ln \frac{3+\sqrt{9-16x^2}}{4x} - \frac{3}{2x}\sqrt{9-16x^2}) \\ \psi_3(x) = -\frac{2}{\pi}(-4x \ln \frac{1+\sqrt{1-x^2}}{x} + \frac{4}{x}\sqrt{1-x^2}) \end{cases}$$

as the candidate of 1-D wavelet function. Apparently, function $\phi(x) :=$

$\int_0^x \psi(x)dx$ is an even function with compactly supported on $[-1, 1]$, and $\phi'(x) = \psi(x)$ holds as well. The graphical description of functions $\psi(x)$ and $\phi(x)$ can be found in Fig. 7.3 in Chapter 7. The 2-D wavelet functions are given by

$$\begin{cases} \psi^1(x, y) := \frac{\partial}{\partial x} \theta(x, y) = \phi'(\sqrt{x^2 + y^2}) \frac{x}{\sqrt{x^2 + y^2}} \\ \psi^2(x, y) := \frac{\partial}{\partial y} \theta(x, y) = \phi'(\sqrt{x^2 + y^2}) \frac{y}{\sqrt{x^2 + y^2}} \end{cases} \quad (8.2)$$

and are graphically displayed in Fig. 7.5 (Chapter 7).

The gradient direction and the amplitude of the wavelet transform are denoted respectively by

$$\nabla W_s f(x, y) := \begin{pmatrix} W_s^1 f(x, y) \\ W_s^2 f(x, y) \end{pmatrix}, \quad (8.3)$$

and

$$|\nabla W_s f(x, y)| := \sqrt{|W_s^1 f(x, y)|^2 + |W_s^2 f(x, y)|^2}. \quad (8.4)$$

Here, $|\nabla W_s f(x, y)|$ is called modulus of the wavelet transform at point (x, y) . The boundary of the shape can be detected by locating the local maxima of the wavelet transform modulus. The details of the selection of the new function and analysis of its property as well as the comparisons with Gaussian function and quadratic spline are presented in [Yang et al., 2003a].

8.2 Characterization of the Boundary of a Shape by Wavelet Transform

In this section, some significant characteristics of the Ribbon-like shape with respect to the local maximum moduli of the wavelet transform based on the above wavelet function will be presented, namely:

- Gray-level invariant — The local maximum moduli of the wavelet transform of the Ribbon-like shape take place at the same points regardless of the different gray-levels of the image.
- Slope invariant: The local maximum moduli of the wavelet transform of the Ribbon-like shape are independent of the slope of the object.

- Width invariant — The maximum moduli of the wavelet transform of the Ribbon-like shape produce two new lines around the original boundary. Moreover, the location of maximum moduli are independent of the width of the shape, namely, the distance between the two new lines depends on only the scale of the wavelet transform rather than width of the shape under certain circumstance.
- Symmetry — The location of maximum moduli of the wavelet transform cover exactly the points of the original boundary if the transform scale s equals to the width of the shape. Moreover, when the scale is bigger than or equals to the width of a shape, the two new lines which formed by the maximum moduli of the wavelet transform are exactly symmetrical with respect to the central line of the shape.

The above properties can be summarized in the following theorem, the proofs can be find in [Yang et al., 2003a].

Theorem 8.1 Let l_d be a straight segment of the Ribbon-like shape with width d and central line l . If the scale of the wavelet transform $s \geq d$, then the local maximum moduli of the wavelet transform using wavelet function of Eq. (8.2) generate two new periphery lines around the original segment, which have the following properties: The two new lines are exactly symmetric with respect to the central line of the segment; the distance between the new lines equals to scale s , in other words, the location of the maximum moduli of the wavelet transform depends completely on scale s , and the location of the central line of the segment. Moreover, the two new lines possess the same gradient direction as the central line does. Furthermore, if and only if scale s equals to the width of the shape, the locations of points of the maximum moduli lie exactly on the boundaries of the shape.

Therefore, the following definitions are natural outcomes:

Definition 8.1 In wavelet transform, if scale s is bigger than or equals to the width of the Ribbon-like shape, the points of the maximum moduli will generate two new lines locating on the periphery of the shape. Moreover, they are local symmetrical with respect to the central line of the shape. This symmetry is called *maximum moduli symmetry (MMS)*.

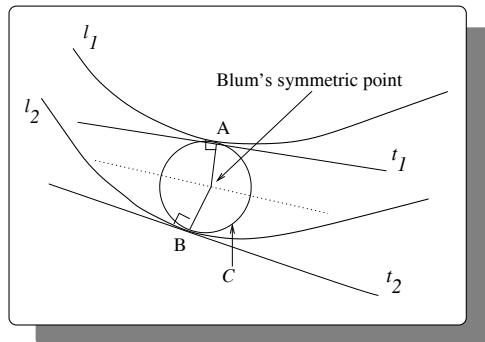
Definition 8.2 The *wavelet skeleton* of the Ribbon-like shape is defined as the curve of all connective midpoints of the segment lines, which are

connected by all pairs of the symmetrical maximum moduli of the wavelet transform of the shape.

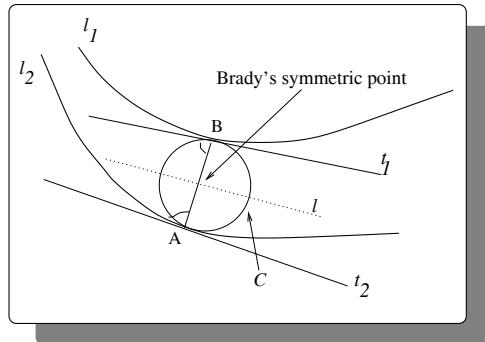
To illustrate the above definitions more clearly, now a brief comparison of our new symmetry analysis with the traditional ones, i.e. Blum's symmetric axis transform (SAT)[Blum, 1967], Brady's smoothed local symmetry (SLS)[Brady, 1983] and Leyton's process-inferring symmetry analysis (PISA) [Leyton, 1988], will be presented.

The graphical descriptions of four symmetry analyses are shown in Fig. 8.1 and Fig. 8.2. Here, l_1 and l_2 are two opposite boundaries of the shape. A circle labelled by C is placed between l_1 and l_2 such that it is simultaneously tangential to both boundaries at points A and B in Figs. 8.1(a), (b) and Fig. 8.2(c). Blum's symmetric axis transform (SAT) is shown in Fig. 8.1(a), and its symmetric point of the local symmetry formed by A and B is defined as the center of circle C . Hence, its corresponding skeleton is defined as the locus of the central points of the maximal inscribed symmetric circles of the shape. It is obvious that the interior of circle C must lie entirely inside either the foreground or the background of the image. Most of the existing skeletonization algorithms are based on the concept of this symmetry analysis, such as the grassfire technique.

In SLC symmetry analysis, Brady defines the symmetric point of the local symmetry by the intersecting point of two lines, the mirror l and segment \overline{AB} , i.e. the midpoint of \overline{AB} . Its corresponding skeleton is the locus of the mid-points of the symmetric lines of the symmetric circles as shown in Fig. 8.1 (b). In Brady's original proposal, a basic criterion, called equal-angle criterion, asks that angle (θ) between symmetric line \overline{AB} and a tangent line t_1 equals to the angle between \overline{AB} and another tangent line t_2 . More precisely, two points, A and B , on a planar curve form a local symmetry if the angle between vector \overrightarrow{AB} and the normal of the curve at B equals to the angle between vector \overrightarrow{BA} and the normal of the curve at A . However, the interior of circle C is not required to stay inside the shape completely. This characteristic can be either an advantage or a disadvantage compared with SAT and PISA methods. The advantage is that a wider variety of situations are considered. For instance, SAT fails to produce the minor axis of an ellipse, which can be generated by SLS [Brady and Asada, 1993]. The disadvantage is that a symmetric axis from SLS may not lie inside the shape entirely [Saint-Marc et al., 1993]. The additional measures have to be used while the perceptually salient axes are



(a) Blum's Symmetric Axis Transform

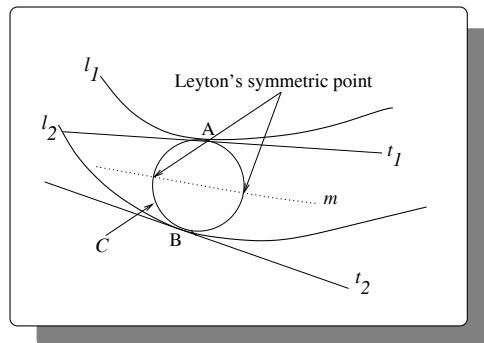


(b) Brady's Smoothed Local Symmetry

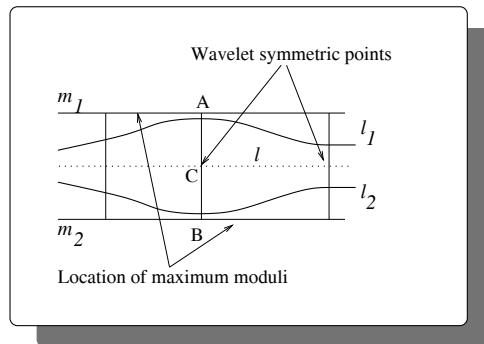
Fig. 8.1 The illustration of Symmetry Analyses.

selected to characterize the shape. In addition, the question of whether the locus of symmetries lies within the shape or not depends not only on the choice of locality of the symmetry point (e.g. centre of circle, mid-chord) but also on whether the bitangent circle is forced to be inscribed (entirely contained within the shape). Typically SLS does not enforce this constraint whilst the SAT of Blum does.

Leyton's symmetric point is defined as the intersecting point of 'mirror' m and symmetric circle C , i.e. the midpoint of the arc \widehat{AB} , and the corresponding skeleton is defined by the locus of the midpoints of the symmetric arcs of the circles, which can be shown in Fig. 8.2(c). Since there are two



(c) Leyton's Process—Inferring Symmetry



(d) Maximum Moduli Symmetry of Wavelet Transform

Fig. 8.2 The illustration of Symmetry Analyses (Continuous).

symmetric arcs, and two symmetric points associated with the symmetric pair A and B , PISA becomes ambiguous when it comes to decide which symmetric point is appropriate to represent the symmetry property of the object.

In our proposed symmetry analysis, which is shown in Fig. 8.2(d), the symmetric point is defined as the midpoint of segment \overline{AB} . Meanwhile, the corresponding skeleton of a shape is defined as the locus of the midpoints of the segments connecting any pair of points, which lie on two opposite symmetric maximum moduli m_1 and m_2 . The first advantage is that a symmetric axis must be inside the shape entirely. Moreover, the skeleton

of the shape can be mathematically centred inside the underlying shape. Another obvious advantage is that its implementation is relatively easy, in other words, the computation of the skeleton of the shape is simple and direct, because its two symmetric maximum moduli lines are easily obtained by applying the wavelet transform to the shape. Moreover, the distance between the two symmetric maximum moduli lines completely depends on the scale of the wavelet transform, thus, it is relatively easy to locate midpoint of the segment in practice. As far as the symmetry analyses in either SAT or SLS or PISA, the implementation is not straightforward and relatively difficult. Because it is hard to fix circle C between l_1 and l_2 such that it is simultaneously tangential to both contours at A and B , which can be found in Figs. 8.1(a)-(b) and Fig. 8.2(c), especially in the discrete domain. Moreover, it also implicates that the above indirect approaches (SAT, SLS and PISA) suffer from the higher computational cost. Our approach is a direct method, in which the points of the modulus maxima of the wavelet transform are just contour points of the shape, therefore, the computational cost is mainly determined by this phase itself. The extra computational cost of searching the tangent circles of the contours in SAT, SLS and PISA can be saved.

8.3 Wavelet Skeletons and Its Implementation

In this section, a set of schemes for extracting wavelet skeletons of the Ribbon-like shapes will be presented, which includes two parts, namely, (1) implementation of the wavelet transform in the discrete domain to generate a primary wavelet skeleton, and (2) modification of the primary wavelet skeleton.

8.3.1 Wavelet Transform in the Discrete Domain

In practice, the images to be processed are in discrete world. In this sub-section, we will give wavelet transform formula in the discrete domain and calculate the corresponding wavelet coefficients. In fact, the wavelet transform formula can be re-written as follows:

$$W_s^i f(n, m) = \int \int f(u, v) \psi_s^i(n - u, m - v) du dv$$

$$\begin{aligned}
&= \sum_{k,l} f(k,l) \int_k^{k+1} \int_l^{l+1} \psi_s^i(n-u, m-v) du dv \\
&= \sum_{k,l} f(n-k-1, m-l-1) \psi_{k,l}^{s,i}, \quad (i = 1, 2),
\end{aligned}$$

where

$$\begin{aligned}
\psi_{k,l}^{s,i} &= \int_k^{k+1} \int_l^{l+1} \psi_s^i(u, v) du dv \\
&= \int_{k/s}^{(k+1)/s} \int_{l/s}^{(l+1)/s} \psi^i(u, v) du dv, \quad (i = 1, 2).
\end{aligned}$$

It is clear that the computation of the above formula is a discrete convolution. For a small scale s , it can be performed directly. But for a large scale s , considering the amount of computation, FNT (fast numerous transform) or other fast algorithms can be utilized instead. The details can be found in [Tang et al., 2000]. In the following, we will discuss the calculation of the coefficients $\{\psi_{k,l}^{s,i}\}$.

For the wavelet functions defined in Eq. (8.2), since

$$\psi^1(u, v) = \phi'(\sqrt{u^2 + v^2}) \frac{u}{\sqrt{u^2 + v^2}} = \psi^2(v, u),$$

we can deduce that

$$\psi_{k,l}^{s,1} = \psi_{l,k}^{s,2}, \quad \psi_{-k,l}^{s,1} = -\psi_{k-1,l}^{s,1}, \quad \psi_{k,-l}^{s,1} = \psi_{k,l-1}^{s,1}, \quad \psi_{-k,-l}^{s,1} = -\psi_{k-1,l-1}^{s,1}.$$

Consequently, we further need to calculate only $\psi_{k,l}^{s,1}$ for all positive integers k and l . By $\psi^1(u, v) = \frac{\partial}{\partial u}[\phi(\sqrt{u^2 + v^2})]$, we have

$$\psi_{k,l}^{s,1} = \phi_{l,k+1}^s + \phi_{l+1,k}^s - \phi_{l+1,k+1}^s - \phi_{l,k}^s,$$

where

$$\phi_{k,l}^s = \int_{k/s}^{\infty} \phi \left(\sqrt{v^2 + \left(\frac{l}{s} \right)^2} \right) dv.$$

Hence, our question turns to calculate $\phi_{k,l}^s$ for all non-negative integers k and l . Since

$$\psi^1(u, v) = \phi'(\sqrt{u^2 + v^2}) \frac{u}{\sqrt{u^2 + v^2}} = \psi(\sqrt{u^2 + v^2}) \frac{u}{\sqrt{u^2 + v^2}},$$

where ψ is defined by Eq. (8.1), for any non-negative integers k and l satisfying $k^2 + l^2 < s^2$, we have

$$\phi_{k,l}^s = \int_{k/s}^{\infty} \phi(\sqrt{v^2 + (l/s)^2}) dv = \int_{\frac{l}{s}\sqrt{k^2+l^2}}^1 \left[\frac{k}{s} - \sqrt{v^2 - (l/s)^2} \right] \psi(v) dv.$$

On the other hand, it is easy to see that $\phi_{k,l}^s = 0$ for all integers k and l satisfying $k^2 + l^2 \geq s^2$ due to the compact support $[-1, 1]$ of $\phi(x)$.

We can calculate all coefficients $\phi_{k,l}^s$ numerically for non-negative integers k and l . The positive integer n in the formula can be set to be so large that the error is smaller than any prior number. The possible nonzero items for $s = 2, 4, 6, 8$ are listed in the following five tables.

$l \setminus k$	$l = 0$	$l = 1$
$k = 0$	0.2500	0.1250
$k = 1$	0.0497	0.0111

Table 8.1 The nonzero coefficients $\{\phi_{k,l}^s\}$ for $s = 2$.

$k \setminus l$	$l = 0$	$l = 1$	$l = 2$	$l = 3$
$k = 0$	0.2500	0.2292	0.1250	0.0208
$k = 1$	0.1468	0.1206	0.0552	0.0060
$k = 2$	0.0497	0.0366	0.0111	0.0003
$k = 3$	0.0047	0.0026	0.0002	0

Table 8.2 The nonzero coefficients $\{\phi_{k,l}^s\}$ for $s = 4$.

$k \setminus l$	$l = 0$	$l = 1$	$l = 2$	$l = 3$	$l = 4$	$l = 5$
$k = 0$	0.2500	0.2438	0.2022	0.1250	0.0478	0.0062
$k = 1$	0.1831	0.1718	0.1333	0.0767	0.0257	0.0025
$k = 2$	0.1106	0.1003	0.0723	0.0367	0.0094	0.0005
$k = 3$	0.0497	0.0436	0.0281	0.0111	0.0017	0.0000
$k = 4$	0.0133	0.0109	0.0056	0.0014	0.0000	0
$k = 5$	0.0011	0.0008	0.0002	0.0000	0	0

Table 8.3 The nonzero coefficients $\{\phi_{k,l}^s\}$ for $s = 6$.

$k \setminus l$	$l = 0$	$l = 1$	$l = 2$	$l = 3$
$k = 0$	0.2500	0.2474	0.2292	0.1849
$k = 1$	0.2006	0.1950	0.1741	0.1358
$k = 2$	0.1468	0.1403	0.1206	0.0902
$k = 3$	0.0935	0.0882	0.0733	0.0517
$k = 4$	0.0497	0.0462	0.0366	0.0236
$k = 5$	0.0199	0.0180	0.0132	0.0072
$k = 6$	0.0047	0.0041	0.0026	0.0011
$k = 7$	0.0004	0.0003	0.0001	0

Table 8.4 The nonzero coefficients $\{\phi_{k,l}^s\}$ for $s = 8$, $l = 0 - 3$.

$k \setminus l$	$l = 4$	$l = 5$	$l = 6$	$l = 7$
$k = 0$	0.1250	0.0651	0.0208	0.0026
$k = 1$	0.0884	0.0433	0.0126	0.0013
$k = 2$	0.0552	0.0244	0.0060	0.0004
$k = 3$	0.0287	0.0107	0.0020	0
$k = 4$	0.0111	0.0032	0.0003	0
$k = 5$	0.0026	0.0004	0	0
$k = 6$	0.0002	0	0	0
$k = 7$	0	0	0	0

Table 8.5 The nonzero coefficients $\{\phi_{k,l}^s\}$ for $s = 8$, $l=4-7$.

8.3.2 Generation of Wavelet Skeleton in the Discrete Domain

Based on the discussion in Section 8.2, a set of schemes to detect the segment of shape can be designed. The result is also valid for general Ribbon-like shapes since a short segment of a shape can be regarded as a straight line. In fact, wavelet transform is essentially a local analysis. Therefore, the result of Theorem 8.1 can be applied to the general Ribbon-like object.

Maximum moduli symmetry is defined in the continuous domain as mentioned in the previous discussion. In fact, the exact symmetries between contour pixels may not exist in the discrete domain. When a continuous shape is digitized, its boundary is sampled by discrete points. For some contour pixels of a digital shape, it may not be possible to find their symmetrical counterparts on the opposite contour to form local symmetries.

Even if a pair of symmetric contour pixels can be found precisely, their corresponding symmetrical central point may not exist.

In addition, in the continuous domain, the gradient direction of a function can be accurately counted in mathematics. However, in contrast, it is difficult to express it exactly in the discrete form. Consequently, we will devote our efforts on finding the best way to locate the symmetric maximum moduli points and the corresponding central points.

First, we determine the gradient direction of a digital signal and its corresponding local maximum. Fortunately, in the equi-spaced sampling, only eight adjacent pixels are around a point. Hence, only the nearest eight points will be taken into account. It is said that the discrete image has eight gradient directions. Therefore, a plane can be divided into eight sectors. A gradient direction is defined by

$$\alpha_s := \arctan\left(\frac{\partial(f * \theta_s)(x, y)}{\partial y} / \frac{\partial(f * \theta_s)(x, y)}{\partial x}\right),$$

where s denotes the wavelet transform scale.

When α_s falls into a sector, it will be quantified to a certain vector, which is represented by a central line of that sector. It indicates a direction of the gradient, and along this direction, the local maximal moduli can be achieved. The effect of any opposite gradient direction is the same. Thus, only 4 codes are needed to be used to code these different directions. The tangent of each direction, which is described by $\operatorname{tg}\alpha_s = \frac{\partial(f * \theta_s)(x, y)}{\partial y} / \frac{\partial(f * \theta_s)(x, y)}{\partial x}$, falls into one of the following intervals:

$$[-1 - \sqrt{2}, 1 - \sqrt{2}), \quad [1 - \sqrt{2}, \sqrt{2} - 1), \quad [\sqrt{2} - 1, \sqrt{2} + 1),$$

$$[\sqrt{2} + 1, +\infty) \cup (-\infty, -1 - \sqrt{2}).$$

The above gradient code is called *Gradient Code of Wavelet Transform (GCWT)*. It will play an important role not only in locating the central points (primary skeleton points), but also in modifying artifacts of the primary wavelet skeleton whereafter.

On the other hand, for every maximum moduli point, even though the corresponding symmetric counterpart cannot be achieved or not exist along its gradient direction, one or more maximum moduli points may appear in a sector, which contains certain GCWT or gradient direction. Hence, from these points, we can select one, where the distance from this point to the

original maximum moduli point approximates to scale s of the wavelet transform, as the symmetric counterpart of the original maximum moduli point. Meantime, the midpoint or approximate midpoint of the segment, which is connected through the pair of maximum moduli symmetric points, is regarded as the symmetric central point, namely, the skeleton point. In practice, this approach can be implemented accurately and easily in mathematics.

An important issue in skeletonization is to reduce the noise. As we have known, the noise always distributes randomly. From the point of view of the statistics, the average value of the noise is nearly a constant in a certain area. In general, we can suppose the value of this constant is zero, and take a weighted mean to the signal in this area. This action can be regarded as a low-pass filtering. In this way, the noise will be eliminated considerably. Virtually, $\theta(x, y)$ is a smoothing operator, which convolutes with $f(x, y)$, thus, the noise can be reduced. In fact, we often have some prior information to identify the difference between the singularities of signal and the noise. The technique of reducing noise based the detection of the wavelet transform maxima has been developed [Canny, 1986; Mallat, 1998; Mallat and Hwang, 1992; Mallat and Zhong, 1992; Tang et al., 2000; You and Tang, 2007]. In our practice, a threshold processing is applied to reduce the noise and remain the edge information. In this processing, after calculating the modulus of the wavelet transform, a threshold is used to decide which point is an noise one to be eliminated. If the modulus of the wavelet transform is less than the threshold, then their modulus will be reset as 0. Even though the threshold cannot be automatically computed, it can be selected by the experiments. In our practice, we have $T = C_T \times M$, where, C_T is a constant to be adjusted manually according to the amount of noise or the distracting background (in our experiments, $C_T \approx 0.35$), M is the maximum modulus of the wavelet transform. Based on the threshold processing, the prospective edge or contour of the shape can be detected effectively and exactly.

8.3.3 Modification of Primary Wavelet Skeleton

Obviously, a few points on the central line produced by the above technique may be lost. Hence, the primary skeleton obtained from the above approach do not resemble somewhat human perceptions. The lost points may be resumed through a slight modification. Namely, for each lost point, we

examine the neighbour points along the normal of the gradient, if they have the same GCWT, then the lost points will be considered to be the skeleton points. Meantime, the typical junction-points and intersections of the shape, such as T-pattern, K-pattern and Cross-pattern intersections, etc. as shown in Fig. 8.3 are applied to the modification of the primary wavelet skeleton.

Let us look at an example as shown in Fig. 8.5(c), by using the proposed foregoing central line searching technique to letter “H”, some skeleton lines in the junctions are lost. In fact, this is a key problem, which is needed to be solved in any skeletonization method. Fortunately, the GCWT plays an extremely important role in our new technique. Six typical junctions are listed in Fig. 8.3, which may arise the pixel-losing in the primary skeleton. Some modifying techniques are proposed in this chapter to reduce the pixel-losing. It will be noted that they still depend completely on the relevant results of the above wavelet transform.

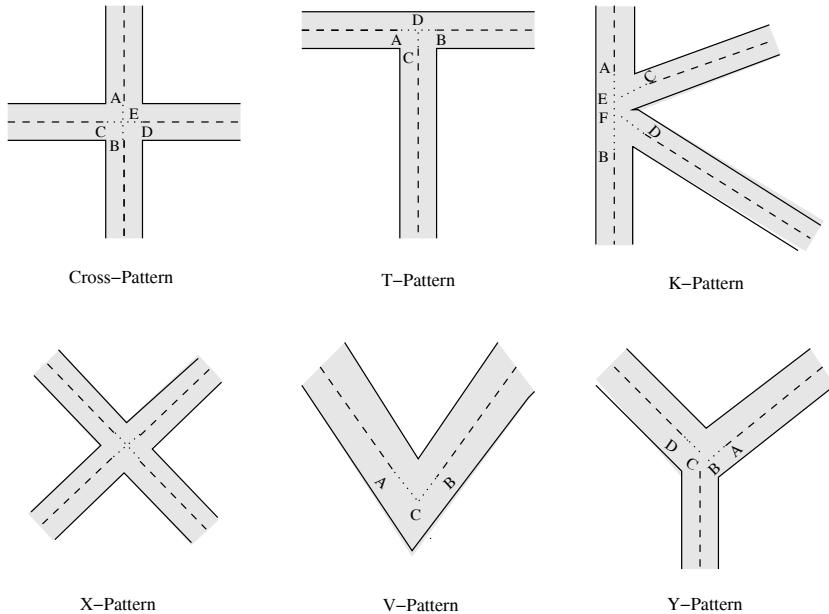


Fig. 8.3 All sorts of junction-points.

For the cross-patterns as shown in Fig. 8.3, the points between *A* and

B of the wavelet skeleton are truncated. Fortunately, it is easy to find these points through the above wavelet transform since they have exactly the same gradient direction or the same GCWT. Based on the foregoing definition of the wavelet symmetry, one can easily accept the fact that the distance between points A and B equals (approximate) to scale s of the wavelet transform. Consequently, in practice, from a terminal point A , we can find another point B along its planar normal direction or the same GCWT. A similar process can be performed between C and D . It is obvious that this modification approach is suitable for the shape with X-pattern intersection as well.

The primary wavelet skeleton of the shape with T-pattern junction is also shown in Fig. 8.3. Here, the points between A and B as well as C and D are lost. Likewise, the lost points between A and B in the primary skeleton possess the same gradient direction as that of points A and B . Moreover, along either its gradient direction or opposite direction, for every such point, one and only one corresponding point has the maximum moduli, and the distance between them is half of scale s . Consequently, we only need to search the lost points, which satisfy the above two conditions, and then retrieve these points from the primary skeleton line to find the modifying skeleton locus. It is noted that if the distance from some lost point to the location of the maximum moduli along the gradient direction or normal direction at the lost point is less than half of the scale s of the wavelet transform, then the retrieve process needs to be stopped. For example, since the distance from the next neighbour of point D along its gradient direction is less than half of scale s for the T-pattern of Fig. 8.3, retrieve process from point C to D has to end at point D .

Virtually, the analogical process can be done for K-pattern, V-pattern and Y-pattern, which are also shown in Fig. 8.3. Apparently, the modified rules depend strongly on the gradient direction and location of the local maximum modulus points. In fact, all retrieve and extend processes are strictly limited within the area enclosed by the local maximum modulus pixels. Therefore, it does not result in the incorrect connection of unrelated contour or create new artificial branch in even closely neighbouring shapes.

Generally, if the terminal or end point lying on the locus of the primary skeleton satisfies one of the following conditions, it will be called unstable terminal point.

- **Condition 1:** For the terminal or end point lying on the locus of

the primary skeleton, along the vertical direction of its gradient, there exists a neighbour point with the same GCWT such that the distance from itself to the nearest location of the maximum moduli is more than scale s .

- **Condition 2:** Along the vertical direction of the gradient, there exists another terminal or end point with the same GCWT lying on another segment of the locus of the primary skeleton, every point between these two end points lies completely inside the shape, and the distance between the two terminal points equals to scale s of the wavelet transform.
- **Condition3:** For the next neighbour point of a terminal point along the vertical direction of its gradient, there exists only one location of maximum moduli point along its gradient direction or opposite direction such that the distance from this point to the location of maximum moduli is half of scale s , further, along the planar normal of the gradient direction at the point, no maximum moduli point exists or the distance from this point to the location of maximum moduli is more than half of scale s of the wavelet transform.

As a result, the above modifying process for the primary wavelet skeleton can be rewritten as the following algorithm :

Input: An image contains the primary wavelet skeleton with corresponding maximum moduli resource.

Output: An image contains the modified wavelet skeleton.

REPEAT

FOR *Every unstable terminal point in the image*

IF *It satisfies Condition 1*

THEN *Retrieve its next neighbour point locating at the vertical direction of its gradient as a modified skeleton;*

ELSE IF *It satisfies Condition 2*

THEN *Retrieve s points between two terminal points as a part of modified skeleton;*

ELSE IF *It satisfies Condition 3*

THEN *Retrieve its next neighbour point locating at the vertical direction of its gradient as one modified skeleton.*

END IF;

END FOR;
UNTIL *No unstable terminal point is detected.*
END.

In addition, it has been noted that a primary wavelet skeleton representation can also be used to recover the contours of the underlying shape. This is because that the contour information of the shape is recorded in its primary representation. In other words, the wavelet skeleton preserves its width and direction information of the shape. This property is often desirable since a shape is completely specified by its contours, and further, by its wavelet skeletons.

8.4 Algorithm and Experiment

To implement the proposed method to extract the wavelet skeletons of the Ribbon-like shapes, in this section, an algorithm will be provided followed by several experiments.

8.4.1 Algorithm

The algorithm based on the maximum moduli analysis of the wavelet transform (MMAWT) method to extract wavelet skeleton of the Ribbon-like shape in an image is designed as follows:

Algorithm 8.1 Let $f(x, y)$ be an image containing Ribbon-like shapes, and scale $s > 0$,

- Step 1** Select the suitable scale of the wavelet transform according to the width of the Ribbon-like shape.
- Step 2** Calculate wavelet transforms $\{W_s^1 f(x, y), W_s^2 f(x, y)\}$ using the wavelets defined by Eq. (8.2).
- Step 3** Calculate modulus $|\nabla W_s f(x, y)|$ of the wavelet transform and the gradient direction $f_{gradient}$.
- Step 4** Take threshold T according to the amount of noise and background in the original image and proceed with threshold on the modulus image (if necessary).
- Step 5** For each point (x, y) , its modulus of the wavelet transform is compared with one of its neighbouring points along its gradient direction, if its modulus arrives at the maximum, it will recorded as the

local modulus maximum f_{locmax} .

- Step 6** For each point (x, y) with local maximum, search the point whose distance to (x, y) along the gradient direction is s . If it is a point with the local maximum, the central point is detected.
- Step 7** The primary skeleton is formed by all the points detected in Step 4.
- Step 8** Modify the above primary wavelet skeleton according to the foregoing modification program to obtain the final wavelet skeleton.

Obviously, the implementation of the above algorithm is easy, simple and fast due to the following reasons: It is well known that performing wavelet transform for an image is easy as long as the discrete transform formula is given. Secondly, only the points of the maximum moduli of the wavelet transform, i.e. contour pixels are considered in Steps 3, 4 and 5. Hence, it is expected to be fast. Finally, the skeletons produced from the proposed algorithm are not sensitive to noise and shape variations such as rotation and scaling.

However, the selection of the proper scale according to the width of the shape is tough and depends on the experience in practice. It is emphasized that the properties of the modulus maximum of the wavelet transform still hold as long as the scale $s \geq d$. Therefore, in practice, the scale can be selected to be bigger than the true value, which is accepted within the certain degree.

8.4.2 *Experiments*

In this subsection, we focus on the verification of the effectiveness of the skeletonization based on the maximum moduli analysis of wavelet transform (MMAWT) method by experiments. Three types of the experiments have been done, namely, (1) production of maximum moduli image of a shape, (2) extraction of primary wavelet skeleton of the Ribbon-like shape, and (3) modification of the primary wavelet skeleton. The images used in the experiments vary in types, noise, grey levels, etc. Some examples will be presented below.

An interesting example is shown in Fig. 8.4. The original image consists of a face drawing with various widths. By carrying out the algorithm of this chapter, the skeleton is extracted, which is shown graphically in Fig. 8.4(c). The second example is illustrated in Fig. 8.5. The original im-

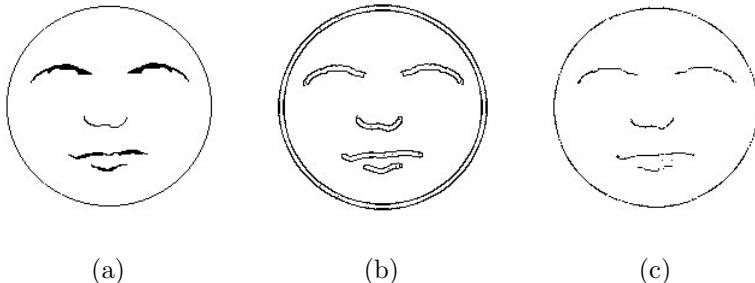


Fig. 8.4 (a) The original image; (b) The location of maximum moduli of the wavelet transform corresponding to $s = 6$; (c) The skeletons extracted by the proposed algorithm.

age, English letter “H” with varied gray levels, is given in Fig. 8.5(a). We take scale $s = 4$, the maximum moduli image is presented in Fig. 8.5(b). The primary wavelet skeleton is obtained from MMAWT and shown in Fig. 8.5(c). Obviously, the skeleton loci which locate in some intersections of the shape are lost. The modified skeleton of English letter “H” is shown in Fig. 8.5(d), where the disappeared skeleton loci, which locate in these intersections of the shape are retrieved successfully.

Some other examples are shown in Figs. 8.6, 8.7 and 8.8, where English printed characters “BesT”, and handwritings “C” and “D” are presented respectively. We take the scales of the wavelet transform, $s = 8$ and $s = 4$, respectively. The original images are shown in Figs. 8.6(a), 8.7(a) and 8.8(a), where each letter contains some strokes with a variety of widths. Fortunately, their locations of the maximum moduli have exactly the same width as shown in Figs. 8.6(b), 8.7(b) and 8.8(b). Some truncations in the primary skeleton loci are apparent in the intersections of the shapes, in addition, some truncations are also discovered in the non-intersection parts, which can be found in Figs. 8.6(c), 8.7(c) and 8.8(c). After applying the modification processes to these images, those truncated points are retrieved successfully in final wavelet skeletons as illustrated in Figs. 8.6(d), 8.7(d) and 8.8(d), which closely resemble human perceptions of the underlying shapes.

Chinese printed characters “Tian'an gate” and “material”, “glory” and “silk” are displayed respectively in Figs. 8.9 and 8.10. Obviously, each Chinese character contains some strokes with a variety of widths as shown in Figs. 8.9(a) and 8.10(a) respectively. Virtually, the width in the same stroke

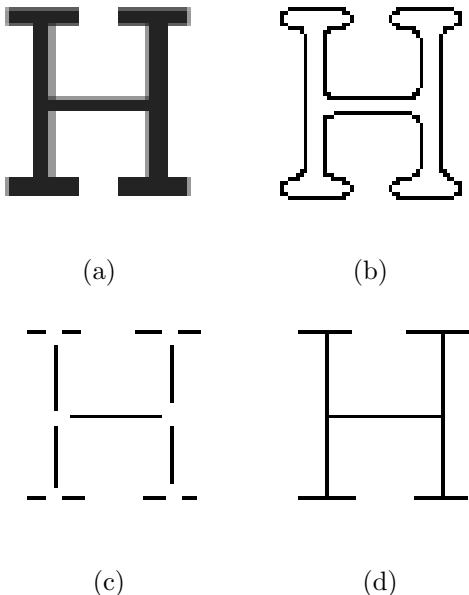


Fig. 8.5 (a) The original image; (b) The location of maximum moduli of the wavelet transform corresponding to $s = 4$; (c) The primary skeletons extracted by the proposed algorithm; (d) The final skeletons obtained by applying the modification algorithm.

maybe different. The locations of the maximum moduli of the wavelet transform with respect to these Chinese characters can be computed, and the results are given in Figs. 8.9(b) and 8.10(b). Likewise, their locations of the maximum moduli of the wavelet transform have the same width. The primary skeletons of these Chinese characters are extracted utilizing Steps 3 and 4 of the above algorithm, and presented in Figs. 8.9(c) and 8.10(c) respectively. Final skeletons are obtained by MMAWT, and shown in Figs. 8.9(d) and 8.10(d). Apparently, all final skeletons are closely resemble human perceptions of the underlying shapes as well as preserve their original topological and geometric properties.

The image of a raw chest x-ray with varying grey level distribution and noise is illustrated in Fig. 8.11(a). The raw modulus image after performing the wavelet transform is presented in Fig. 8.11(b). The corresponding output of edge detection obtained from non-threshold and threshold processing are shown respectively in Figs. 8.11(c) and (d). Applying MMAWT to Figs. 8.11(c) and (d) produces the primary skeleton as shown in Figs. 8.11(e)

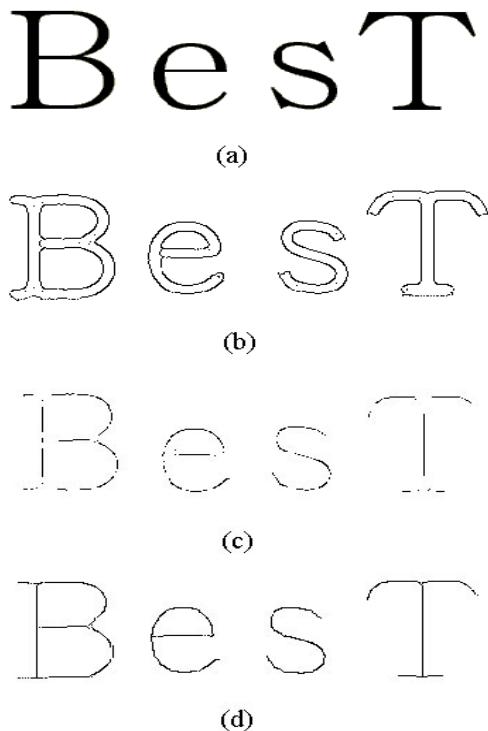


Fig. 8.6 (a) The original image of English letters; (b) The location of maximum moduli of the wavelet transform corresponding to $s = 8$; (c) The primary skeletons extracted by the proposed algorithm; (d) The final skeletons modified by applying the modification program.

and (f). By comparing the result of the non-threshold processing shown in Fig. 8.11(c) with that of the threshold one shown in Fig. 8.11(d), we can conclude that the threshold processing of wavelet transform modulus is robust against the noise. Finally, it is clear that the modified result shown in of Fig. 8.11(g) is much better than the primary skeletons shown in Fig. 8.11(f), which can verify the desired effect.

The experiment results in Figs. 8.12 and 8.13 show that the proposed approach is robust against both the noise and the affine transformation. Image in Fig. 8.12(a) contains three affine transformed patterns of the same Chinese character. In Fig. 8.13(a), the image is harmed by both the affine transformation and white “salt and pepper” noise. The wavelet

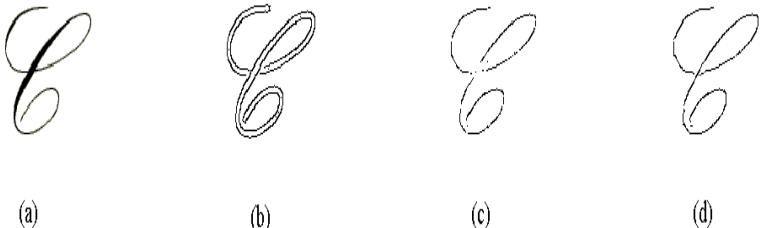


Fig. 8.7 (a) The original image of a English handwriting; (b) the location of maximum moduli of the wavelet transform corresponding to $s = 4$; (c) The primary skeletons extracted by the proposed algorithm; (d) The final skeletons modified by applying the modification program.

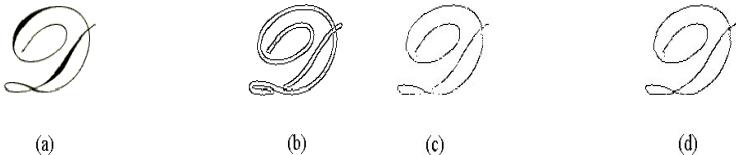


Fig. 8.8 (a) The original image; (b) The location of maximum moduli of the wavelet transform corresponding to $s = 4$; (c) The primary skeletons extracted by the proposed algorithm; (d) The final skeletons modified by applying the modification program.

transform with scale $s = 4$ is applied to the images in Figs. 8.12(a) and 8.13(a) and their corresponding raw outputs of modulus images are placed in Figs. 8.12(b) and 8.13(b) respectively. To reduce the noise affection to the edge detection, we set threshold $T = 0.39 \times M$ for eliminating the points, which have relatively weak modulus values and were caused by the noise or the distracting background. Namely, for each point, if the modulus value of the wavelet transform is less than 0.39 times of the maxima modulus, its modulus value will be reset to 0. After the threshold processing, the raw edge for the further skeleton extraction can be detected and shown in Figs. 8.12(c) and 8.13(c). The primary and modified skeletons of three Chinese characters are extracted and shown respectively in Figs. 8.12(d)-(e) and 8.13(d)-(e).

Finally, to evaluate the performance of our method, we compare it with two typical skeletonization method: ZSM method [Lam et al., 1992] and CYM method [Chang and Yan, 1999]. Our method is suitable for removing both periphery and intersection artifacts, but the ZSM and CYM methods



Fig. 8.9 (a) The original image of Chinese characters; (b) The location of maximum moduli of the wavelet transform corresponding to $s = 4$; (c) The primary skeletons extracted by the proposed algorithm; (d) The final skeletons modified by applying the modification program.

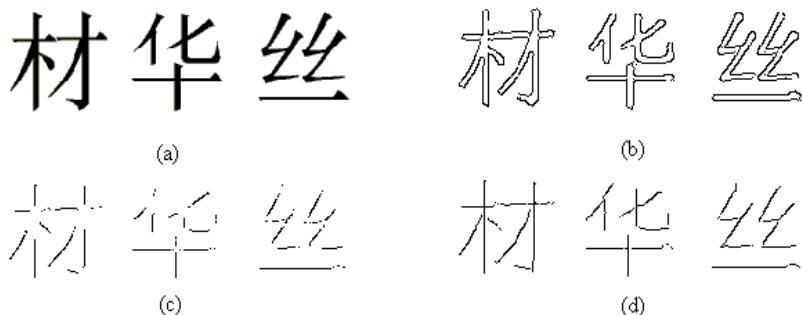


Fig. 8.10 (a) The original image of Chinese characters; (b) The location of maximum moduli of the wavelet transform corresponding to $s = 4$; (c) The primary skeletons extracted by the proposed algorithm; (d) The final skeletons modified by applying the modification program.

can remove only one type of artifact. The results of the comparison are illustrated in Fig. 8.14. The original image, which is a Chinese character, is shown in Fig. 8.14(a). The skeletons obtained from the proposed method, ZSM and CYM are shown in Figs. 8.14(b), (c) and (d) respectively. Apparently, the result obtained from the proposed method is much better than these from the ZSM and CYM, and it resembles closer human perception than those obtained from ZSM and CYM.

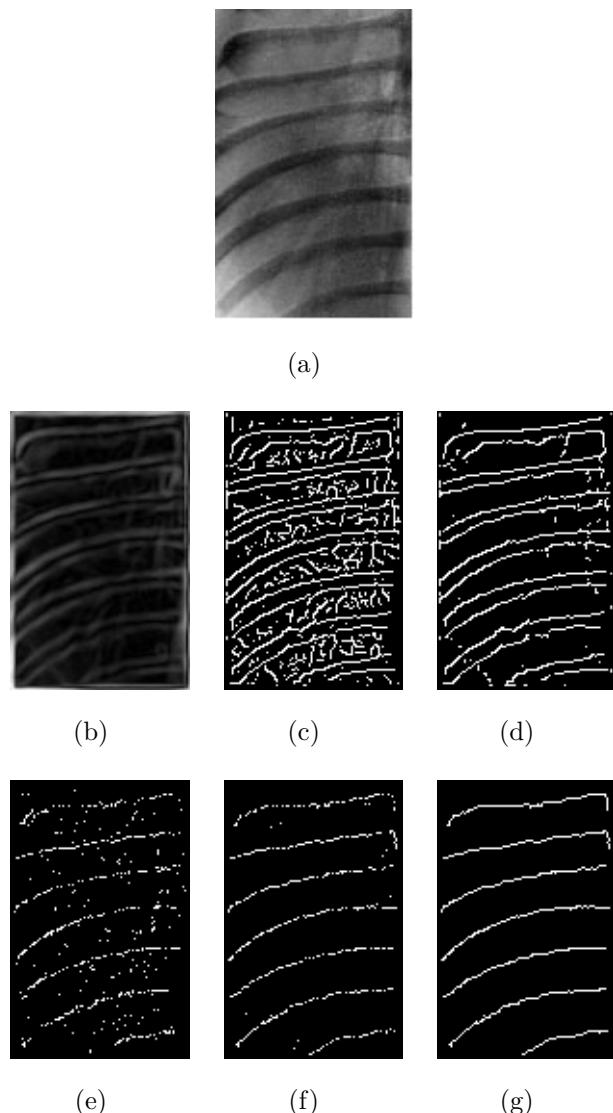


Fig. 8.11 (a) The chest x-ray image; (b) Raw of modulus of wavelet transform; (c) Non-threshold processing on modulus image, raw output of edge obtained by MMAWT; (d) The output of edge obtained after threshold processing on modulus image; (e) Raw skeleton obtained from (b); (f) The primary skeleton obtained from (c); (g) The final modified skeleton from (e).

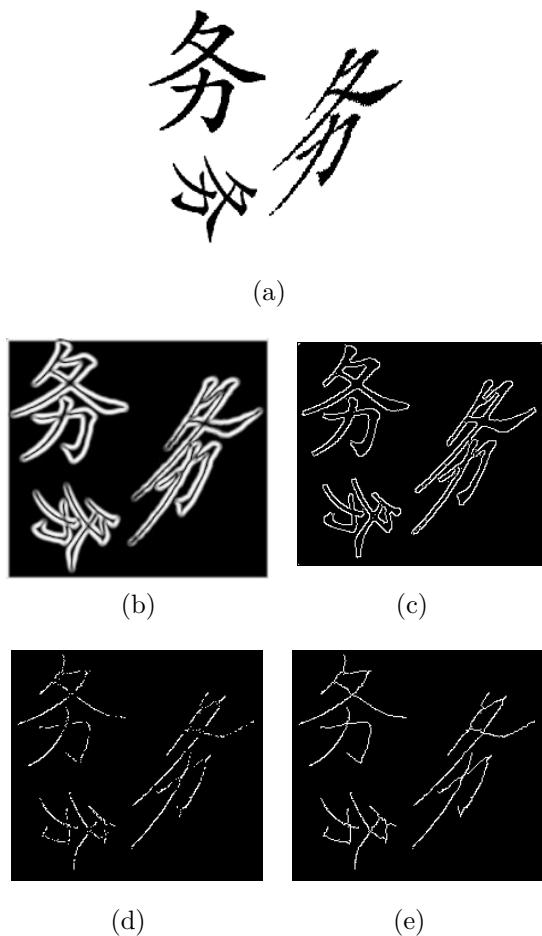


Fig. 8.12 (a) The original image of chinese character with affine transforms; (b) Raw of modulus of wavelet transform; (c) Raw output of edge detection through the modulus maximum of wavelet transform; (d) The primary skeleton extracted by the proposed algorithm; (e) The final modified skeletons.

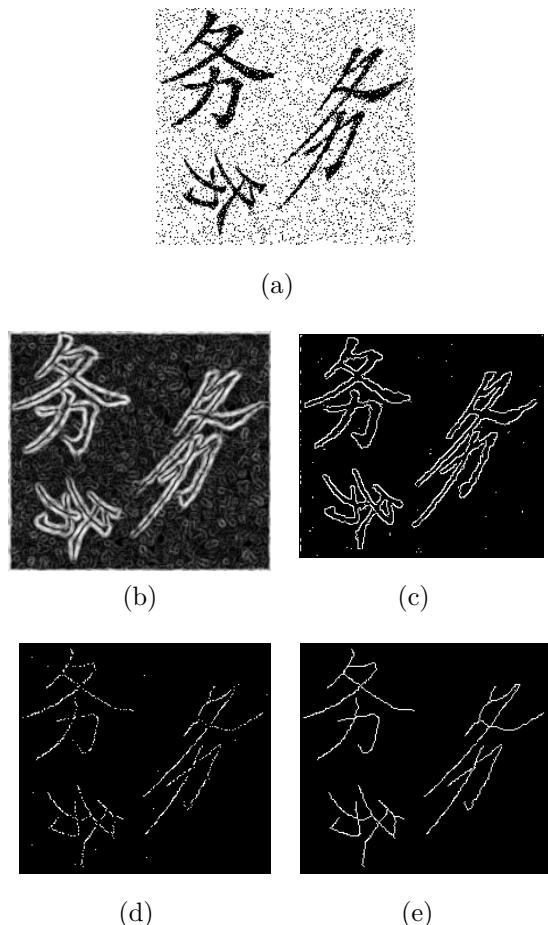


Fig. 8.13 (a) The original image which is harmed by both the affine transform and noise; (b) Raw of modulus of wavelet transform; (c) Raw output of edge detection obtained by the modulus maximum of wavelet transform; (d) The output of primary skeleton extracted by the proposed algorithm; (e) The modified skeletons.

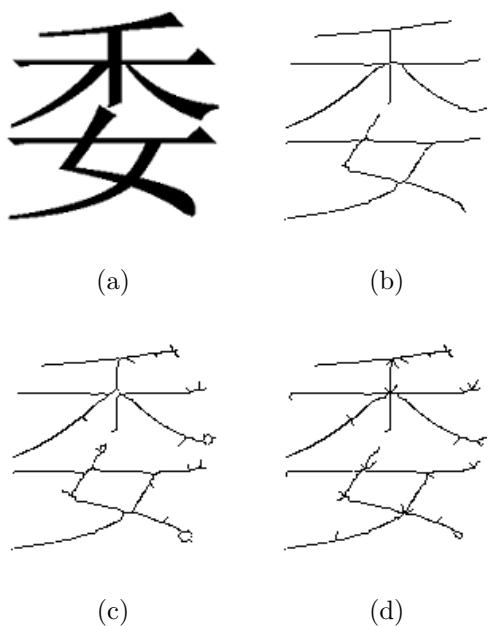


Fig. 8.14 (a) The original image of a chinese character; (b) The skeleton extracted by the proposed algorithm; (c) The skeleton obtained from ZSM; (d) The skeleton obtained from CYM.

This page intentionally left blank

Chapter 9

Feature Extraction by Wavelet Sub-Patterns and Divider Dimensions

Feature extraction is the heart of a pattern recognition system. In pattern recognition, features are utilized to identify one class of pattern from another. The pattern space is usually of high dimensionality. The objective of the feature extraction is to characterize the object, and further, to reduce the dimensionality of the measurement space to a space suitable for the application of pattern classification techniques. In the feature extraction phase, only the salient features necessary for the recognition process are retained such that the classification can be implemented on a vastly reduced feature space.

Feature extraction can be viewed as a mapping, which maps a pattern space into a feature space. Pattern space \mathbf{X} may be described by a vector of m pattern vectors such that

$$\mathbf{X} = \begin{vmatrix} X_1^T \\ X_2^T \\ \vdots \\ X_m^T \end{vmatrix} = \begin{vmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & & & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{vmatrix}, \quad (9.1)$$

where the superscript T for each vector stands for its transpose, the $X_i^T = (x_{i1}, x_{i2}, \dots, x_{in})$, $i = 1, \dots, m$ represent pattern vectors.

The objective of the feature extraction functions as the dimensionality reduction. It maps the pattern space (i.e. original data) into the feature space (i.e. feature vectors). The dimensionality of the feature space has to be smaller than that of the pattern space. Obviously, the feature vectors

can be represented by

$$X_i^T = (x_{i1}, x_{i2}, \dots, x_{ir}), i = 1, \dots, m, r < n. \quad (9.2)$$

Note that since the feature space is in a smaller dimension, the maximum value of r has to be smaller than that of n , i.e. $r < n$.

There are many methods to feature selection, however they can be categorized into the followings [Tou and Gonzalez, 1974]:

- Entropy minimization - Entropy is a statistical measure of uncertainty. It can be used as a suitable criterion in the design of optimum feature selection. The entropy minimization approach is based on the assumption that the pattern classes under consideration are normally distribution.
- Orthogonal expansion - When the assumption of the normally distribution is not valid, the method of orthogonal expansion offers an alternative approach to the feature extraction. The principal advantage of this method is that it does not require knowledge of the various probability densities.
- Functional approximation - If the features of a class of objects can be characterized by a function that is determined on the basis of observed data, the feature extraction can be considered to be a problem of functional approximation.
- Divergence - It can be used to determine feature ranking and to evaluate the effectiveness of class discrimination. The divergence can also be used as a criterion function for generating an optimum set of features.

In this chapter, we present a novel approach to extract features in pattern recognition that utilizes ring-projection-wavelet-fractal signatures (RPWFS). This approach can be categorized into the second method listed in the preceding paragraph. In particular, this approach reduces the dimensionality of a two-dimensional pattern by way of a ring-projection method, and thereafter, performs Daubechies' wavelet transform on the derived one-dimensional pattern to generate a set of wavelet transformed sub-patterns, namely, curves that are non-self-intersecting. Further from the resulting non-self-intersecting curves, the divider dimensions are readily computed. These divider dimensions constitute a new feature vector for the original two-dimensional pattern, defined over the curves' fractal dimensions.

Wavelet analysis and its applications have become the fastest growing research areas in the recent years. Advanced research and development in wavelet analysis has found numerous applications in such areas as signal processing, image processing, and pattern recognition with many encouraging results [Daubechies, 1990; IEEE, 1992; IEEE, 1993; Mallat and Hwang, 1992; Tang et al., 1997a; Tang et al., 1998a]. During this fast growth in theories and applications, the theoretical development of high-dimensional wavelet analysis is somewhat lagging behind as compared to that of the one-dimensional wavelet. As has been shown in several real-life applications, the two-dimensional wavelet analysis has not been as effectively applied as the one-dimensional analysis.

The goal of this chapter is to, through mathematically sound derivations, reduce the problem of two-dimensional patterns into that of one-dimensional ones, and thereafter, utilize the well-established 1-D wavelet transform coupled with fractal theory to extract the one-dimensional patterns' feature vectors for the purpose of pattern recognition.

It is a well-known fact that in many real-life pattern recognition situations such as optical character recognition (OCR), patterns are often found to be rotated due to experimentation constraints or errors. This implies that the new pattern recognition method must be developed that is invariant to rotations. In 1991, Tang [Tang et al., 1991] first proposed a method of transforming two-dimensional patterns into one-dimensional patterns through so-called *ring projections*. As the projections are done in the form of rings, the one-dimensional pattern obtained from ring-projection is invariant to rotations.

Drawing on the aforementioned work by Tang [Tang et al., 1991], the present work further explores the use of ring-projection in reducing the dimensionality of two-dimensional patterns such as alpha-numeric symbols into one-dimensional ones. Consequently, we can perform Daubechies' wavelet transform on the derived one-dimensional patterns to generate a set of *wavelet transformed sub-patterns* which are curves with non-self-intersecting. We then compute the *divider dimensions* for the individual curves. This allows us to map the set of non-self-intersecting curves into a feature vector defined over the *curves' fractal dimensions*, which is also the feature vector corresponding to the original two-dimensional patterns. An overall description of this approach can be illustrated by a diagram shown in Fig. 9.1.

Following the descriptions of the theoretical constructs of the ring-

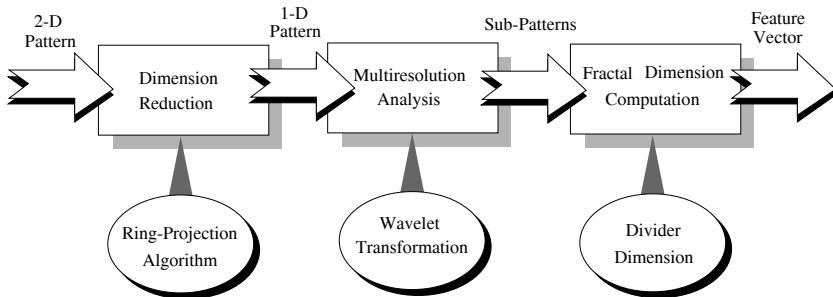


Fig. 9.1 Diagram of Ring-projection-wavelet-fractal method.

projection-wavelet-fractal signatures (RPWFS) method, we shall also present the results of experiments in which two-dimensional patterns were tested. The experiment data consist of a sub-set of Chinese characters and a set of printed alphanumeric symbols, including 26 alphabets in both upper and lower cases, 10 numeric digits, and additional 10 ASCII symbols. Our method yielded 100 % correct classification rate, even when the orientations or fonts of these characters were altered.

The overall sequence of presentation of this chapter is illustrated below:

- Dimensionality reduction of two-Dimensional Patterns with a Ring-Projection
- Wavelet orthonormal decomposition to produce sub-patterns
- Wavelet-fractal scheme
 - Basic concepts of fractal dimension
 - The divider dimension of one-dimensional patterns
- Experiments
 - Experimental procedure
 - Experimental results

9.1 Dimensionality Reduction of Two-Dimensional Patterns with Ring-Projection

This section provides an overview of the ring-projection method for reducing the dimensionality of two-dimensional patterns. First, suppose that a two-dimensional pattern such as an alphanumeric symbol has been repre-

sented into a binary image. Taking a letter A as an example, its grey-scale image, $p(x, y)$, can be discretized into binary values as follows:

$$p(x, y) = \begin{cases} 1 & \text{if } (x, y) \in \mathcal{D} \\ 0 & \text{otherwise} \end{cases} \quad (9.3)$$

where domain \mathcal{D} corresponds to the white region of letter A , as shown in Fig. 9.2. The above multivariate function $p(x, y)$ can also be viewed as a two-dimensional density function of mass distribution on the plane. From Eq. (9.3), it is readily noted that the corresponding density function is a uniform distribution. That is also to say, the mass is homogeneously distributed over the region \mathcal{D} . From this uniform mass distribution, we can derive the centroid of the mass for the region \mathcal{D} , as denoted by $m(x_0, y_0)$, and subsequently, translate the origin of our reference frame to this centroid, as has been illustrated in Fig. 9.2.

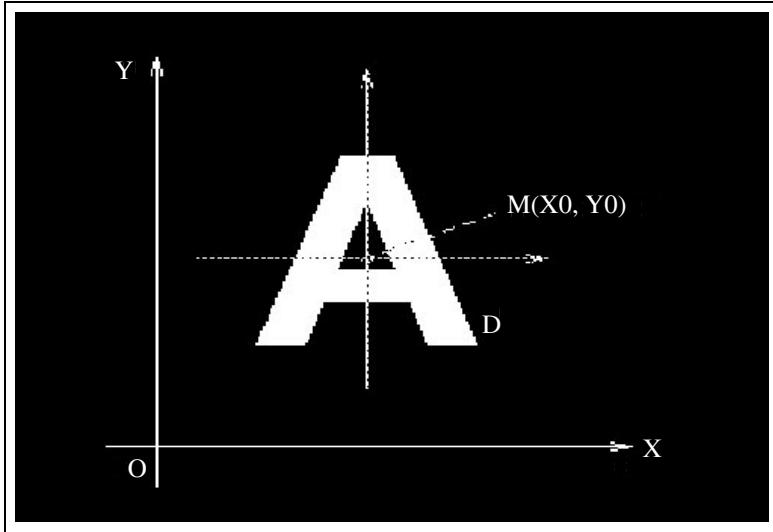


Fig. 9.2 A binary-image of letter “A” whose centroid is used to place the origin of a new reference frame.

Next, we let

$$M = \max_{N \in \mathcal{D}} | N(x, y) - m(x_0, y_0) |$$

where $| N(x, y) - m(x_0, y_0) |$ represents the Euclidean distance between two

points, N and m , on the plane. Further, we transform the original reference Cartesian frame into a polar frame based on the following relations:

$$\begin{cases} x = \gamma \cos \theta \\ y = \gamma \sin \theta \end{cases} \quad (9.4)$$

Hence,

$$p(x, y) = p(\gamma \cos \theta, \gamma \sin \theta)$$

where $\gamma \in [0, \infty)$, $\theta \in (0, 2\pi]$.

For any fixed $\gamma \in [0, M]$, we then compute the following integral:

$$f(\gamma) = \int_0^{2\pi} p(\gamma \cos \theta, \gamma \sin \theta) d\theta \quad (9.5)$$

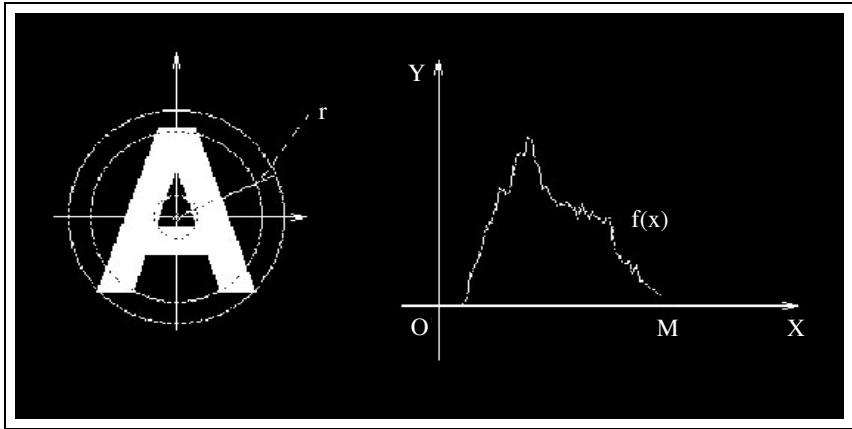


Fig. 9.3 An illustration of the ring-projection for letter *A*.

The resulting $f(\gamma)$ is in fact equal to the total mass as distributed along circular rings, as shown in Fig. 9.3. Hence, the derivation of $f(\gamma)$ is also termed as a ring-projection of the planar mass distribution. The single-variate function $f(\gamma)$, $\gamma \in [0, M]$, sometimes also denoted as $f(x)$, $x \in [0, M]$, can be viewed as an one-dimensional pattern that is directly transformed from the original two-dimensional pattern through a ring-projection. Owing to the facts that the centroid of the mass distribution is

invariant to rotation and that the projection is done along circular rings, the derived one-dimensional pattern will be invariant to the rotations of its original two-dimensional pattern. In other words, the ring-projection is rotation-invariant.

From a practical point of view, the images to be analyzed by a recognition system are most often stored in discrete formats. Catering to such discretized two-dimensional patterns, we shall modify Eq. (9.5) into the following expression (see Fig. 9.4):

$$f(\gamma) = \sum_{k=0}^M p(\gamma \cos \theta_k, \gamma \sin \theta_k) \quad (9.6)$$

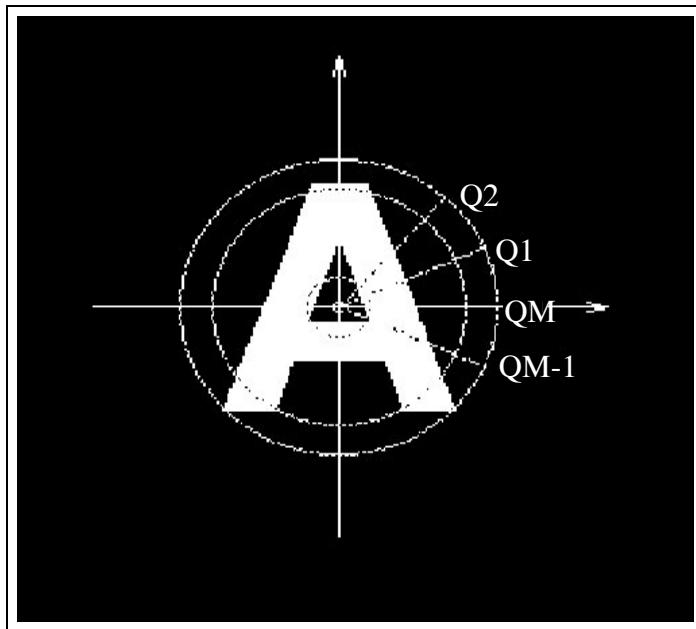


Fig. 9.4 Projection of a two-dimensional pattern along rings in a discrete manner.

Two examples of dimension reduction for two-dimensional patterns can be found in Figs. 9.5 and 9.6. In the first example, the image of Chinese character “big” is a two-dimensional pattern as shown on the left side of Fig. 9.5. After applying the operation of the ring-projection, an one-dimensional signal is obtained and illustrated on the right side of Fig. 9.5.

In the second example, the original pattern is a two-dimensional object, an aircraft, which is displayed on the left side of Fig. 9.6. As the ring-projection operation is applied to it, an one-dimensional signal is obtained, and presented on the right side in Fig. 9.6.

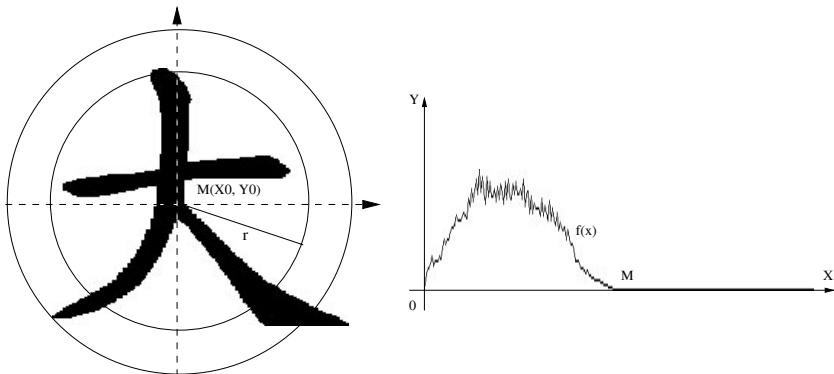


Fig. 9.5 An example of dimension reduction for Chinese character “big”.

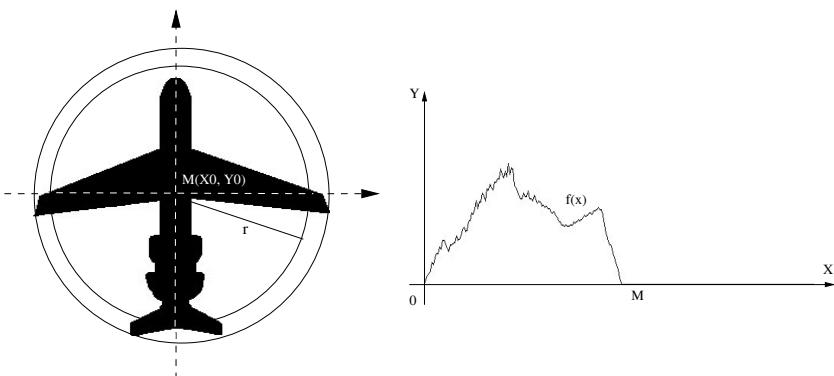


Fig. 9.6 An example of dimension reduction for a two-dimensional pattern.

Interested readers are referred to [Tang et al., 1991] for a thorough discussion on ring-projection. In the succeeding sections, we are concerned mainly with how to extract as most information as possible from the obtained ring-projection, i.e., an one-dimensional pattern, by way of wavelet transform. This, as will be described later, enables us to obtain a set of

wavelet transformed sub-patterns – curves that are non-self-intersecting, from which feature vectors defined over the curves' fractal dimensions can easily be computed.

9.2 Wavelet Orthonormal Decomposition to Produce Sub-Patterns

Let $\{V_j\}$ be an orthonormal MRA, then V_{j-1} can be decomposed orthogonally as follows:

$$V_{j-1} = W_j \oplus V_j. \quad (9.7)$$

Note that the order of V_j and V_{j-1} here, in this chapter, differs from that in the previous chapters, namely,

- we use the subspaces in $\{V_j\}$ by the order of

$$\cdots \subset V_{j+1} \subset V_j \subset V_{j-1}$$

in the present chapter;

- we use the order of

$$\cdots \subset V_{j-1} \subset V_j \subset V_{j+1}$$

in the previous chapters.

Actually, these two representations have no different.

As can be realized, any real-world instruments for measuring physical data are capable of acquiring information with limited precision. In other words, the signals obtained using such instruments will have limited resolution. The same is also true for the one-dimensional pattern that we have derived in the preceding sections, $f(x)$; it must belong to a closed space V_j . Mathematically, we can express this as follows:

$$P_j : L^2(\mathbb{R}) \implies V_j \quad \text{projective operator from } L^2(\mathbb{R}) \text{ to } V_j$$

$$Q_j : L^2(\mathbb{R}) \implies W_j \quad \text{projective operator from } L^2(\mathbb{R}) \text{ to } W_j$$

Since $f(x) \in V_j \subset L^2(\mathbb{R})$, we arrive at

$$\begin{aligned} f(x) &= P_j f(x) = \sum_{k \in \mathbb{Z}} c_{j,k} \varphi_{j,k}(x) \\ &= P_{j+1} f(x) + Q_{j+1} f(x) \end{aligned}$$

$$= \sum_{m \in \mathbb{Z}} c_{j+1,m} \varphi_{j+1,m}(x) + \sum_{m \in \mathbb{Z}} d_{j+1,m} \psi_{j+1,m}(x) \quad (9.8)$$

where

$$c_{j+1,m} = \langle P_{j+1}f, \varphi_{j+1,m} \rangle = \sum_{k \in \mathbb{Z}} c_{j,k} \hat{h}_{k-2m} \quad (9.9)$$

$$d_{j+1,m} = \langle Q_{j+1}f, \psi_{j+1,m} \rangle = \sum_{k \in \mathbb{Z}} c_{j,k} \hat{g}_{k-2m} \quad (9.10)$$

and \hat{h}_{k-2m} and \hat{g}_{k-2m} denote the conjugate vectors of h_{k-2m} and g_{k-2m} , respectively. When both h_{k-2m} and g_{k-2m} are real, we have $\hat{h}_{k-2m} = h_{k-2m}$ and $\hat{g}_{k-2m} = g_{k-2m}$.

It should be mentioned that both Eqs. (9.9) and (9.10) compute the sum of an infinite number of terms. However, from a computational point of view, we wish to see only a finite number of non-zero terms, h_k , hence reducing the problem of infinite summation to that of finite summation. In doing so, we carry out the following procedure, in order to find the expression

$$m_0(\omega) = \sum_{k \in \mathbb{Z}} \frac{1}{\sqrt{2}} h_k e^{-ik\omega},$$

such that the number of non-zero h_k is finite.

First, let us consider:

$$\begin{aligned} 1 &= \left(\cos^2 \frac{\omega}{2} + \sin^2 \frac{\omega}{2} \right)^3 \\ &= \cos^6 \frac{\omega}{2} + 3 \cos^4 \frac{\omega}{2} \sin^2 \frac{\omega}{2} + 3 \cos^2 \frac{\omega}{2} \sin^4 \frac{\omega}{2} + \sin^6 \frac{\omega}{2} \end{aligned} \quad (9.11)$$

Suppose that

$$|m_0(\omega)|^2 = \cos^6 \frac{\omega}{2} + 3 \cos^4 \frac{\omega}{2} \sin^2 \frac{\omega}{2}, \quad (9.12)$$

hence,

$$|m_0(\omega + \pi)|^2 = \sin^6 \frac{\omega}{2} + 3 \sin^4 \frac{\omega}{2} \cos^2 \frac{\omega}{2}. \quad (9.13)$$

Since,

$$\frac{1 + e^{-i\omega/2}}{2} = \frac{e^{-i\omega/2}(e^{i\omega/2} + e^{-i\omega/2})}{2} = e^{-i\omega/2} \cos \frac{\omega}{2} \quad (9.14)$$

$$\frac{1 - e^{-i\omega/2}}{2} = \frac{e^{-i\omega/2}(e^{i\omega/2} - e^{-i\omega/2})}{2} = ie^{-i\omega/2} \sin \frac{\omega}{2} \quad (9.15)$$

we have:

$$\begin{aligned} m_0(\omega) &= (e^{-i\omega/2} \cos \frac{\omega}{2})^2 e^{-i\omega/2} (\cos \frac{\omega}{2} + i\sqrt{3} \sin \frac{\omega}{2}) \\ &= \left(\frac{1 + e^{-i\omega/2}}{2} \right)^2 \left(\frac{1 + e^{-i\omega/2}}{2} + \sqrt{3} \frac{1 - e^{-i\omega/2}}{2} \right) \\ &= \frac{1}{8} \left[(1 + \sqrt{3}) + (3 + \sqrt{3})e^{-i\omega} \right] \\ &\quad + \frac{1}{8} \left[(3 - \sqrt{3})e^{-i\omega 2} + (1 - \sqrt{3})e^{-i\omega 3} \right] \end{aligned} \quad (9.16)$$

at the same time, we know:

$$m_0(\omega) = \sum_{k=0}^3 \frac{1}{\sqrt{2}} h_k e^{-ik\omega} \quad (9.17)$$

By comparing Eqs. (9.16) and (9.17), we can immediately obtain the following:

$$\begin{aligned} h_0 &= \frac{1 + \sqrt{3}}{4\sqrt{2}} = 0.4829629131445341 \\ h_1 &= \frac{3 + \sqrt{3}}{4\sqrt{2}} = 0.8365163037378077 \\ h_2 &= \frac{3 - \sqrt{3}}{4\sqrt{2}} = 0.2241438680420134 \\ h_3 &= \frac{1 - \sqrt{3}}{4\sqrt{2}} = -0.1294095225512603 \end{aligned}$$

In other words, if the above terms are chosen as h_k , $k = 0, 1, 2, 3$, and

$$m_0(\omega) = \sum_{k=0}^3 \frac{1}{\sqrt{2}} h_k e^{-ik\omega}$$

will satisfy the following conditions:

$$|m_0(\omega)|^2 + |m_0(\omega + \pi)|^2 = 1$$

$$m_0(0) = 1$$

Now, if we define that

$$\hat{\varphi}(\omega) = \prod_{j=1}^{\infty} m_0(\omega/2^j),$$

then the resulting scaling function, $\varphi(x)$, can be said to have a compact support [Daubechies, 1988]. Thus, we can simplify the wavelet transform expressions of Eqs. (9.9) and (9.10) into the following:

$$c_{j+1,m} = \sum_{k=0}^3 h_k c_{j,k+2m} \quad (9.18)$$

$$d_{j+1,m} = \sum_{k=0}^3 g_k c_{j,k+2m} \quad (9.19)$$

Next, $m_0(\omega)$ is also determined following the above mentioned steps, except that we examine the expression of

$$1 = \left(\cos^2 \frac{\omega}{2} + \sin^2 \frac{\omega}{2} \right)^5$$

to obtain

$$m_0(\omega) = \sum_{k=0}^5 \frac{1}{\sqrt{2}} h_k e^{-ik\omega}.$$

In this case, we can simplify the wavelet transform expressions of Eqs. (9.9) and (9.10) into the following:

$$c_{j+1,m} = \sum_{k=0}^5 h_k c_{j,k+2m} \quad (9.20)$$

$$d_{j+1,m} = \sum_{k=0}^5 g_k c_{j,k+2m} \quad (9.21)$$

In [Daubechies, 1992], I. Daubechies used the following notations:

$$nM_0(\xi) = \frac{1}{\sqrt{2}} \sum_{n=0}^{2N-1} N h_n e^{-in\xi} \quad (9.22)$$

where $nM_0(\xi)$ is equivalent to $m_0(\omega)$ in our case, and Nh_n is equivalent to our h_k . Daubechies also provided the exact values of h_k (or Nh_n) for $N = 2$ to 10 in tabular forms. The tables might be of interest to those readers who want to use Eqs. (9.9) and (9.10) but not to know how h_k 's are computed. When N is large enough, the computation of h_k could become problem-dependent. Generally speaking, the larger the N value, the higher the resolution of the wavelet orthonormal decomposition will be, and at the same time, the more costly the computation will become.

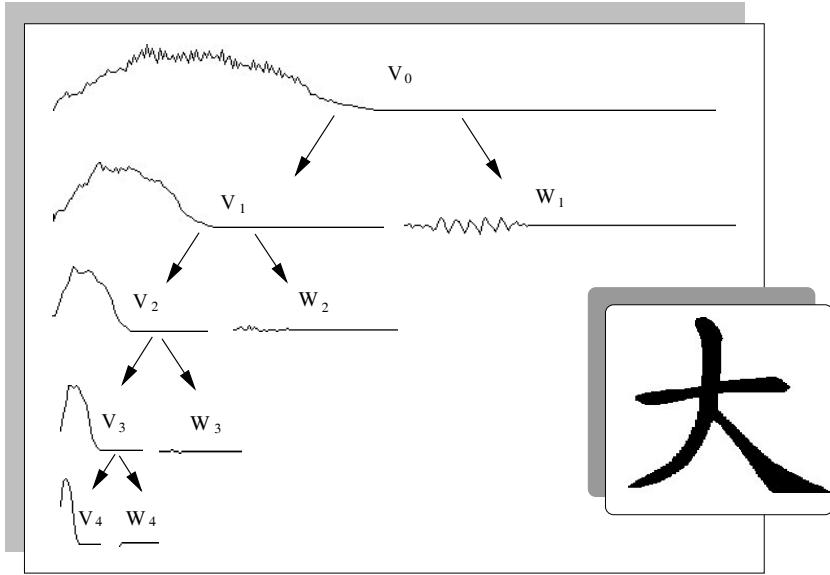


Fig. 9.7 Examples of the wavelet transform sub-patterns from a Chinese character.

According to the wavelet orthonormal decomposition as shown in Eq. (9.7), first V_j is decomposed orthogonally into a high-frequency sub-space W_{j+1} and a low-frequency one V_{j+1} using the wavelet transform Eqs. (9.20) and (9.21). The low-frequency sub-space V_{j+1} is further decomposed into W_{j+2} and V_{j+2} . Afterwards, V_{j+2} is broken into W_{j+3} and V_{j+3} , and these processes can be continued. The above wavelet orthonormal decomposition

can be represented by

$$\begin{aligned}
 V_j &= W_{j+1} \oplus V_{j+1} \\
 &\quad \Downarrow \\
 &(W_{j+2} \oplus V_{j+2}) \\
 &\quad \Downarrow \\
 &(W_{j+3} \oplus V_{j+3}) \\
 &\quad \Downarrow \\
 &\dots\dots
 \end{aligned}$$

In the view of pattern recognition, V_j can be considered to be an original pattern, while V_{j+1} , V_{j+2} , ..., W_{j+1} , W_{j+2} , ... can be regard as the sub-patterns which are referred as *wavelet transform sub-patterns*.

In Fig. 9.7, we have shown the one-dimensional pattern resulted from the ring-projection [Tang et al., 1991] of a two-dimensional pattern, for example, a Chinese character, and its corresponding wavelet transform sub-patterns. In this figure, V_0 denotes the resulting one-dimensional pattern from the ring-projection, which is a dimension reduction operation. V_j denotes the wavelet transform sub-pattern resulted from the j th wavelet transform based on Eq. (9.20). W_j denotes the wavelet transform sub-pattern resulted from the j th wavelet transform based on Eq. (9.21). Since

$$\begin{aligned}
 V_j &= W_{j+1} \oplus V_{j+1} \\
 &= W_{j+1} \oplus W_{j+2} \oplus V_{j+2} \\
 &= W_{j+1} \oplus W_{j+2} \oplus W_{j+3} \oplus V_{j+3} \\
 &= \dots\dots,
 \end{aligned}$$

it can be seen that in Fig. 9.7,

$$\begin{aligned}
 V_0 &= W_1 \oplus V_1 \\
 &= W_1 \oplus W_2 \oplus V_2 \\
 &= W_1 \oplus W_2 \oplus W_3 \oplus V_3 \\
 &= W_1 \oplus W_2 \oplus W_3 \oplus W_4 \oplus V_4.
 \end{aligned}$$

Another example of the wavelet transform sub-patterns from a two-dimensional object using the wavelet orthonormal decomposition is illustrated in Fig. 9.8. The original pattern is an aircraft, which is a 2-D object. The dimension reduction operation is applied to it, and an 1-D pattern is produced which is a curve labeled by V_0 in Fig. 9.8. After utilizing the

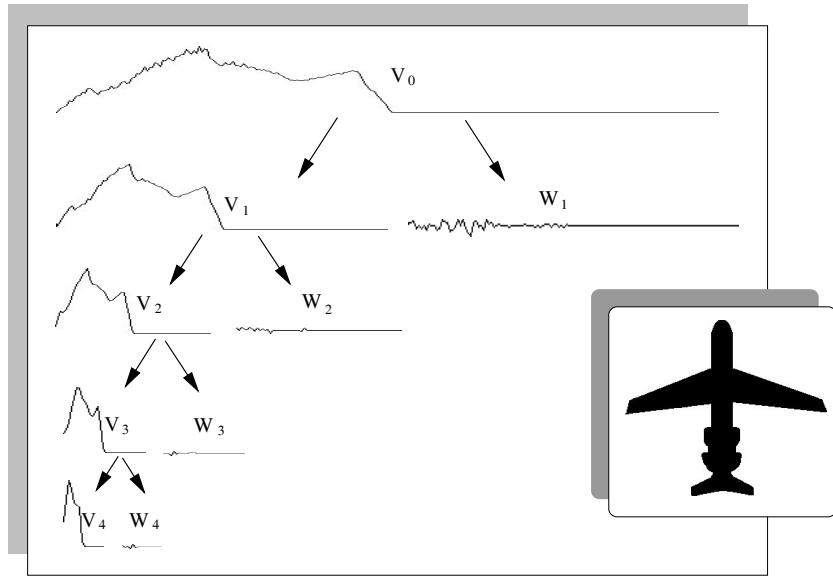


Fig. 9.8 Examples of the wavelet transform sub-patterns from a pattern.

wavelet orthonormal decomposition, eight wavelet transform sub-patterns are obtained, i.e. $V_1 - V_4$ and $W_1 - W_4$ in Fig. 9.8.

In Fig. 9.9, we have shown the one-dimensional pattern resulted from the ring-projection of an alphabet, A , and its corresponding wavelet transformation sub-patterns. In the figure, $S_{20}A$ denotes the resulting one-dimensional pattern. $S_{2j}A$ denotes the wavelet sub-pattern resulted from the j th wavelet transformation of $S_{20}A$ based on Eq. (9.20). $W_{2j}A$ denotes the wavelet sub-pattern resulted from the j th wavelet transformation of $S_{20}A$ based on Eq. (9.21).

The above wavelet orthonormal decomposition can be represented by

$$\begin{aligned}
 S_{20}A = & \quad W_{21}A \quad \oplus \quad S_{21}A \\
 & \quad \Downarrow \\
 & (W_{22}A \quad \oplus \quad V_{22}A) \\
 & \quad \Downarrow \\
 & (W_{23}A \quad \oplus \quad V_{23}A) \\
 & \quad \Downarrow \\
 & (W_{24}A \quad \oplus \quad V_{24}A)
 \end{aligned}$$

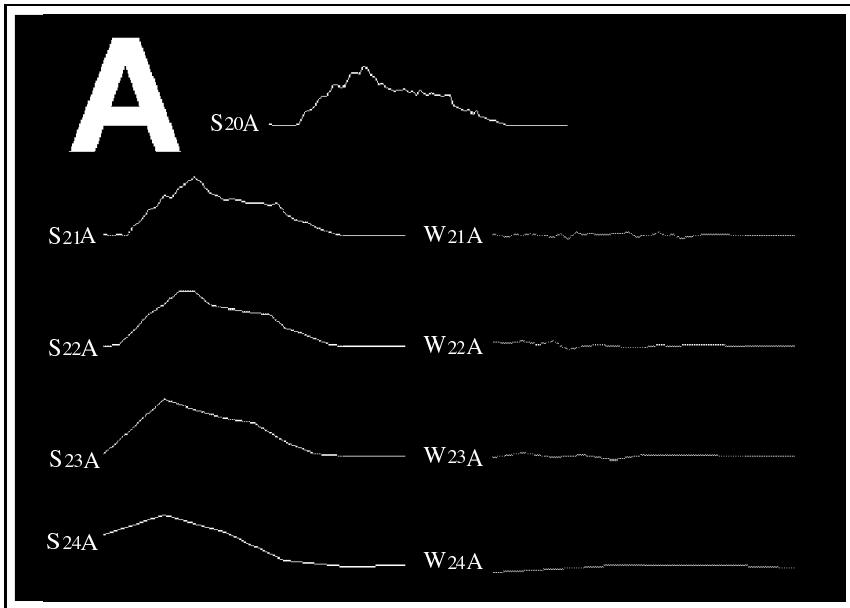


Fig. 9.9 The one-dimensional pattern obtained after ring-projection, and its corresponding wavelet transformation sub-patterns.

In the view of pattern recognition, $S_{20}A$ can be considered to be an original pattern, while $V_{2j}A$ and $W_{2j}A$ can be regard as the sub-patterns.

9.3 Wavelet-Fractal Scheme

In 1975, B. Mandelbrot first developed the notion of *fractal* [Mandelbrot, 1982]. For the past decades, the theories of fractal geometry have been widely applied in a variety of domains, such as in computer science, physics, chemistry, biology, material science, geography, geology, and even in social science and humanities. Of equal significance is its applications in pattern recognition. Among others, in 1994, Tang et al. examined the essence of Minkowski fractional dimension, and applied the Minkowski blanket technique in deriving fractal signatures for document processing, as part of a new method for page layout analysis [Ma et al., 1995; Tang et al., 1995b; Tang et al., 1997b].

9.3.1 Basic Concepts of Fractal Dimension

In this sub-section, the basic concepts of *fractals* will be introduced.

Definition 9.1 A collection \mathfrak{R} of subsets of the space X is said to have order $m + 1$ if some point of X lies in $m + 1$ elements of \mathfrak{R} , and no point of X lies in more than $m + 1$ elements of \mathfrak{R} . Given a collection \mathfrak{R} of subsets of X , a collection \mathfrak{S} is said to *refine* \mathfrak{R} , or to be a *refinement* of \mathfrak{R} , if each element B of \mathfrak{S} is contained in at least one element of \mathfrak{R} .

Now we define what we mean by the *topological dimension* of a space X .

Definition 9.2 A space X is said to be *finite-dimensional* if there is some integer m such that for every open covering \mathfrak{R} of X , there is an open covering \mathfrak{S} of X that refines \mathfrak{R} and has order at most $m + 1$. The *topological dimension* of X is defined to be the smallest value of m for which this statement holds. We use notation $\dim_T X$ to represent the topological dimension of X .

The topological dimension is a complicated and advanced mathematical topic, more details about it can be found in [Munkres, 1975].

What are fractals? There are many definitions, because it is very difficult to define fractal strictly. B. Mandelbrot gave two definitions in 1982 and 1986 respectively.

(1) The first definition from his original essay (1982) says:

Definition 9.3 A set F is called fractal set if its *Hausdorff dimension* ($\dim_H F$) greater than the *Topological dimension* ($\dim_T F$), namely:

$$\dim_H F > \dim_T F$$

(2) In 1986, B. Mandelbrot defined the fractal as:

Definition 9.4 Fractal is a compound object, which contains several sub-objects. The global characteristic of this object is similar to the local characteristics of each sub-object.

(3) A more precise definition of the fractal set F can be provided below:

Definition 9.5 A set F is called fractal set if the following conditions are satisfied:

- The global characteristic of the set F is self-similar to the local characteristics of each sub-set, namely:

$$\mathfrak{S}(F) \sim \mathfrak{S}(f_i), \quad f_i \supset F,$$

where $\mathfrak{S}(\cdot)$ stands for the characteristic of (\cdot) .

- The set F is infinitely separable, i.e.

$$F = \{f_1^1, f_2^1, \dots, f_i^1, \dots, f_n^1\},$$

$$f_i^1 = \{f_1^2, f_2^2, \dots, f_k^2, \dots, f_n^2\},$$

.....

$$f_k^m = \{f_1^{m+1}, f_2^{m+1}, \dots, f_k^{m+1}, \dots, f_n^{m+1}\}, \quad m+1 \rightarrow \infty.$$

- Usually, the fractal dimension of the set F is a fraction, and greater than the *Topological dimension* $\dim_T F$, namely:

$$\dim_H F > \dim_T F$$

- In many cases the definition of F is recursive.

Let U be a non-empty subset of n-dimensional Euclid space \mathbb{R}^n , and the diameter of U is defined as

$$|U| = \sup\{|x - y| : x, y \in U\},$$

where $\sup\{\cdot\}$ stands for the supremum of $\{\cdot\}$. Thus, the diameter of U is the greatest distance apart of any pair of points in U . If $\{U_i\}$ is a countable collection of sets of diameter at most δ that cover F , such that

$$\mathfrak{R}(\delta) = \{U_i\} = \{U_i : i = 1, 2, \dots\},$$

and

$$F \subset \bigcup_{i=1}^{\infty} U_i, \quad 0 < |U_i| \leq \delta,$$

we say that $\{U_i\}$ is a δ -cover of F .

Suppose that $F \subset I\!\!R^n$ and s is a real number and $s \geq 0$. For any $\delta > 0$, we define

$$H_\delta^s(F) = \inf_{\Re(\delta)} \left\{ \sum_{i=1}^{\infty} |U_i|^s : \{U_i\} \text{ is a } \delta\text{-cover of } F \right\}, \quad (9.23)$$

where the symbol $\inf\{\cdot\}$ indicates the infimum of $\{\cdot\}$. As δ decreases, the class of permissible covers of F in Eq. (9.23) is reduced. Consequently, the infimum $H_\delta^s(F)$ increases, and so approaches a limit as $\delta \rightarrow 0$. We have the following definition:

Definition 9.6 When $\delta \rightarrow 0$, the limit of $H_\delta^s(F)$ exists for any subset F of $I\!\!R^n$, and the limiting value can be (and usually) 0 or ∞ . The s -dimensional *Hausdorff measure* of F can be defined by:

$$\begin{aligned} H^s(F) &= \lim_{\delta \rightarrow 0} H_\delta^s(F) \\ &= \lim_{\delta \rightarrow 0} \left[\inf_{\Re(\delta)} \left\{ \sum_{i=1}^{\infty} |U_i|^s : \{U_i\} \text{ is a } \delta\text{-cover of } F \right\} \right]. \end{aligned} \quad (9.24)$$

We can clearly prove that H^s is a measure. Specifically, $H^s(\phi) = 0$, and if $E \subset F$ then $H^s(E) \leq H^s(F)$. If $\{F_i\}$ is any countable collection of *Borel set*, such that

$$\bigcap_{i=1}^{\infty} F_i = \phi,$$

we have

$$H^s \left(\bigcup_{i=1}^{\infty} F_i \right) = \sum_{i=1}^{\infty} H^s(F_i).$$

Furthermore, if F is a Borel subset of $I\!\!R^n$, then the n-dimensional *Hausdorff measure* of F can be deduced as:

$$H^n(F) = \frac{\pi^{\frac{n}{2}}}{2^n \Gamma(\frac{n+2}{2})} vol^n(F),$$

where

- $H^n(F)$ stands for the n-dimensional *Hausdorff measure* of F .
- $vol^n(F)$ represents the n-dimensional volume of F which is called *Lebesgue measure* of F .

- vol^1 is length,
- vol^2 is area,
- vol^3 is the usual 3-dimensional volume.

Consequently, Hausdorff measures generalize the familiar ideas of length, area and volume.

Let us review Eq. (9.23). For any set F and $\delta < 1$, $H_\delta^s(F)$ is a non-increasing function of s . According to Eq. (9.24), it can be shown that $H^s(F)$ is also a non-increasing function of s . In fact, the stronger conclusion is that if $t > 0$ and $\{U_i\}$ is a δ -cover of F , we have

$$H_\delta^t(F) \leq \sum_i |U_i|^t \leq \delta^{t-s} \sum_i |U_i|^s. \quad (9.25)$$

We take the infimum, that is

$$H_\delta^t(F) \leq \delta^{t-s} H_\delta^s(F).$$

Definition 9.7 Let $\delta \rightarrow 0$, if $H^s(F) < \infty$, then $H^t(F) = 0$ for $s < t$. Therefore, there exists a critical value of s , such that $H^s(F)$ jumps from ∞ to 0 at this point. This critical value is called *Hausdorff Dimension* of F , and it is symbolized by $\dim_H F$.

Formally, we have

$$\dim_H F = \inf \{s : H^s(F) = 0\} = \sup \{s : H^s(F) = \infty\}, \quad (9.26)$$

and

$$H^s(F) = \begin{cases} \infty & \text{if } s < \dim_H F \\ 0 & \text{if } s > \dim_H F \end{cases}$$

If $s = \dim_H F$, probably $H^s(F)$ is 0 or ∞ , or may satisfy

$$0 < H^s(F) < \infty.$$

A Borel set is called an *s-set* if the latter condition as shown in the above is satisfied. More details about Hausdorff dimension can be found in [Falconer, 1990; Falconer, 1985]. Hausdorff dimension is the oldest and probably the most important one of the fractal dimensions. It has the following advantages: (1) Hausdorff dimension can be defined for any set. (2) It is mathematically convenient. (3) It is based on measures, which are relatively easy to manipulate. However, the major disadvantage of the Hausdorff dimension is that it is difficult to compute or to estimate in many cases. In

practice, *box-computing dimension* is convenient to apply. Therefore, our study will focus on the box-computing dimension.

Fundamental to most definitions of dimension is the idea of measurement at scale δ . For each δ , a set can be measured in a way that ignores irregularities of size less than δ , and we see how these measurements behave as $\delta \rightarrow 0$ [Falconer, 1990].

Suppose F is a plane curve, the measurement $M_\delta(F)$ denotes the number of sets (with length δ) which divide the set F . A dimension of F is determined by the power law obeyed by $M_\delta(F)$ as $\delta \rightarrow 0$. If

$$M_\delta(F) \sim \mathcal{K}\delta^{-s}, \quad (9.27)$$

where \mathcal{K} and s are constants, we might say that F has dimension s , and \mathcal{K} can be considered as “ s -dimensional length” of F .

Taking the logarithm of both sides in Eq. (9.27) yields the formula:

$$\log_2 M_\delta(F) \simeq \log_2 \mathcal{K} - s \log_2 \delta,$$

in the sense that the difference of the two sides tends to 0 with δ , we have

$$s = \lim_{\delta \rightarrow 0} \frac{\log_2 M_\delta(F)}{-\log_2 \delta}. \quad (9.28)$$

From the above equation, s can be regarded as a slope on a log-log scale [Falconer, 1990].

Box-computing dimension or box dimension is one of the most widely used fractal dimensions. The popularity of the Box-computing dimension is largely due to its relative ease of mathematical calculation and empirical estimation.

Let F be a non-empty and bounded subset of \mathbb{R}^n , $\xi = \{\omega_i : i = 1, 2, 3, \dots\}$ be covers of the set F as shown in Fig. 9.10. $N_\delta(F)$ denotes the number of covers, such that

$$N_\delta(F) = |\xi : d_i \leq \delta|,$$

where d_i stands for the diameter of the i -th cover. This equation means that $N_\delta(F)$ is the smallest number of subsets which cover the set F , and their diameters d_i 's are not greater than δ (Fig. 9.10).

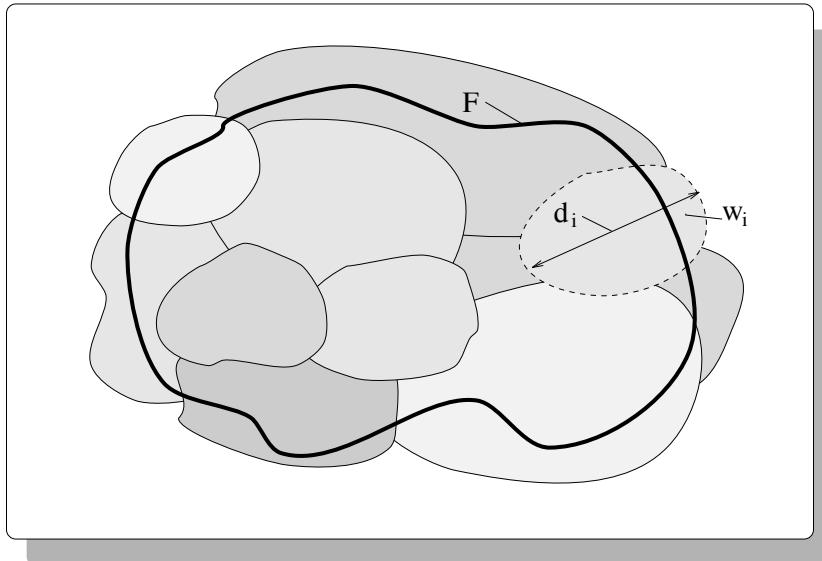


Fig. 9.10 Opening covers with diameters d_i 's covering F .

The upper and lower bounds of the box-computing dimension of F can be defined by the following formulas:

$$\underline{\dim}_B F = \liminf_{\delta \rightarrow 0} \frac{\log_2 N_\delta(F)}{-\log_2 \delta}, \quad (9.29)$$

$$\overline{\dim}_B F = \overline{\lim}_{\delta \rightarrow 0} \frac{\log_2 N_\delta(F)}{-\log_2 \delta}, \quad (9.30)$$

where the over line of $\dim_B F$ stands for the upper bound of dimension while the under line of $\dim_B F$ for lower bound. An example of the upper bound and lower bound is shown in Fig. 9.11

Definition 9.8 If both the upper bound $\overline{\dim}_B F$ and the lower bound $\underline{\dim}_B F$ are equal, i.e.

$$\liminf_{\delta \rightarrow 0} \frac{\log_2 N_\delta(F)}{-\log_2 \delta} = \overline{\lim}_{\delta \rightarrow 0} \frac{\log_2 N_\delta(F)}{-\log_2 \delta},$$

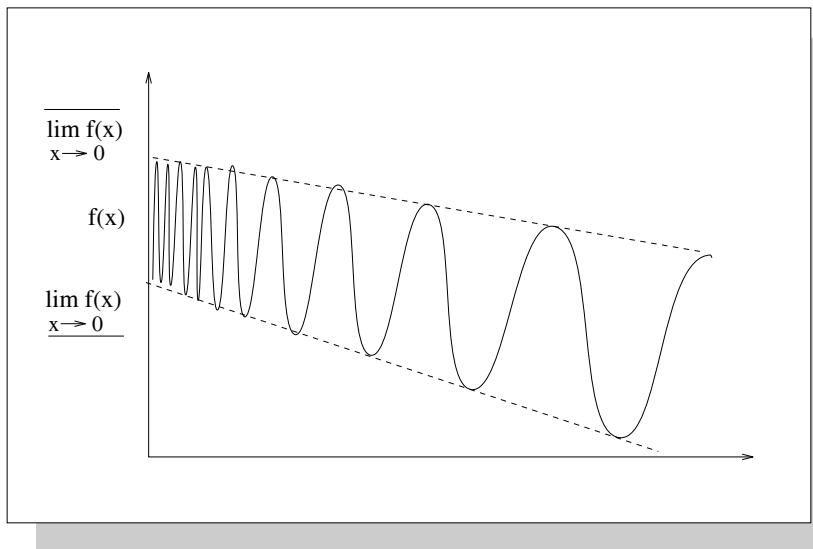


Fig. 9.11 Upper and lower bounds of a function.

the common value is called *box-computing dimension* or *box dimension* of F , namely:

$$\dim_B F = \lim_{\delta \rightarrow 0} \frac{\log_2 N_\delta(F)}{-\log_2 \delta}. \quad (9.31)$$

Further discussions on fractal theory can be found in [Edgar, 1990; Falconer, 1990].

9.3.2 The Divider Dimension of One-Dimensional Patterns

In the preceding sections, we have shown how to carry out ring-projections to reduce an original two-dimensional pattern, such as alphanumeric symbols, into an one-dimensional pattern, and furthermore described how to apply wavelet transform to the resulting one-dimensional patterns in order to obtain a set of wavelet transformed sub-patterns. Such sub-patterns are in fact non-self-intersecting curves. In this subsection, we shall address the problem of computing the divider dimension of those curves, and thereafter, use the computed divider dimension to construct a feature vector for the original two-dimensional pattern in question for pattern recognition. In

what follows, we shall first formally define the notion of divider dimension of a non-self-intersecting curve.

Definition 9.9 Suppose that C is a non-self-intersecting curve, and $\delta > 0$. Let $M_\delta(C)$ be the maximum number of ordered sequence of points x_0, x_1, \dots, x_M on curve C , such that $|x_k - x_{k-1}| = \delta$ for $k = 1, 2, \dots, M$. The divider dimension, $\dim_D C$, of curve C is defined as follows:

$$\dim_D C = \lim_{\delta \rightarrow 0} \frac{\log M_\delta(C)}{-\log \delta} \quad (9.32)$$

where $|x_k - x_{k-1}|$ represents the magnitude of the difference between two vectors x_k and x_{k-1} , as illustrated in Fig. 9.12.

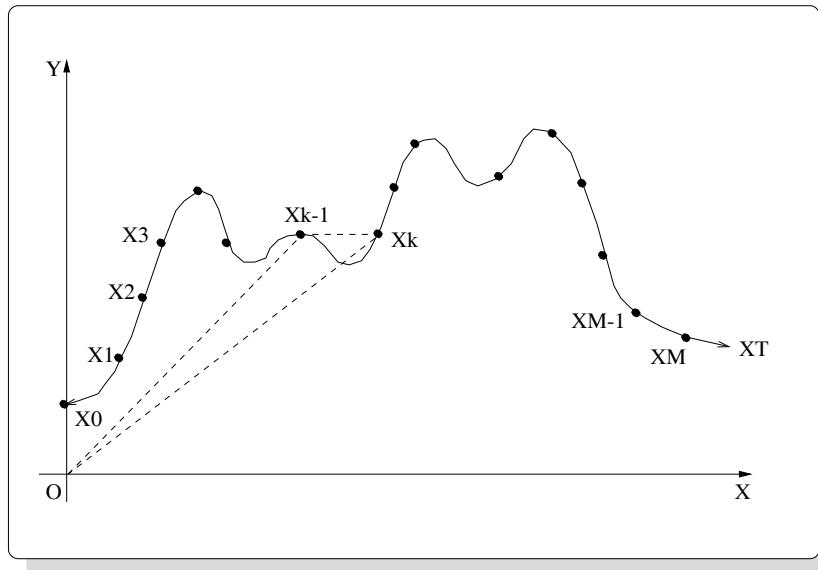


Fig. 9.12 The difference between two vectors on curve C .

It should also be mentioned that x_M is not necessarily the end point of curve C , x_T , but $|x_T - x_M| < \delta$. Furthermore, $(M_\delta(C) - 1)\delta$ may be viewed as the “length” of curve C as measured using a pair of divider that are set δ distance apart.

If the readers are interested in the details of the divider dimension, who can refer to [Falconer, 1990].

9.4 Experiments

This section presents the procedure as well as the results of our experiments that aim at recognizing a set of two-dimensional patterns.

9.4.1 Experimental Procedure

The key steps of the experimental procedure consist of the following:

Step 1. Ring-projection of two-dimensional patterns:

We denote each of the two-dimensional patterns in question by $p(x, y)$. Thus, the ring-projection of $p(x, y)$ can be expressed as follows:

$$f(x_k) = \sum_{i=0}^M p(x_k \cos \theta_i, x_k \sin \theta_i) \quad (9.33)$$

Step 2. Wavelet transform of the one-dimensional patterns:

Let $f(x_k) = c_{j,k}$, where $k = 0, 1, \dots, 2N - 1$. and

$$S_{20}A = c_{j,k} = \{c_{j,0}, c_{j,1}, \dots, c_{j,2N-1}\}.$$

Thus, the expressions for the wavelet transform of $S_{20}A$ can be written as follows:

$$\begin{aligned} c_{j+1,m} &= \sum_{k=0}^5 h_k c_{j,k+2m} \\ d_{j+1,m} &= \sum_{k=0}^5 g_k c_{j,k+2m} \end{aligned} \quad (9.34)$$

where $m = 0, 1, \dots, N - 1$.

Hence, the wavelet transformed sub-patterns of $S_{20}A$ obtained using Eq. (9.34) will become:

$$S_{21}A = \{c_{j+1,0}, c_{j+1,1}, \dots, c_{j+1,N-1}\}$$

$$W_{21}A = \{d_{j+1,0}, d_{j+1,1}, \dots, d_{j+1,N-1}\}$$

Step 3. Computation of divider dimensions for wavelet transformed sub-patterns:

Since the divider dimensions of non-self-intersecting curves are asymptotic values, we can derive their approximations based on the following expression:

$$\frac{\log M_\delta(C)}{-\log \delta}$$

when δ is set small enough.

In our experiments, we performed three consecutive wavelet transformations for each one-dimensional pattern. Hence, the 1-D pattern, such as the ring-projection of letter A , will yield seven non-self-intersecting curves, namely:

$$S_{20}A, S_{21}A, S_{22}A, S_{23}A, W_{21}A, W_{22}A, W_{23}A$$

For each of the seven curves, we further compute its divider dimension, and therefore, relate each symbol with a feature vector.

9.4.2 *Experimental Results*

This sub-section presents the results of our experiments that aim at recognizing a set of two-dimensional patterns, including a sub-set of Chinese characters, 52 upper and lower case English letters, and ten numeric digits plus additional ten ASCII symbols. All samples were rotated at different angles. Examples of Chinese character “Xin” (in English it means heart) and English letter “g” rotated at 0° , 75° , 150° , 225° , and 300° are illustrated in Fig. 9.13. The 52 upper and lower case alphabets with five different fonts were considered in the feature-vector computation.

In our experiments, we performed three consecutive wavelet transform for each one-dimensional pattern. Hence, the one-dimensional pattern, such as the ring-projection of Chinese character “Da” (in English it means big), will yield seven non-self-intersecting curves, namely,

$$V_0, V_1, V_2, V_3, W_1, W_2, W_3,$$

since

$$\begin{aligned} V_0 &= W_1 \oplus V_1 \\ &= W_1 \oplus W_2 \oplus V_2 \\ &= W_1 \oplus W_2 \oplus W_3 \oplus V_3. \end{aligned}$$

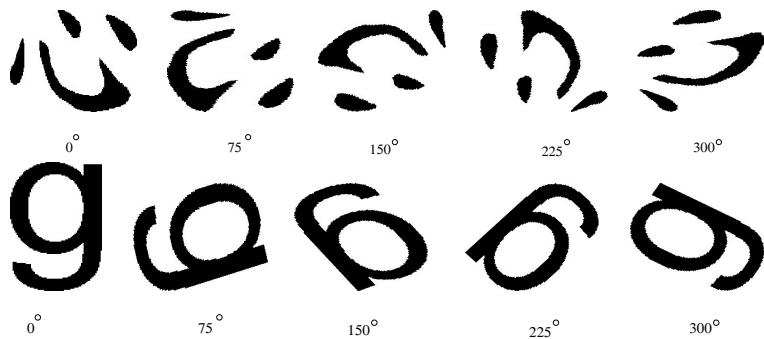


Fig. 9.13 Examples of Chinese character “Xin” (Heart) and English letter “g” rotated at 0° , 75° , 150° , 225° , and 300° .

The graphical illustration is shown in Fig. 9.14, where the curves (a), (c), (g), (i) and (d), (h), (j) represent $V_0 - V_3$ through $W_1 - W_3$ respectively. Note that $f(x)$ is the one-dimensional pattern of Chinese character “Da” (Big), and V_0 is $f(x)$ itself in this case. Fig. 9.14(b) is the frequency spectrum of V_0 , while Fig. 9.14(e) is the frequency spectrum of V_1 . By comparing these two frequency spectra, we can find that V_0 contains all frequency components of $f(x)$, while V_1 contains only one half of frequency components of $f(x)$. The reason is that V_0 is divided into two parts:

$$V_0 = V_1 \oplus W_1,$$

As for the first part, V_1 , only the low-frequency components of V_0 are kept, and the high-frequency components are lost. In addition, only the high-frequency components of V_0 are kept, and the low-frequency components are lost in W_1 . Similarly, V_1 is further divided into two parts:

$$V_1 = V_2 \oplus W_2,$$

where, the low-frequency components of V_1 remain in V_2 , while the high-frequency components of V_1 are kept in W_2 . Hence, V_2 contains only one half of frequency components of V_1 , that is only one quarter of frequency

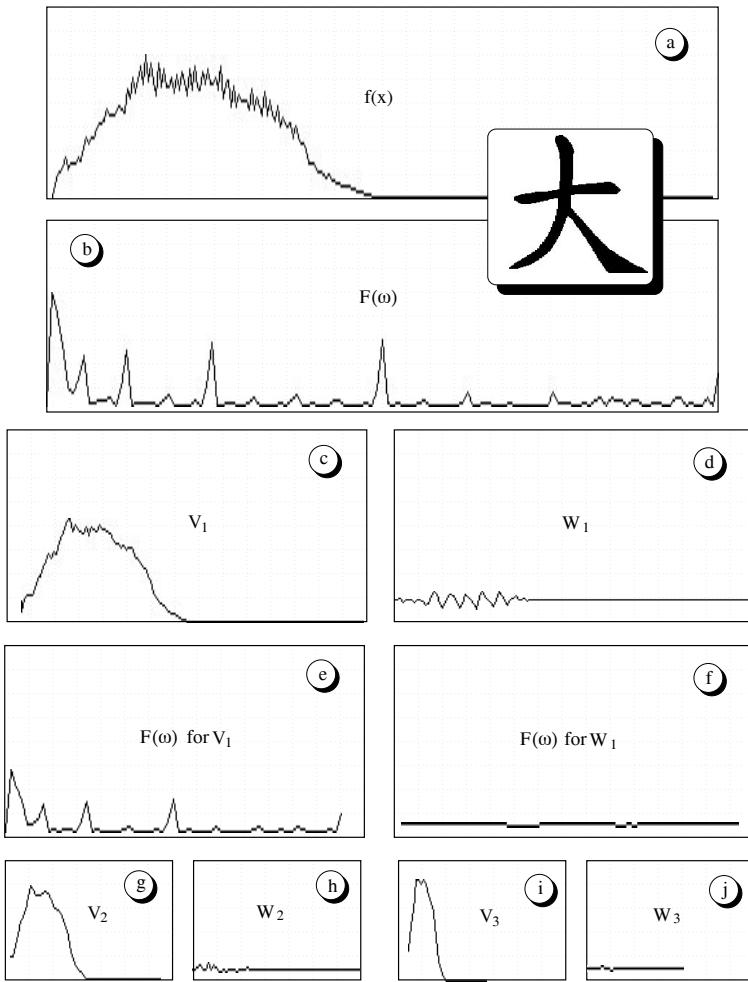


Fig. 9.14 Wavelet transform sub-patterns of Chinese character “Da” (Big) and their frequency analyses.

components of $f(x)$. Again for the same reason, V_3 contains only one eighth of frequency components of $f(x)$, etc..

A series of experiments have been conducted to verify the proposed method. The results are shown in Figs. 9.14 - 9.19.

After wavelet decomposition, a pattern has produced seven sub-patterns. For each of the seven curves, we further computed its divider Dimension,

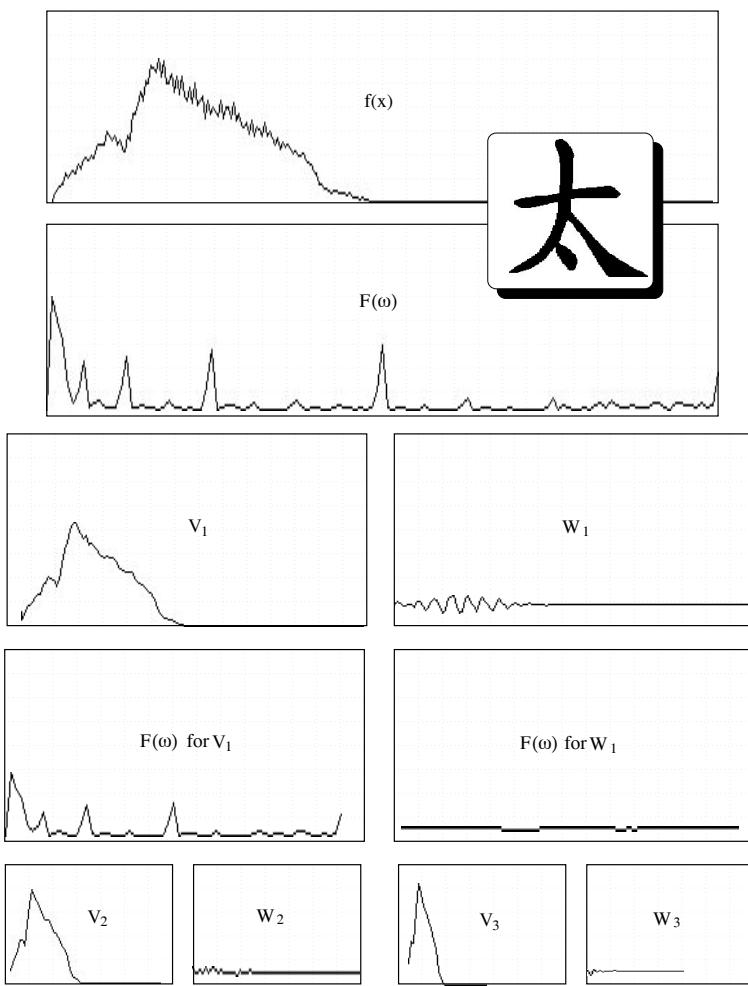


Fig. 9.15 Wavelet transform sub-patterns of Chinese character “Tai” (Too much) and their frequency analyses.

and therefore, related each pattern with a feature vector. The experiment data consist of a sub-set of Chinese characters and a set of printed alphanumeric symbols, including 26 alphabets in both upper and lower cases, 10 numeric digits, and additional 10 ASCII symbols. For the printed alphanumeric symbols, their feature vectors are presented in Tables 9.1 - 9.4, where the values in each row indicate the divider dimensions of seven sub-patterns

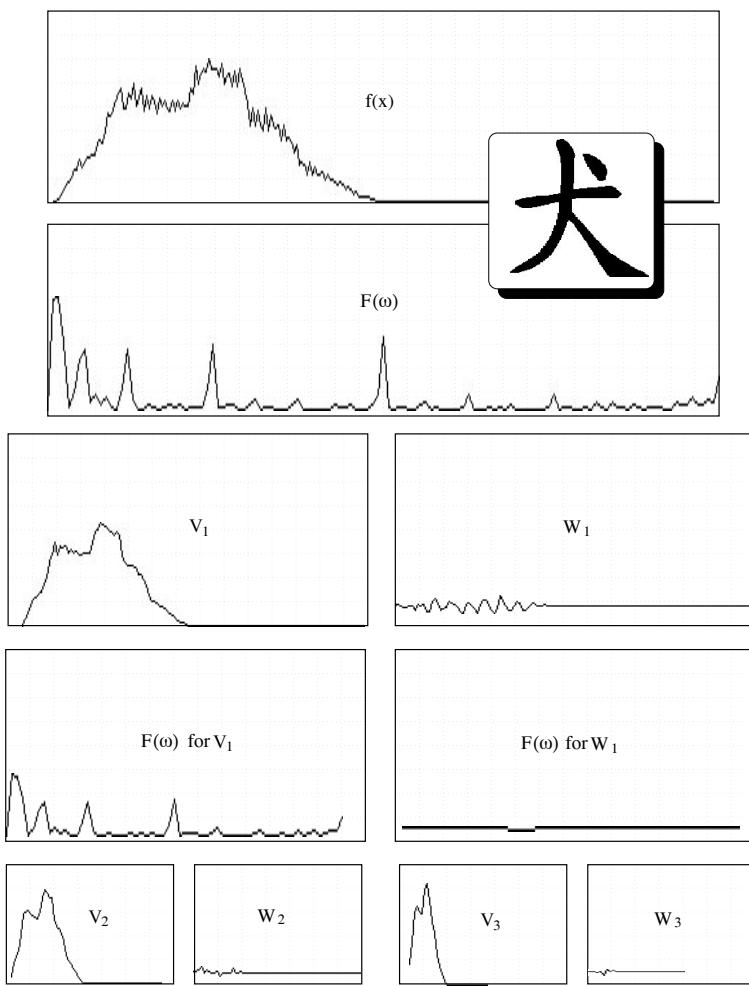


Fig. 9.16 Wavelet transform sub-patterns of Chinese character “Quan” (Dog) and their frequency analyses.

for each symbol.

char	$S_{20}(\ast)$	$S_{21}(\ast)$	$W_{21}(\ast)$
A	1.1139433384	1.1047174931	1.0648593903
B	1.1238559484	1.1136001348	1.0730702877
C	1.1167527437	1.1132092476	1.0624803305
D	1.1197948456	1.1136578321	1.0700432062
E	1.1139442921	1.1020903587	1.0629045963
F	1.1100301743	1.0874330997	1.0602608919
G	1.1257215738	1.1170243025	1.0674185753
H	1.1142077446	1.1060614586	1.0664664507
I	1.0840704441	1.0549546480	1.0441081524
J	1.1017495394	1.0808098316	1.0476298332
K	1.1139196157	1.0974267721	1.0708941221
L	1.1026504040	1.0669544935	1.0534558296
M	1.1212527752	1.1126092672	1.0775213242
N	1.1176242828	1.1083389521	1.0729970932
O	1.1214458942	1.1196649075	1.0770056248
P	1.1121571064	1.1039747000	1.0661952496
Q	1.1247720718	1.1199330091	1.0716108084
R	1.1206896305	1.1102559566	1.0806628466
S	1.1185953617	1.1158298254	1.0743304491
T	1.1036618948	1.0789839029	1.0521930456
U	1.1125218868	1.1064934731	1.0647976398
V	1.1079084873	1.0890588760	1.0515108109
W	1.1243741512	1.1173660755	1.0764572620
X	1.1142315865	1.0871822834	1.0648663044
Y	1.1017326117	1.0773953199	1.0506635904
Z	1.1157565117	1.1043380499	1.0640560389
0	1.1168053150	1.1078009605	1.0619374514
1	1.1028971672	1.0694137812	1.0519375801
2	1.1089236736	1.0893353224	1.0689998865
3	1.1099339724	1.0979336500	1.0579220057
4	1.1061652899	1.0899257660	1.0559248924
5	1.1143358946	1.0999945402	1.0730379820
6	1.1117491722	1.0993975401	1.0590016842
7	1.1040900946	1.0803054571	1.0495930910
8	1.1123523712	1.1005206108	1.0584080219
9	1.1148401499	1.1010701656	1.0597313643

Table 9.1 Divider dimensions computed for 72 printed symbols

char	$S_{22}(\ast)$	$W_{22}(\ast)$	$S_{23}(\ast)$	$W_{23}(\ast)$
A	1.0620796680	1.0203253031	1.0321298838	1.0096461773
B	1.0757583380	1.0170320272	1.0424953699	1.0116459131
C	1.0778726339	1.0138599873	1.0455238819	1.0122945309
D	1.0791871548	1.0267117023	1.0460337400	1.0130724907
E	1.0550003052	1.0216575861	1.0250078440	1.0069218874
F	1.0459311008	1.0240161419	1.0195240974	1.0061128139
G	1.0895853043	1.0255262852	1.0553485155	1.0214713812
H	1.0645886660	1.0131975412	1.0322223902	1.0091987848
I	1.0211712122	1.0151667595	1.0079869032	1.0055073500
J	1.0463556051	1.0200892687	1.0200742483	1.0079109669
K	1.0503554344	1.0194057226	1.0224356651	1.0070123672
L	1.0315659046	1.0158586502	1.0114238262	1.0061960220
M	1.0746159554	1.0180855989	1.0394244194	1.0094848871
N	1.0676525831	1.0181691647	1.0326002836	1.0105568171
O	1.0961915255	1.0249220133	1.0584391356	1.0235477686
P	1.0630793571	1.0239624977	1.0319234133	1.0084927082
Q	1.0963153839	1.0238838196	1.0569688082	1.0244978666
R	1.0684889555	1.0262739658	1.0370421410	1.0095887184
S	1.0754898787	1.0259306431	1.0420234203	1.0142017603
T	1.0403193235	1.0152760744	1.0144587755	1.0082896948
U	1.0712003708	1.0221848488	1.0397174358	1.0074933767
V	1.0530000925	1.0157818794	1.0261570215	1.0078532696
W	1.0759001970	1.0372149944	1.0386097431	1.0088177919
X	1.0430511236	1.0131111145	1.0191410780	1.0054205656
Y	1.0399550200	1.0197407007	1.0164992809	1.0062215328
Z	1.0620642900	1.0264353752	1.0310798883	1.0084336996
0	1.0728080273	1.0246663094	1.0417411327	1.0079085827
1	1.0326262712	1.0114061832	1.0116872787	1.0071448088
2	1.0527602434	1.0210211277	1.0254948139	1.0066089630
3	1.0563023090	1.0109890699	1.0268583298	1.0093988180
4	1.0551825762	1.0126544237	1.0278050900	1.0051388741
5	1.0546742678	1.0228652954	1.0254955292	1.0088582039
6	1.0559306145	1.0179107189	1.0281671286	1.0088255405
7	1.0430243015	1.0153788328	1.0195813179	1.0059862137
8	1.0582846403	1.0142199993	1.0291059017	1.0085783005
9	1.0582153797	1.0221956968	1.0298372507	1.0067170858

Table 9.2 Divider dimensions computed for 72 printed symbols (continued)

char	$S_{20}(\ast)$	$S_{21}(\ast)$	$W_{21}(\ast)$
a	1.1124289036	1.1017379761	1.0647013187
b	1.1129053831	1.1025065184	1.0577483177
c	1.1096138954	1.1018588543	1.0524585247
d	1.1138308048	1.1023379564	1.0613058805
e	1.1173223257	1.1131840944	1.0712224245
f	1.0966352224	1.0739761591	1.0401916504
g	1.1138812304	1.1035263538	1.0677573681
h	1.1104953289	1.0909165144	1.0611625910
i	1.0898797512	1.0654842854	1.0428867340
j	1.0985819101	1.0733737946	1.0505572557
k	1.1036610603	1.0771027803	1.0567890406
l	1.0866196156	1.0642234087	1.0423823595
m	1.1117404699	1.1028811932	1.0665664673
n	1.1099185944	1.1014550924	1.0498541594
o	1.1197787523	1.1124036312	1.0888708830
p	1.1132911444	1.1032136679	1.0608453751
q	1.1141122580	1.1080610752	1.0727815628
r	1.0962741375	1.0674076080	1.0423336029
s	1.1148400307	1.1055955887	1.0563476086
t	1.0966463089	1.0725744963	1.0399602652
u	1.1121745110	1.1033401489	1.0550986528
v	1.1025810242	1.0756613016	1.0482001305
w	1.1197830439	1.1085958481	1.0760765076
x	1.1038902998	1.0788321495	1.0512022972
y	1.1015509367	1.0733027458	1.0484067202
z	1.1092675924	1.0997101068	1.0636947155
+	1.0951038599	1.0682361126	1.0504192114
-	1.0777190924	1.0501787663	1.0364004374
*	1.0885618925	1.0625029802	1.0379626751
/	1.0820037127	1.0522413254	1.0374629498
=	1.1034380198	1.0821037292	1.0473917723
ł	1.0986634493	1.0686403513	1.0433781147
ı	1.1005434990	1.0753053427	1.0443359613
(1.0821937323	1.0506196022	1.0377899408
)	1.0795989037	1.0480804443	1.0329642296
?	1.1018882990	1.0856221914	1.0461981297

Table 9.3 Divider dimensions computed for 72 printed symbols (continued)

char	$S_{22}(\ast)$	$W_{22}(\ast)$	$S_{23}(\ast)$	$W_{23}(\ast)$
a	1.0639376640	1.0190188885	1.0330584049	1.0155755281
b	1.0619244576	1.0232038498	1.0326099396	1.0063673258
c	1.0650064945	1.0149598122	1.0355489254	1.0124291182
d	1.0654503107	1.0228716135	1.0351454020	1.0064482689
e	1.0756344795	1.0332411528	1.0424871445	1.0176811218
f	1.0408747196	1.0150277615	1.0157804489	1.0071839094
g	1.0618500710	1.0178625584	1.0322690010	1.0100055933
h	1.0562689304	1.0192553997	1.0275653601	1.0072581768
i	1.0308871269	1.0149402618	1.0106828213	1.0073467493
j	1.0384023190	1.0228306055	1.0125319958	1.0095375776
k	1.0393686295	1.0122373104	1.0165767670	1.0052092075
l	1.0321862698	1.0211771727	1.0108287334	1.0083882809
m	1.0571626425	1.0170327425	1.0283771753	1.0088099241
n	1.0639086962	1.0118888617	1.0350456238	1.0062968731
o	1.0819115639	1.0584974289	1.0469623804	1.0224372149
p	1.0617285967	1.0226594210	1.0317831039	1.0102237463
q	1.0720965862	1.0300858021	1.0401749611	1.0091934204
r	1.0329773426	1.0121468306	1.0128583908	1.0058519840
s	1.0681239367	1.0374056101	1.0327115059	1.0203156471
t	1.0337879658	1.0192551613	1.0106018782	1.0087591410
u	1.0648975372	1.0165529251	1.0357654095	1.0063078403
v	1.0425969362	1.0102100372	1.0186051130	1.0057344437
w	1.0638923645	1.0303624868	1.0287992954	1.0120517015
x	1.0377725363	1.0120788813	1.0158277750	1.0059341192
y	1.0364419222	1.0107214451	1.0150331259	1.0061452389
z	1.0546982288	1.0181393623	1.0260168314	1.0094790459
+	1.0358840227	1.0237274170	1.0138664246	1.0080053806
-	1.0249313116	1.0114963055	1.0089241266	1.0068209171
*	1.0311700106	1.0164752007	1.0129623413	1.0083116293
/	1.0214143991	1.0115058422	1.0073471069	1.0066345930
=	1.0502055883	1.0121105909	1.0243492126	1.0072437525
;	1.0380367041	1.0120087862	1.0165723562	1.0057575703
ı	1.0435506105	1.0124135017	1.0198223591	1.0060656071
(1.0224382877	1.0079113245	1.0076332092	1.0052670240
)	1.0189806223	1.0097336769	1.0072559118	1.0047732592
?	1.0515812635	1.0236015320	1.0239137411	1.0104774237

Table 9.4 Divider dimensions computed for 72 printed symbols (continued)

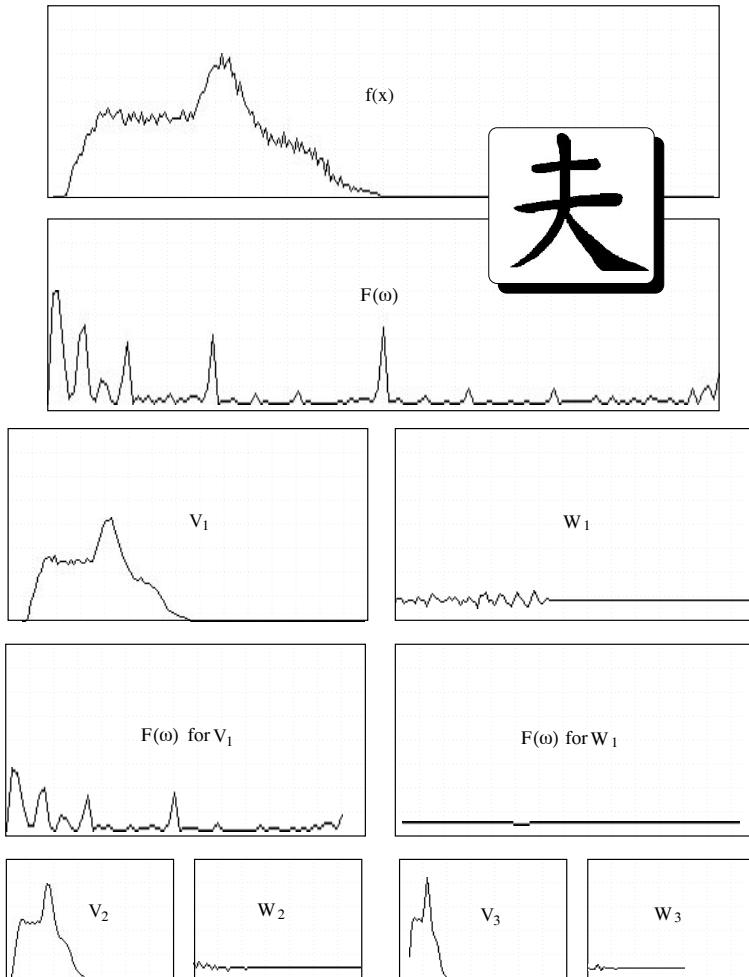


Fig. 9.17 Wavelet transform sub-patterns of Chinese character "Fu" (Husband) and their frequency analyses.

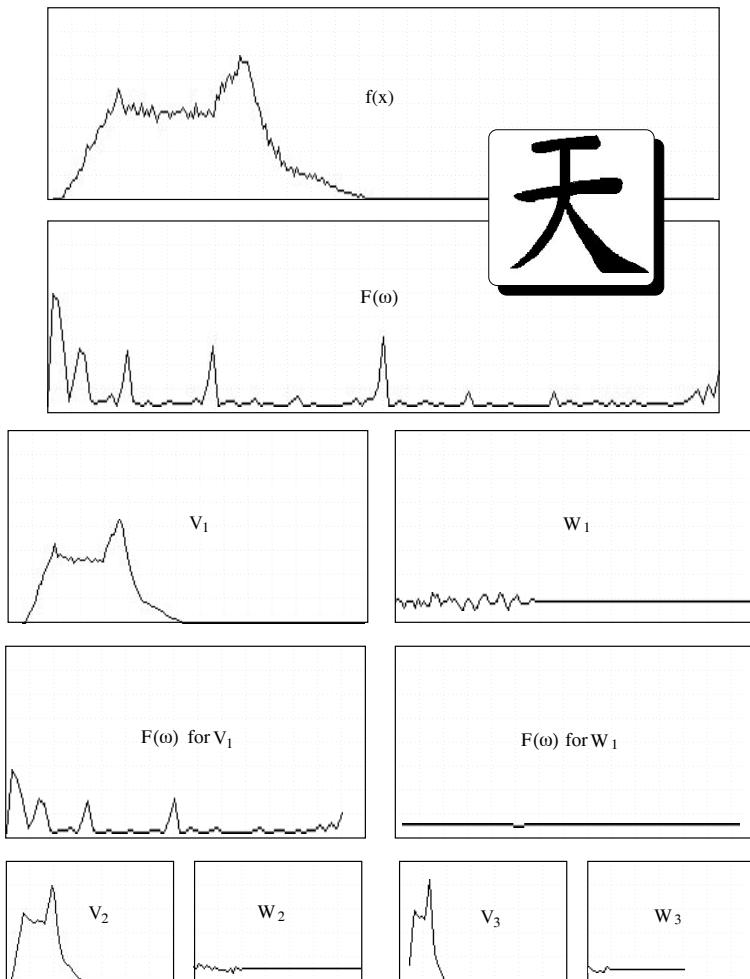


Fig. 9.18 Wavelet transform sub-patterns of Chinese character “Tian” (Sky) and their frequency analyses.

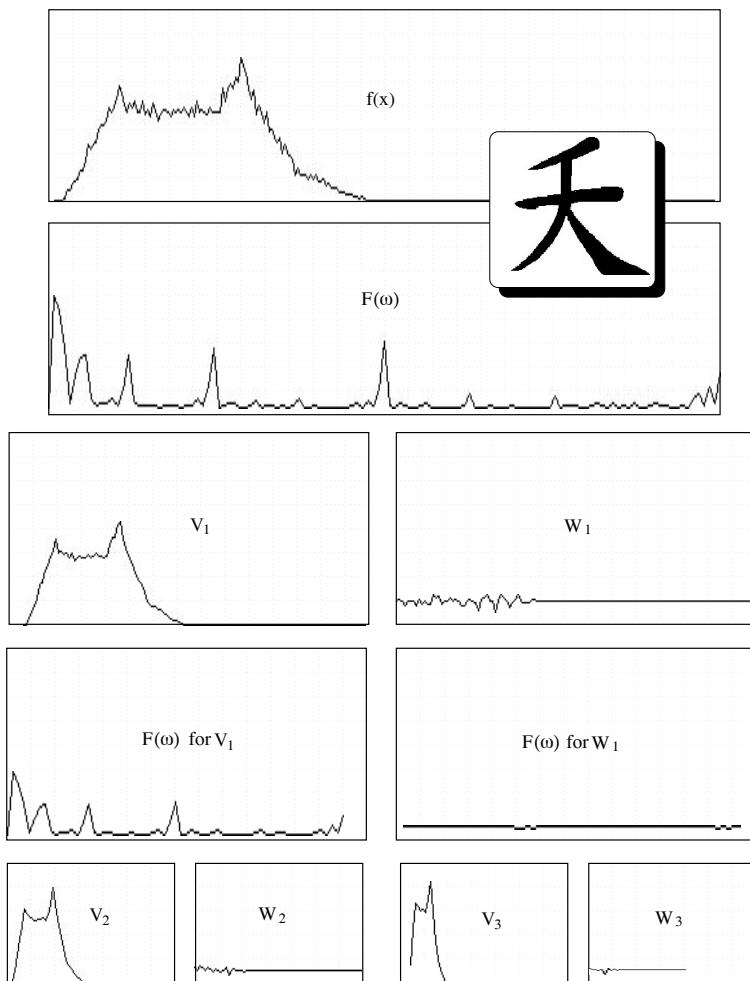


Fig. 9.19 Wavelet transform sub-patterns of Chinese character "Yao" (Die prematurely) and their frequency analyses.

This page intentionally left blank

Chapter 10

Document Analysis by Reference Line Detection with 2-D Wavelet Transform

Document processing has become a very active topic in areas of pattern recognition, office automation, artificial intelligence and knowledge engineering for a decade [Tang et al., 1994]. The acquisition of knowledge from documents by an information system can involve an extensive amount of hand-crafting which can be time-consuming. Actually, it is a bottleneck of information systems. Thus, automatic knowledge acquisition from documents is an important subject, and many researchers are trying to find new techniques of processing documents. In reality, it is very difficult to develop a general system that can process all kinds of documents, such as technical reports, government files, newspapers, books, journals, magazines, letters, bank cheques, etc. As the first step, many researchers concentrated they studies to developing specific ones to treat some specific types of documents. After carefully studying the major characteristics of different types of documents, the specific properties of form documents have been analyzed in our earlier work [Tang et al., 1995c]. According to these properties, we have taken the approach of building a simpler system for processing form documents, instead of a complex one for all sorts of documents.

First, let us consider the major characteristics of forms:

- In general, a form consists of straight lines, which are oriented mostly in horizontal and vertical directions. These lines are referred to as *reference lines*, which can be found in an example, a Canadian cheque, shown in Fig. 10.1(a).
- The reference lines are pre-printed to guide the users to complete the form. A typical example is a bank cheque, where reference lines are printed to guide the users to write in the name of the payee,

THE TORONTO-DOMINION BANK
 1400 SHERBROOKE ST. W. & BISHOP STREET
 MONTREAL, QUEBEC
 H3C 1K4

Nov 12 1990

PAY TO THE ORDER OF Computer Science .210⁰⁰
Two hundred ten DOLLARS

Elaine Henry

147601-00412 (a)

(1) Date	<u>Nov 12 1990</u>
(2) Numeric Amount	<u>.210</u> ⁰⁰
(3) Legal Amount	<u>Two hundred ten</u>
(4) Signature	<u>Elaine Henry</u>

(b)

Fig. 10.1 (a) A Canadian cheque image, and (b) the item images extracted from (a).

amount, and date in the appropriate places.

- Not all the information in a form is useful. The information that should be entered to computer and processed is usually the filled data. For example, in Fig. 10.1(a), the date, numeric amount, legal amount and signature are the filled data. These data shown in Fig. 10.1(b) need to be extracted from the cheque image Fig. 10.1(a).
- In order to indicate the filling position, the reference lines can be used and the filled information usually appears either above, beneath, or beside these reference lines. Thus, in form processing, the reference lines have to be detected first, then we can find the useful information from a form based on them, and thereafter, enter to the computers.

Consequently, extraction of such reference lines plays a very important role in the form processing. However, the extraction of the reference lines from a complex-background document is a difficult task. A traditional method to detect lines from images is the Hough transform. In 1994, the journal IEEE PAMI published a new method called SLIDE (Subspace-Based Line Detection) proposed by Aghajan and Kailath [Aghajan and Kailath, 1994]. The SLIDE yields closed-form and high resolution estimates for line parameters, and its computational complexity and storage requirement are far less than those of the Hough transform.

In this chapter, a novel wavelet-based method will be presented. In this method, two-dimensional multiresolution analysis (MRA), wavelet decomposition algorithm , and compactly supported orthonormal wavelets are used to transform a document image into several sub-images. Based on these sub-images, the reference lines of a complex-background document can be extracted, and knowledge about the geometric structure of the document can be acquired. Particularly, this approach appears to be more efficient in processing form documents with multi-grey level background.

10.1 Two-Dimensional MRA and Mallat Algorithm

The principle of the 1-D multiresolution analysis (MRA) presented in Chapter 3 can be extended directly to 2-D multiresolution by replacing $L^2(\mathbb{R})$ with $L^2(\mathbb{R}^2)$.

First of all, we declare that, as in the last chapter we use the subspaces in $\{V_j\}$ by the order of

$$\cdots \subset V_{j+1} \subset V_j \subset V_{j-1}$$

when we deal with MRAs in this chapter.

Definition 10.1 Let $\{V_j\}_j \in \mathbb{Z}$ be a sequence of closed subspaces in $L^2(\mathbb{R})$. The 2-D MRA can be constructed by *tensor product space* $\{V_j^2\}_{j \in \mathbb{Z}}$ if and only if $\{V_j\}_{j \in \mathbb{Z}}$ is a 1-D MRA in $L^2(\mathbb{R})$, and

$$V_j^2 = V_j \otimes V_j. \quad (10.1)$$

The scaling function $\Phi(x, y)$ for 2-D MRA has the form of

$$\Phi(x, y) = \varphi(x)\varphi(y),$$

where φ is the real scaling function of the 1-D multiresolution analysis $\{V_j\}_{j \in \mathbb{Z}}$. For each $j \in \mathbb{Z}$, the orthonormal bases of V_j^2 can be produced by the function system of

$$\{\Phi_{j,k_1,k_2} = \varphi_{j,k_1}(x)\varphi_{j,k_2}(y) | (k_1, k_2) \in \mathbb{Z}^2\}.$$

Such multiresolution analysis $\{V_j^2\}_{j \in \mathbb{Z}}$ in space $L^2(\mathbb{R}^2)$ is called *divisible MRA*.

We define a wavelet space $W_j^2 = (V_j^2)^\perp$, i.e. $V_j^2 \oplus W_j^2 = V_{j-1}^2$. Thus the wavelet function consists of three basic wavelet functions: ψ^1 , ψ^2 and ψ^3 . The orthonormal bases of the wavelet space W_j^2 can be obtained from the “single wavelets” ψ^1 , ψ^2 and ψ^3 by a *binary dilation* i.e. dilation by 2^j and *dyadic translation* (of $k/2^j$). More precisely, we have the following theorem:

Theorem 10.1 Let $\{V_j^2\}_{j \in \mathbb{Z}}$ be a divisible MRA in $L^2(\mathbb{R}^2)$: $V_j^2 = V_j \otimes V_j$, where $\{V_j\}_{j \in \mathbb{Z}}$ is a 1-D MRA in the space $L^2(\mathbb{R})$ with scaling function φ and wavelet function ψ . We define the following three functions:

$$\begin{cases} \psi^1(x, y) = \varphi(x)\psi(y) \\ \psi^2(x, y) = \psi(x)\varphi(y) \\ \psi^3(x, y) = \psi(x)\psi(y) \end{cases}$$

For any $j \in \mathbb{Z}$, the orthonormal bases of the space W_j^2 can be obtained from the following function system:

$$\begin{cases} \Psi_{j,k,m}^1 = \varphi_{j,k}(x)\psi_{j,m}(y) \\ \Psi_{j,k,m}^2 = \psi_{j,k}(x)\varphi_{j,m}(y) \\ \Psi_{j,k,m}^3 = \psi_{j,k}(x)\psi_{j,m}(y) \end{cases}$$

Therefore, the function system

$$\{\Psi_{j,k,m}^e | e = 1, 2, 3; j, k, m \in \mathbb{Z}\} \quad (10.2)$$

becomes a set of orthonormal bases of $L^2(\mathbb{R}^2)$.

Proof

From Eq. (10.1) a very significant formula can be produced

$$\begin{aligned} V_{j-1}^2 &= V_{j-1} \otimes V_{j-1} \\ &= (V_j \oplus W_j) \otimes (V_j \oplus W_j) \\ &= (V_j \otimes V_j) \oplus (V_j \otimes W_j) \oplus (W_j \otimes V_j) \oplus (W_j \otimes W_j). \end{aligned} \quad (10.3)$$

It can be rewritten as

$$\begin{aligned}
 V_{j-1}^2 &= V_{j-1} \otimes V_{j-1} \\
 &= (V_j \otimes V_j) \oplus (V_j \otimes W_j) \oplus (W_j \otimes V_j) \oplus (W_j \otimes W_j) \\
 &= V_j^2 \oplus \underbrace{[(V_j \otimes W_j) \oplus (W_j \otimes V_j) \oplus (W_j \otimes W_j)]}_{W_j^2}
 \end{aligned} \tag{10.4}$$

Since $\{\varphi_{j,k}|k \in \mathbb{Z}\}$ is an orthonormal base for V_j , the set $\{\psi_{j,k}|k \in \mathbb{Z}\}$ becomes the orthonormal bases of the space W_j . Therefore,

- $\{\Psi_{j,k,m}^1|m \in \mathbb{Z}\}$ is an orthonormal base of $V_j \otimes W_j$;
- $\{\Psi_{j,k,m}^2|m \in \mathbb{Z}\}$ is an orthonormal base of $W_j \otimes V_j$;
- $\{\Psi_{j,k,m}^3|m \in \mathbb{Z}\}$ is an orthonormal base of $W_j \otimes W_j$.

Consequently, Eq. (10.4) indicates that the function represented by Eq. (10.2) has constituted the orthonormal bases of W_j^2 . ■

Let P_j , D_j^1 , Q_j^2 and Q_j^3 be projection operators from $L^2(\mathbb{R}^2)$ to its subspaces $(V_j \otimes V_j)$, $(V_j \otimes W_j)$, $(W_j \otimes V_j)$ and $(W_j \otimes W_j)$ respectively. In practice, the document image $f(x, y) \in V_{j_1}^2$ has a limited resolution, namely: j_1 is a certain integer. We have

$$\begin{aligned}
 f(x, y) &= P_{j_1} f(x, y) \\
 &= \sum_{k_1 \in \mathbb{Z}} \sum_{k_2 \in \mathbb{Z}} c_{j_1, k_1, k_2} \Phi_{j_1, k_1, k_2} \\
 &= \sum_{k_1 \in \mathbb{Z}} \sum_{k_2 \in \mathbb{Z}} c_{j_1, k_1, k_2} \varphi_{j_1, k_1}(x) \varphi_{j_1, k_2}(y) \\
 &= P_{j_1+1} f + Q_{j_1+1}^1 f + Q_{j_1+1}^2 f + Q_{j_1+1}^3 f,
 \end{aligned} \tag{10.5}$$

where

$$c_{j_1, k_1, k_2} = \langle P_{j_1+1} f(x, y), \varphi_{j_1, k_1}(x) \varphi_{j_1, k_2}(y) \rangle,$$

and

- $P_{j_1} f \in V_{j_1} \otimes V_{j_1}$;
- $P_{j_1+1} f \in V_{j_1+1} \otimes V_{j_1+1}$;
- $Q_{j_1+1}^1 f \in V_{j_1+1} \otimes W_{j_1+1}$;
- $Q_{j_1+1}^2 f \in W_{j_1+1} \otimes V_{j_1+1}$;
- $Q_{j_1+1}^3 f \in W_{j_1+1} \otimes W_{j_1+1}$.

In Eq. (10.5), $P_{j_1+1}f$ and $Q_{j_1+1}^\beta f$ ($\beta = 1, 2, 3$) can be computed by

$$\begin{aligned} P_{j_1+1}f &= \sum_{m_1 \in \mathbb{Z}} \sum_{m_2 \in \mathbb{Z}} c_{j_1+1, m_1, m_2} \Phi_{j_1+1, m_1, m_2} \\ &= \sum_{m_1 \in \mathbb{Z}} \sum_{m_2 \in \mathbb{Z}} c_{j_1+1, m_1, m_2} \varphi_{j_1+1, m_1}(x) \varphi_{j_1+1, m_2}(y), \end{aligned}$$

and

$$\begin{aligned} Q_{j_1+1}^\beta f &= \sum_{m_1 \in \mathbb{Z}} \sum_{m_2 \in \mathbb{Z}} d_{j_1+1, m_1, m_2}^\beta \Psi_{j_1+1, m_1, m_2}^\beta \\ &= \sum_{m_1 \in \mathbb{Z}} \sum_{m_2 \in \mathbb{Z}} d_{j_1+1, m_1, m_2}^\beta \psi_{j_1+1, m_1}^\beta(x) \psi_{j_1+1, m_2}^\beta(y) \quad (\beta = 1, 2, 3). \end{aligned}$$

Eq. (10.5) can be written below:

$$\begin{aligned} f(x, y) &= P_{j_1+1}f + Q_{j_1+1}^1f + Q_{j_1+1}^2f + Q_{j_1+1}^3f \\ &= \sum_{m_1 \in \mathbb{Z}} \sum_{m_2 \in \mathbb{Z}} c_{j_1+1, m_1, m_2} \varphi_{j_1+1, m_1}(x) \varphi_{j_1+1, m_2}(y) \\ &= \sum_{m_1 \in \mathbb{Z}} \sum_{m_2 \in \mathbb{Z}} d_{j_1+1, m_1, m_2}^1 \varphi_{j_1+1, m_1}(x) \psi_{j_1+1, m_2}(y) \\ &= \sum_{m_1 \in \mathbb{Z}} \sum_{m_2 \in \mathbb{Z}} d_{j_1+1, m_1, m_2}^2 \psi_{j_1+1, m_1}(x) \varphi_{j_1+1, m_2}(y) \\ &= \sum_{m_1 \in \mathbb{Z}} \sum_{m_2 \in \mathbb{Z}} d_{j_1+1, m_1, m_2}^3 \psi_{j_1+1, m_1}(x) \psi_{j_1+1, m_2}(y). \end{aligned}$$

An iterative algorithm called *Mallat algorithm* is presented as follows:

$$\left\{ \begin{array}{l} c_{j_1+1, m_1, m_2} = \sum_{k_1 \in \mathbb{Z}} \sum_{k_2 \in \mathbb{Z}} h_{k_1-2m_1} h_{k_2-2m_2} c_{j_1, k_1, k_2} \\ d_{j_1+1, m_1, m_2}^1 = \sum_{k_1 \in \mathbb{Z}} \sum_{k_2 \in \mathbb{Z}} h_{k_1-2m_1} g_{k_2-2m_2} c_{j_1, k_1, k_2} \\ d_{j_1+1, m_1, m_2}^2 = \sum_{k_1 \in \mathbb{Z}} \sum_{k_2 \in \mathbb{Z}} g_{k_1-2m_1} h_{k_2-2m_2} c_{j_1, k_1, k_2} \\ d_{j_1+1, m_1, m_2}^3 = \sum_{k_1 \in \mathbb{Z}} \sum_{k_2 \in \mathbb{Z}} g_{k_1-2m_1} g_{k_2-2m_2} c_{j_1, k_1, k_2} \end{array} \right. . \quad (10.6)$$

Let $H_r = (H_{k_1, m_1})$, $H_c = (H_{k_2, m_2})$, $G_r = (G_{k_1, m_1})$, and $G_c = (G_{k_2, m_2})$ be matrices. The subscript “r” indicates an operation on the rows of the matrix, while the subscript “c” for operation on the column. Thus, Eq. (10.6) can be represented by a simple form:

$$\begin{cases} C_{j_1+1} = H_r H_c C_{j_1} \\ Q_{j_1+1}^1 = H_r G_c C_{j_1} \\ Q_{j_1+1}^2 = G_r H_c C_{j_1} \\ Q_{j_1+1}^3 = G_r G_c C_{j_1} \end{cases}. \quad (10.7)$$

$P_{j_1+1}f$ will be decomposed in the same way, and $P_{j_1+2}f$ and $Q_{j_1+2}^\beta f$ will be produced. This iteration procedure will be continued. After $j_2 - j_1$ steps, we arrive at

$$f(x, y) = P_{j_2}f(x, y) + \sum_{j=j_1+1}^{j_2} \sum_{\beta=1}^3 Q_j^\beta f(x, y) \quad (10.8)$$

10.2 Detection of Reference Line from Sub-Images by the MRA

In this section, the basic idea of form analysis by the MRA will be presented. First, we will introduce the properties of the wavelet transformed sub-images. A document image can be transformed into four sub-images by applying the Mallat algorithm, namely, (1) LL sub-image, (2) LH sub-image, (3) HL sub-image, and (4) HH sub-image. According to Eq. (10.3), these sub-images possess the following properties:

- LL sub-image: both horizontal and vertical directions have low-frequencies. It corresponds to $(V_j \otimes V_j)$, and its orthonormal basis is $\{\Phi_{j,k,m} | k, m \in \mathbb{Z}\}$.
- LH sub-image: the horizontal direction has low-frequencies, and the vertical one has high-frequencies. It corresponds to $(V_j \otimes W_j)$, and its orthonormal basis is $\{\Psi_{j,k,m}^1 | k, m \in \mathbb{Z}\}$.
- HL sub-image: the horizontal direction has high-frequencies, and the vertical one has low-frequencies. It corresponds to $(W_j \otimes V_j)$, and its orthonormal basis is $\{\Psi_{j,k,m}^2 | k, m \in \mathbb{Z}\}$.
- HH sub-image: both horizontal and vertical directions have high-frequencies. It corresponds to $(W_j \otimes W_j)$, and its orthonormal basis

is $\{\Psi_{j,k,m}^3 | k, m \in \mathbb{Z}\}$.

An example is illustrated in Fig. 10.2. The original image is a square with grey level shown in Fig. 10.2(a). It has been transformed into four sub-images by the MRA, illustrated in Fig. 10.2(b). The LL sub-image is the result from a filter, which allows lower frequencies to pass through along the horizontal direction as well as the vertical direction. That is a “smoothing” effect on both directions. The HH sub-image comes from a filter where the higher frequency components can cross it along both directions. That is an “enhancing” effect on the horizontal and vertical directions.

We are interested in the LH and HL sub-images. The LH sub-image is achieved from a filter which allows lower frequency components to reach across along the horizontal direction, as well as the higher frequencies along the vertical direction. That is an “enhancing” effect on the vertical, and “smoothing” effect on the horizontal. The result of the HL sub-image is opposite to that of the LH one. In this way, the horizontal direction of the filter opens for the higher frequencies, and the vertical direction for lower frequency components. That is an “enhancing” effect on the horizontal, and “smoothing” effect on the vertical.

Another example can be shown in Figs. 10.3 and 10.4. The input image is illustrated in Fig. 10.3, which contains several lines with complex-background including texts. Its LH and HL sub-images are depicted in Figs. 10.4(a) and (b) respectively.

From Fig. 10.4(a), it is clear that only horizontal lines remain in the LH sub-image, while only vertical ones remain in the HL sub-image as shown in Fig. 10.4(b). Both the LH and HL sub-images keep only straight lines, the grey-level background is removed. We are interested in Figs. 10.4(b) and (c), because these important properties can be used to extract the horizontal and vertical lines in a form document, which has complex-background. In form documents, the information that should be entered to the computer and processed is usually the filled data. In order to indicate the filling position, some pre-printed reference lines should be extracted. The useful information, in general, is either above, beneath, or beside these specific lines. In our work, the LH and HL sub-images have been used to extract such lines.

To perform the above proposed method, an MRA algorithm and compactly supported orthonormal wavelets are used.

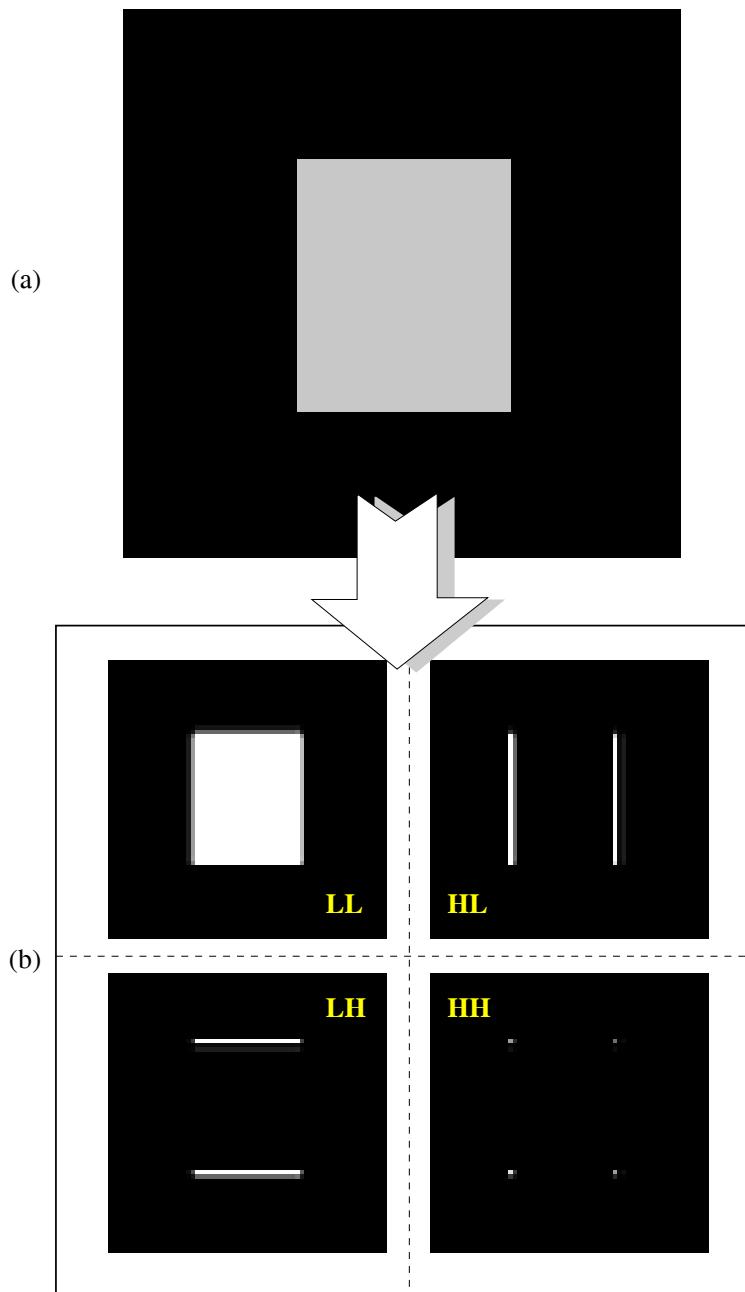


Fig. 10.2 A square image and its sub-images.

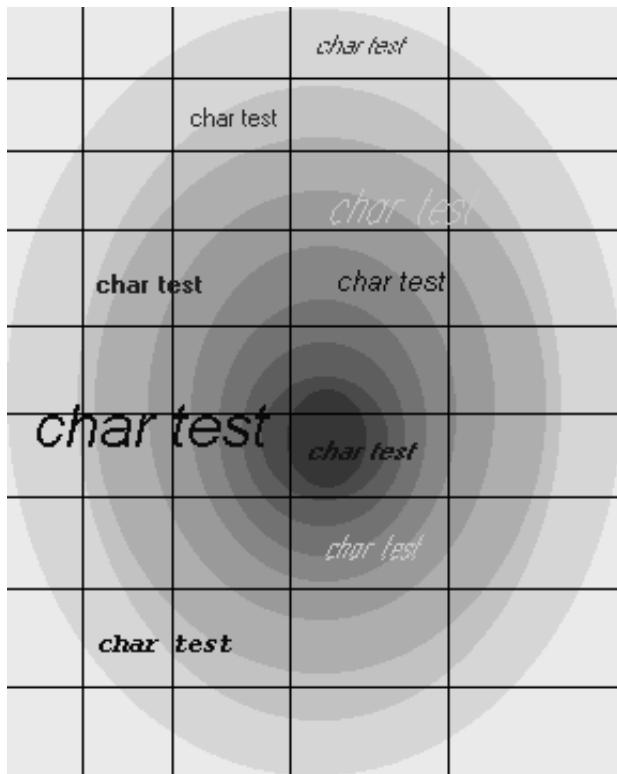


Fig. 10.3 Input image.

10.3 Experiments

Using sub-images of wavelet, experiments have been conducted to process financial documents by a personal computer system. The document is entered into the system by an optical scanner and converted to a digital image. A document can be digitized into a color image, monochrome grey scale image or binary image according to different requirements and applications. In our experiments, an HP scanner is employed to capture the image of the document. The resolution of digitization in our experiments can vary over a range of 200-300 DPI. All documents are converted to gray scale images in the experiments. We have examined a variety of form documents originated from United States, Canada, and China. They contain alpha-

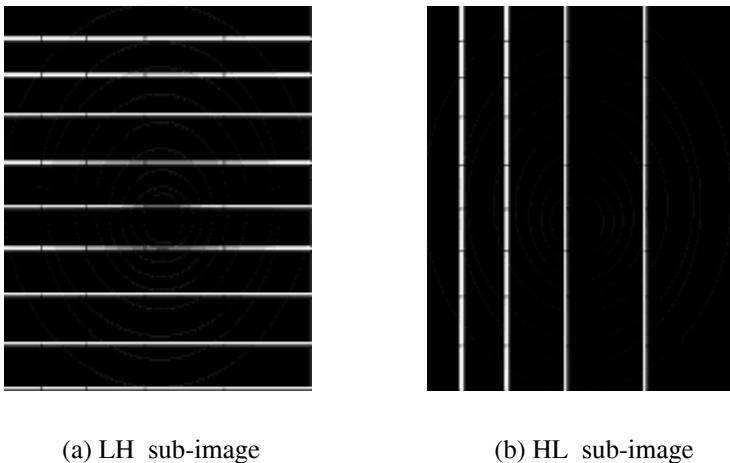


Fig. 10.4 LH and HL sub-images.

numerics, Chinese characters, some special symbols and graphics. Both simple financial documents such as Canadian bank cheques and complex documents such as Canadian Federal tax return forms and deposit forms of Chinese bank have been tried by our new method. Because of the page limit in this chapter, only a few examples will be presented, namely:

- (1) Cheque of a Canadian bank (TD Bank) (Fig. 10.5);
- (2) Portion of a Federal tax return form (Fig. 10.6);
- (3) Deposit form of Chinese bank (Fig. 10.8).

In these experiments, the 2-D MRA algorithm has been applied. The compactly supported orthonormal wavelets used in our study have been chosen from [Daubechies, 1988], and the values of h_k 's are listed in Tables 10.1, 10.2, 10.3, and 10.4.

For the Canadian Bank cheque shown in Fig. 10.5(a), there exists a grey-level background on it. To remove the grey-level background, the special properties of the sub-images can be used. Precisely, the grey-level background has been removed in both LH sub-image and HL sub-image.

The LH sub-image results from a filter which allows lower frequencies to pass through along the horizontal direction, and higher frequencies along the vertical direction. That is an “enhancing” effect on the vertical, and “smoothing” effect on the horizontal. As a result, only horizontal lines

$h_N(n)$			$h_N(n)$		
N=2	n=0	0.482 962 913 145 1 2 3 -0.129 409 522 551	N=3	n=0	0.332 670 552 950 1 2 3 4 5 -0.135 011 020 010 -0.085 441 273 882 0.035 226 291 882
N=4	n=0	0.230 377 813 309 1 2 3 4 5 6 7 -0.027 983 769 417 -0.187 034 811 719 0.030 841 381 836 0.032 883 011 667 -0.010 597 401 785	N=5	n=0	0.160 102 397 974 1 2 3 4 5 6 7 8 9 0.603 829 269 797 0.724 308 528 438 0.138 428 145 901 -0.242 294 887 066 -0.032 244 869 585 0.077 571 493 840 -0.006 241 490 213 -0.012 580 751 999 0.003 335 725 285

Table 10.1 Values of $h_N(n)$'s, $N = 2 \sim 5$.

remain in the LH sub-image. The HL sub-image is opposite to the LH one. The high frequency spectrum can pass through along the horizontal direction, while low frequency spectrum along the vertical direction. That is an “enhancing” effect on the horizontal, and “smoothing” effect on the vertical. Thus, only vertical lines remain in the HL sub-image.

In our experiments, the special property of the LH sub-image has been used. To produce the LH sub-image, the two-dimensional multiresolution analysis (MRA) algorithm and compactly supported orthonormal wavelets have been applied to the documents. The LH sub-image of Fig. 10.5(a) has been obtained and shown in Fig. 10.5(b). Since only horizontal lines remain in the LH sub-image, reference lines, which guides the writer to fill data in the proper location, can be extracted.

A point worth noting to mention is that the number of pixels contained in any sub-image is one fourth of that of the original document image. Thus, the size of LH sub-image shown in Fig. 10.5(b) is one fourth of that of the original cheque image shown in Fig. 10.5(a). To achieve a high quality image and correctly map the filled data with reference lines, the

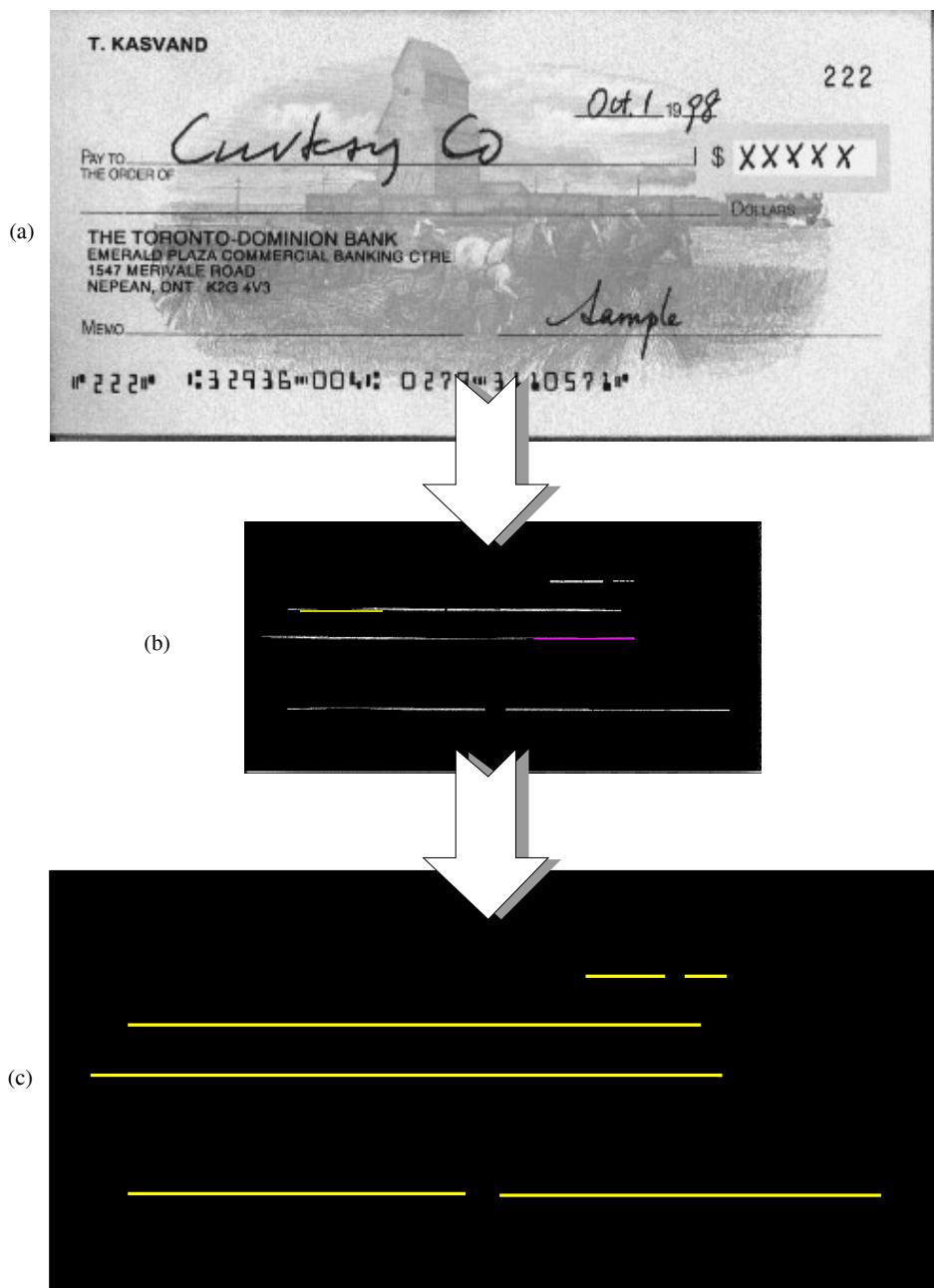


Fig. 10.5 A bank cheque and its original and enhanced LH sub-images.

$h_N(n)$			$h_N(n)$		
N=6	n=0	n=1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11	N=7	n=0	n=1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13
		0.111 540 743 350 0.494 623 890 398 0.751 133 908 021 0.315 250 351 709 -0.226 264 693 965 -0.129 766 867 567 0.097 501 605 587 0.027 522 865 530 -0.031 582 039 318 0.000 553 842 201 0.004 777 257 511 -0.001 077 301 085			0.077 852 054 085 0.396 539 319 482 0.729 132 090 846 0.469 782 287 405 -0.143 906 003 929 -0.224 036 184 994 0.071 309 219 267 0.080 612 609 151 -0.038 029 936 935 -0.016 574 541 631 0.012 550 998 556 0.000 429 577 937 -0.001 801 640 704 0.000 353 713 800

Table 10.2 Values of $h_N(n)$'s, $N = 6 \sim 7$.

LH sub-image has been processed by regular enhancement and smoothing techniques, and has been scaled up to the same size as the original image. The result can be found in Fig. 10.5(c).

For the form shown in Fig. 10.6, there are altogether 20 horizontal lines and many text strings in it. The result of extracting the pre-printed reference lines from the LH sub-image of wavelet is presented in Fig. 10.7.

Note that the extracted lines in both Fig. 10.5(b) and Fig. 10.7(a) are gray-levels, since the document images to be processed in our method are gray scale images.

The last example is a deposit form of Chinese bank as shown in Fig. 10.8, which is rather complicated. The original image of the Chinese bank deposit form is given in Fig. 10.8(a). As applying wavelet decomposition algorithm, the reference lines are detected, and shown in Fig. 10.8(b). Based on these reference lines, the images of the some items which contain the useful information are extracted, the results are displayed in Fig. 10.8(c).

$h_N(n)$			$h_N(n)$		
N=8	n=0	0.054 415 842 243	N=9	n=0	0.038 077 947 364
	1	0.312 871 590 914		1	0.243 834 674 613
	2	0.675 630 736 297		2	0.604 823 123 690
	3	0.585 354 683 654		3	0.657 288 078 051
	4	-0.015 829 105 256		4	0.133 197 385 825
	5	-0.284 015 542 962		5	-0.293 273 783 279
	6	0.000 472 484 574		6	-0.096 840 783 223
	7	0.128 747 426 620		7	0.148 540 749 338
	8	-0.017 369 301 002		8	0.030 725 681 479
	9	-0.044 088 253 931		9	-0.067 832 829 061
	10	0.013 981 027 917		10	0.000 250 947 115
	11	0.008 746 094 047		11	0.022 361 662 124
	12	-0.004 870 352 993		12	-0.004 723 204 758
	13	-0.000 391 740 373		13	-0.004 281 503 682
	14	0.000 675 449 406		14	0.001 847 646 883
	15	-0.000 117 476 784		15	0.000 230 385 764
				16	-0.000 253 963 189
				17	0.000 039 347 320

Table 10.3 Values of $h_N(n)$'s, $N = 8 \sim 9$.

$h_N(n)$			$h_N(n)$		
N=10	n=0	0.026 670 057 901	N=10	n=10	-0.029 457 536 822
	1	0.188 176 800 078		11	0.033 212 674 059
	2	0.527 201 188 932		12	0.003 606 553 567
	3	0.688 459 039 454		13	-0.010 733 157 483
	4	0.281 172 343 661		14	0.001 395 351 747
	5	-0.249 846 424 327		15	0.001 992 405 295
	6	-0.195 946 274 377		16	-0.000 685 856 695
	7	0.127 369 340 336		17	-0.000 116 466 855
	8	0.093 057 364 604		18	0.000 093 588 670
	9	-0.071 394 147 166		19	-0.000 013 264 203

Table 10.4 Values of $h_N(n)$'s, $N = 10$.

	Allowable credit	414
Add lines 410, 412, and 414	416	
line 406; if negative, enter "0")		
the amount from line 406 in Table B. If you are using		
Add lines 417 and 419. This is your net federal tax credit		
428 in the guide)		
Add lines 420 through 428. This is your total tax		
guide)	437	3256
437)	438	449
look up the amount from line 406 in 0 of that schedule)	439	347
	440	7001

Fig. 10.6 Portion of a Federal tax return form.

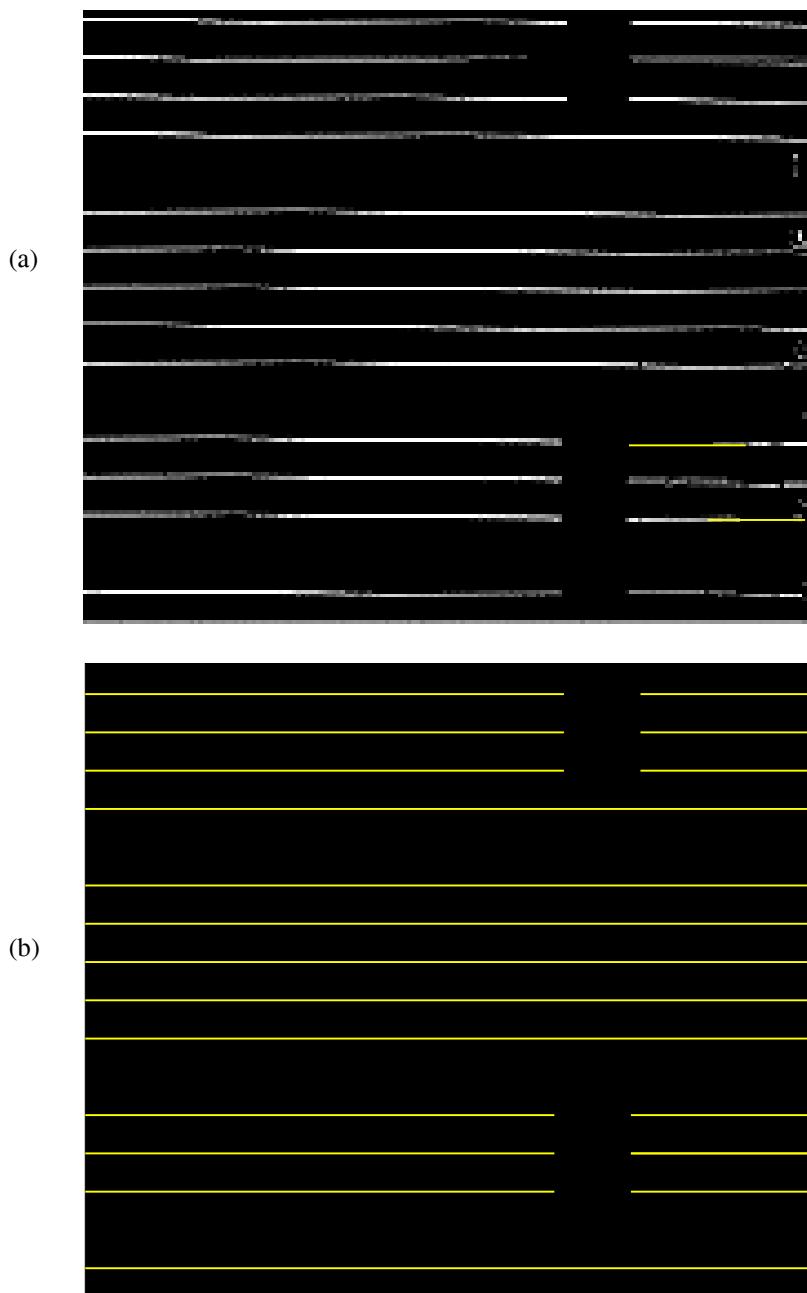
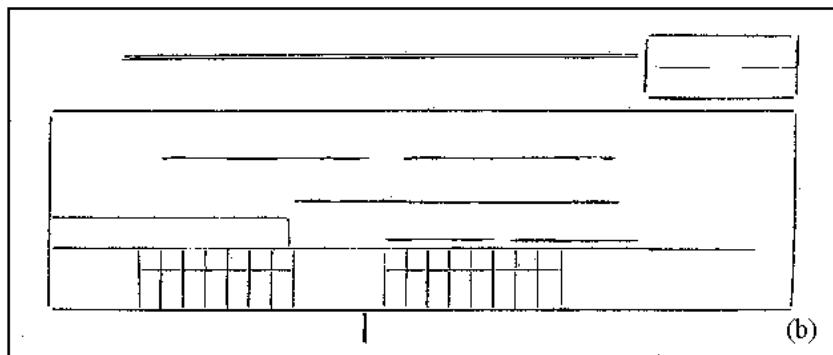


Fig. 10.7 (a) LH sub-image of the portion of a Federal tax return form shown in Fig. 10.6; (b) the enhanced image.

代号3066 中国工商银行 活期储蓄存款凭条(代收入传票)

科目(收入)	1990年4月23日	凭证号	号
		字第	号
帐号 96745 户名 赵万		附件 张	
存入人民币(大写) 肆佰壹拾贰元			
以下各栏由银行填写 (小写) 元: 412.00			
存款余额	利息余额	地址(开户时填写)	
或积数			
事后监查:		储蓄所: 直接	出纳
		记帐	(a)



1990年4月23日
96745 赵万
肆佰壹拾贰元 412.00
(c)

Fig. 10.8 (a) An image of the deposit form of Chinese bank; (b) the reference lines detected from (a); and (c) the item images extracted from (a) according to the reference lines.

Chapter 11

Chinese Character Processing with B-Spline Wavelet Transform

Chinese character recognition is a very significant branch in pattern recognition, and Chinese character processing is an important technology within it. Chinese character processing is to operate and modify the Chinese characters including generation, storage, display, printing, transferring, geometric transformation, etc. with the modern computer.

In this chapter, several algorithms for Chinese character processing are studied, based on cubic B-spline wavelet transform, namely:

- (1) Compression of Chinese character,
- (2) Arbitrary enlargement of the typeface size of Chinese character,
- (3) Generation of Chinese type style.

The outline of Chinese character processing by wavelet transform presented in this chapter is graphically illustrated in Fig. 11.1. It consists of three stages, precisely, (1) pre-processing, (2) wavelet transform, and (3) objective processing. They will be presented below:

(1) Pre-processing:

The original character is entered into a computer system by scanning, and it is further undergone by a contour extraction operator. Thus, a Chinese character is converted to one or more contours, which can be considered to be curves. Each curve is represented by an array of coordinate points, $Q_i(i = 1, 2, \dots, M)$. It is then sent to the next stage.

(2) Wavelet transform:

After the pre-processing stage, a Chinese character is already represented by its contours (curves), and each of them is formed by an array

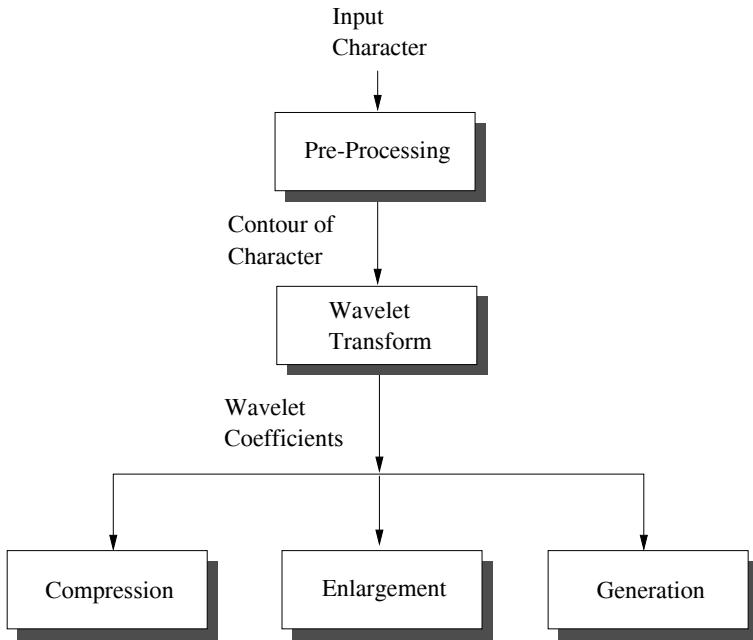


Fig. 11.1 Procedure of Chinese character processing with wavelet transform.

of coordinate points $Q_i (i = 1, 2, \dots, M)$. In this stage, the cubic B-spline function is employed to interpolate such a parameter curve. Suppose $S^n(t)$ indicates the interpolation curve, which is produced in accordance with the cubic B-spline function, and passes through these coordinate points.

The interpolation curve $S^n(t)$ using cubic B-spline can be written as

$$S^n(t) = (s_0^n(t), s_1^n(t), \dots, s_{M-2}^n(t)), \quad (11.1)$$

where n stands for the level number of decompositions in wavelet analysis, which will be presented later. $s_i^n(t)$ denotes the i^{th} interpolation sub-curve with cubic B-spline and can be expressed by

$$s_i^n(t) = \sum_{l=1}^4 c_{i+l-1}^n \cdot N_{l,3}(t), \quad 0 \leq i \leq M-2, \quad 0 \leq t \leq 1, \quad (11.2)$$

where $N_{l,3}(t)$ is a basis function of the cubic B-spline interpolation, which means polynomial spline function with equally spaced sample knots. $C^n = (c_0^n, c_1^n, \dots, c_{M+1}^n)$ is a coefficient sequence, which is called the control point

of the cubic B-spline interpolation curve. These control points correspond to the coordinate points (sometimes, those points are called vertices of eigen-polygon of B-spline).

The relationship between the coordinate points and the control points is

$$\frac{(c_{i-1}^n + 4c_i^n + c_{i+1}^n)}{6} = Q_i, \quad i = 1, 2, \dots, M, \quad (11.3)$$

there are M equations in (11.3), but there are $M + 2$ unknown variables. Hence, two boundary conditions are needed to supply. For the interpolation curve with cubic B-spline, we have the following at the endpoints:

$$c_0^n = c_1^n, \quad c_{M+1}^n = c_M^n.$$

Consequently, we can obtain a linear equation as follows:

$$\begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 \\ \frac{1}{6} & \frac{4}{6} & \frac{1}{6} & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \frac{1}{6} & \frac{4}{6} & \frac{1}{6} \\ 0 & 0 & 0 & 0 & -1 & 1 \end{bmatrix} \cdot \begin{bmatrix} c_0^n \\ c_1^n \\ \dots \\ c_M^n \\ c_{M+1}^n \end{bmatrix} = \begin{bmatrix} 0 \\ Q_1 \\ \dots \\ Q_M \\ 0 \end{bmatrix} \quad (11.4)$$

According to the above analysis, it is clear that

- Owing to (11.2) and (11.4), we can find the relationship among the coordinate points Q_i , sub-curves s_i^n and control points c_i^n .
- A cubic B-spline curve constructed by $C^n = (c_0^n, c_1^n, \dots, c_{M+1}^n)$ passes through the coordinate points $Q_i (i = 1, 2, \dots, M)$.
- By (11.2), four control points produce a sub-curve. M coordinate points express $M - 1$ sub-curves, i.e., a curve needs $M + 2$ control points.
- The coordinate points can be converted to the control points by sloving equation (11.4).

Next, the multiresolution analysis in wavelet theory is applied to describe a curve $S^n(t)$. In this way, the control points, by which the curve is formed, are transformed with wavelet function. For effectively decomposing curve, the number of sub-curves is usually 2 to the power n , i.e., curve $S^n(t)$ should have $2^n + 3$ control points. When the control points are less than $2^n + 3$, the coordinate points must be extended as well the control points. The performance of the extension can be periodical spinning out or

adding zeros or re-sampling. For a Chinese character, which is composed of several closed curves, the periodic extension is usually utilized.

When the number of the control points is equal to $2^n + 3$, after the first decomposition by wavelet transform, the sequence of the control points C^n becomes C^{n-1} and D^{n-1} . Thereafter, C^{n-1} is decomposed into C^{n-2} and D^{n-2} again. This process continues, as a total number of n decompositions, finally, we arrive at C^0 and D^0 . This process can be illustrated as follows:

$$\begin{aligned}
 C^n &= D^{n-1} \oplus C^{n-1} \\
 &\quad \Downarrow \\
 &(D^{n-2} \oplus C^{n-2}) \\
 &\quad \Downarrow \\
 &(D^{n-3} \oplus C^{n-3}) \\
 &\quad \Downarrow \\
 &\dots\dots\dots \\
 &\quad \Downarrow \\
 &(D^0 \oplus C^0)
 \end{aligned}$$

Consequently, it is clear that the sequence of the control point C^n can be represented by wavelet coefficients:

$$\begin{aligned}
 C^n &\implies C^{n-1}, C^{n-2}, \dots, C^1, C^0, \\
 &\quad D^{n-1}, D^{n-2}, \dots, D^1, D^0.
 \end{aligned}$$

Finally, we obtain

$$C^n = D^{n-1} \oplus D^{n-2} \oplus \dots \oplus D^1 \oplus D^0 \oplus C^0.$$

The curve S^n can be decomposed as

$$S^n \implies S^{n-1}, S^{n-2}, \dots, S^1, S^0.$$

An example of this procedure is shown in Fig. 11.2, here we know $n = 4$. Therefore, four wavelet decomposition layers occur. Owing to Fig. 11.2, we can further find the relationship between the sub-curves and the control points as well as the relationship between the wavelet coefficients and the sub-curves at the different layers. In this example, we obtain

$$\begin{aligned}
 C^4 &\implies C^3, C^2, C^1, C^0, \\
 &\quad D^3, D^2, D^1, D^0 \\
 S^4 &\implies S^3, S^2, S^1, S^0
 \end{aligned}$$

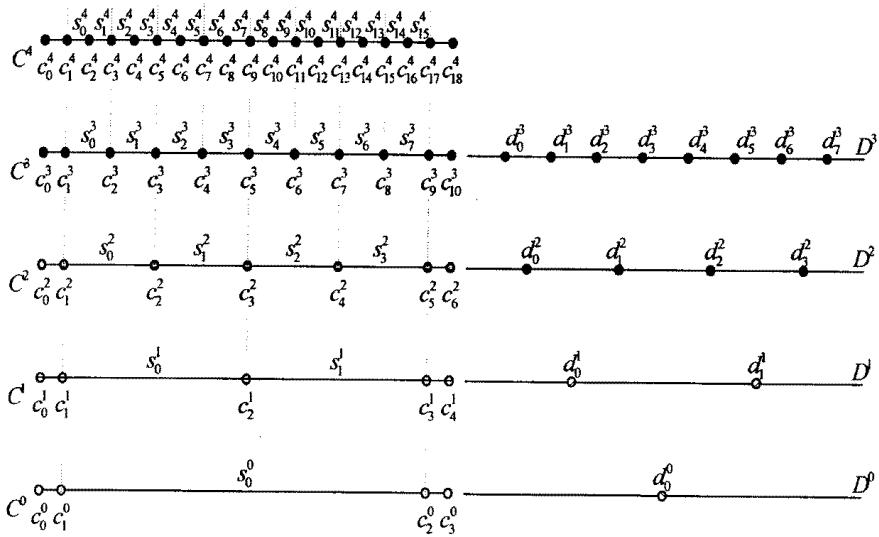


Fig. 11.2 Decomposition of cubic B-spline curve.

and

$$C^4 = D^3 \oplus D^2 \oplus D^1 \oplus D^0 \oplus C^0.$$

where the highest frequency in C^3 must be greater than that in C^2 . In turn, the highest frequency in C^2 must be greater than that in C^1 , etc. In general, we have

$$F_{max}(C^k) > F_{max}(C^{k-1})$$

where $F_{max}(C^k)$ denotes the highest frequency in C^k . The same thing appears in D^i and S^i . Therefore,

$$F_{max}(D^k) > F_{max}(D^{k-1}), \quad F_{max}(S^k) > F_{max}(S^{k-1}).$$

If we consider the curve to be a 1-D signal, this implies that S^k (or C^k or D^k) contains more particulars than S^{k-1} (or C^{k-1} or D^{k-1}) has.

(3) Objective processing:

In the stage of wavelet transform, a curve is decomposed to several layers, and the details of the curve at each layer can be described by wavelet coefficients. Therefore, by adequately choosing and processing these wavelet

coefficients, the different objective processing of the Chinese character can be performed. In this book, compression of characters, arbitrary scale of the typeface and generation of Chinese type styles are introduced using the cubic B-spline wavelet transform. In the rest of this chapter, we emphasize these objective operations and implementation.

11.1 Compression of Chinese Character

Compression of Chinese character is a very important aspect in Chinese information processing with computers. It can benefit reducing the amount of character storage memory, speeding up the processing, decreasing the cost, etc.

In the first three stages in Fig. 11.1, each Chinese character is converted to several contours (closed curves) by the process of extracting strokes and detecting edges, and the individual curve is estimated by a cubic B-spline expression, which can be represented by a sequence of control points. The control points of B-spline curve are transformed into the different layers of details $D^{n-1}, D^{n-2}, \dots, D^1, D^0$ plus C^0 or the different layers of control points $C^{n-1}, C^{n-2}, \dots, C^1, C^0$ by wavelet transform. Therefore, each Chinese character can be described with wavelet transform coefficients (or control points) at different decomposition layers. According to the characteristics of wavelet transform coefficients (or control points), two approaches of compressing Chinese character are presented in this section, namely, (1) global approach, and (2) local approach.

11.1.1 Algorithm 1 (Global Approach)

The method of this algorithm is similar to image compression by wavelet transform. The basic idea is that, under certain accuracy or error range, we consider the entire curve to delete some details D^i and reserve less coefficients on the basis of the influence of different layer details on the curve. More precisely, in the result of the wavelet decomposition, a threshold J is chosen, so that the certain accuracy is satisfied, and thereafter, based on this threshold, some wavelet transform coefficients are removed, and some are kept. This can be described as follows:

$$D^n = \underbrace{D^{n-1} \oplus D^{n-2} \oplus \dots \oplus D^{J+1}}_{Deleted} \oplus \underbrace{D^J \oplus D^{J-1} \oplus \dots \oplus D^1 \oplus D^0 \oplus C^0}_{Kept}$$

$$\approx \underbrace{D^J \oplus D^{J-1} \oplus \cdots \oplus D^1 \oplus D^0 \oplus C^0}_N.$$

In this algorithm, a method of the non-fixed length code will be used in encoding these remained details as follow:

- Give a quantification level L and the maximum absolute value (\max) in the residual details;
- Calculate the quantified value x of the residual details by the equation: $x = [(L * V_D) / \max + 0.5]$, where V_D stands for the value the residual details, and $[.]$ denotes obtaining the integer value;
- In encoding the quantified value x , define two expressions $l = [\log_2(|x|)]$ and $y = (|x| - 2^l)$ except the value x equating zero. There are two situations to be discussed. If $x = 0$, the code is represented by one byte. If $x < 0$ or $x > 0$, the x code is composed of three parts such as $\underbrace{0, 0, \dots, 0}_{l+1}, \underbrace{1, 0, (y)_2}_l$. Here, the first part shows the code unique feature, and the second part $flag$ is positive/negative flag, and the final part is a binary value with l representing the quantified value x .

Quantified value	Positive value code	Negative value code
0	1	1
1	010	011
2	00100	00110
3	00101	00111
4	0001000	0001100
5	0001001	0001101
.....
8	000010000	000011000
.....
15	000010111	000011111
.....
x	$\underbrace{0, 0, \dots, 0}_{l+1}, \underbrace{1, 0, (y)_2}_l$	$\underbrace{0, 0, \dots, 0}_{l+1}, \underbrace{1, 1, (y)_2}_l$

Table 11.1 Non-fixed length code

As the encoding method mentioned above, some examples are listed in

Table 11.1. At the same time, the compressed data in Algorithm record information as follows:

- Firstly, record the length of the original control points and the number of decomposition layers N ;
- Then, store the quantification level L and the value max ;
- Finally, record the code of the details at different layers and the low-frequency components of the curve C^0 .

11.1.2 Algorithm 2 (Local Approach)

The second algorithm is a specific one for the curves. The control points of a curve are decomposed to $C^{n-1}, C^{n-2}, \dots, C^0$ by the wavelet transform, which approximate the original curve S^n . Those points at the different decomposition layers reflect the characteristics of the sub-curve approximation. The control points in C^k are more close to the original curve than these in C^{k-1} . In a cubic B-spline expression, four control points describe a sub-curve. Suppose that at the j^{th} layer, s_i^j is a sub-curve, which is determined by four control points $c_i^j, c_{i+1}^j, c_{i+2}^j, c_{i+3}^j$ ($0 \leq j \leq (n-1), 0 \leq i \leq 2^j$), the sub-curve s_i^j corresponds to exactly two (more detailed) curves s_{2i}^{j+1} and s_{2i+1}^{j+1} at the $(j+1)^{th}$ layer. Consequently, in this case, the number of control points are added up to 5 from 4. In Fig. 11.2, the sub-curve s_0^2 at the second layer corresponds to two curves both s_0^3 and s_1^3 at the third layer, i.e., the curve s_0^2 is the approximation of two curves both s_0^3 and s_1^3 . As the layers increase, the approximation to the original curve become better. Hence, under a certain error, the entire original curve can be estimated by many sub-curves established by control points separately at the different layers. The steps of algorithm are listed as follows:

- Step-1** Given an error, the approximation process starts from the lowest layer;
- Step-2** At the j^{th} layer, the error between the approximated sub-curves and the original ones is checked at each segment;
- Step-3** If the error is less than the given error in a certain segment, then the control points of this segment remain, otherwise, the control points at the $(j+1)^{th}$ layer are chosen, and we repeat the second step, and so on, until the whole curve can be described with control points at the different layers.

In this algorithm, the compressed data are organized as follows:

- The first byte records the length of curve N , i.e. how many segments (sub-curves) on the entire curve have been divided in the approximation.
- The successive bits are used to store a position information of the control points at each decomposition layer. For example, the approximated curve described by control points of k^{th} segment at the j^{th} layer, k is the position information. According to incremental change of the position information, the information code of each layer is performed below: (1) To compute the difference value of the position information k ; (2) To code the difference value according to the non-fixed length code (Table 11.1).
- The last few bits are utilized to quantize and code the control points of each decomposition layer.

11.1.3 Experiments

The process of our experiment are listed as follows:

- (1) To extract the contours (closed curves) from the bitmap image of a Chinese character.
- (2) To check the number of coordinate points (or control points). The number of points is usually 2 to the power n , exactly, curve $S^n(t)$ should have $2^n + 3$ control points. When the control points are less than $2^n + 3$, the coordinate points must be extended as well the control points. The performance of the extension can be periodical spinning out or adding zeros or re-sampling.
- (3) To convert each contour into a cubic B-spline curve.
- (4) To apply the wavelet transform to control points C^n , and the wavelet transform coefficients are produced.
- (5) To choose adequately coefficients according to the above algorithms.
- (6) To organize compression data by the above approaches.

In this experiment, 120×120 bitmap images of characters (1800 bytes) with four typeface fonts (Song, Fang Song, Kai and Hei) are chosen. In Algorithm 1, the quantification level L mainly affects compression ratio. In the second algorithm, the given error also have enormous effect on compression ratio. Table 11.2 shows the compression results of Algorithms 1

Font (1800)	CR(Algorithm 1)	CR(Algorithm 1)	CR(Algorithm 2)
	$L = 32$	$L = 64$	$\varepsilon = 2$
Songti	6.59	6.12	8.45
F-songti	7.28	6.41	9.20
Kaishu	5.46	4.85	7.69
Black	6.29	5.47	7.79
Average	6.41	5.71	8.28

Font (1800)	CR(Algorithm 2)	Method of white and black	Method of outline link-code
	$\varepsilon = 1$		
Songti	5.96	2.25	3.50
F-songti	6.45	1.93	2.61
Kaishu	5.90	2.03	3.74
Black	5.69	2.67	4.83
Average	6.00	2.22	3.67

Table 11.2 Results of **Algorithms 1** and **2** as well as the traditional methods. Note CR=Compression Ratio.

and 2 as well as the traditional methods. In this table, the quantization level in the first algorithm are 32 or 64. The given error range in the second algorithm are 1 or 2 pixels. Examples of the compression of Chinese typeface are illustrated in Figs. 11.3 and 11.4. The compression result using the first algorithm is shown in Fig. 11.3, and that using the second algorithm is displayed in Fig. 11.4.

11.2 Enlargement of Type Size with Arbitrary Scale Based on Wavelet Transform

A font has a particular size and style of type. The size modification of Chinese character, including the reduction and enlargement, is one of the most useful technologies in Chinese character processing systems. Compared with the font zooming out, the font enlargement is more difficult to be realized. Recently, some effective methods, such as the homogeneous coordinate, the logical equation interpolation, etc., have been proposed to deal with the enlargement of type size of Chinese character with arbitrary (smooth) scale. This section aims at a new method for the font enlargement

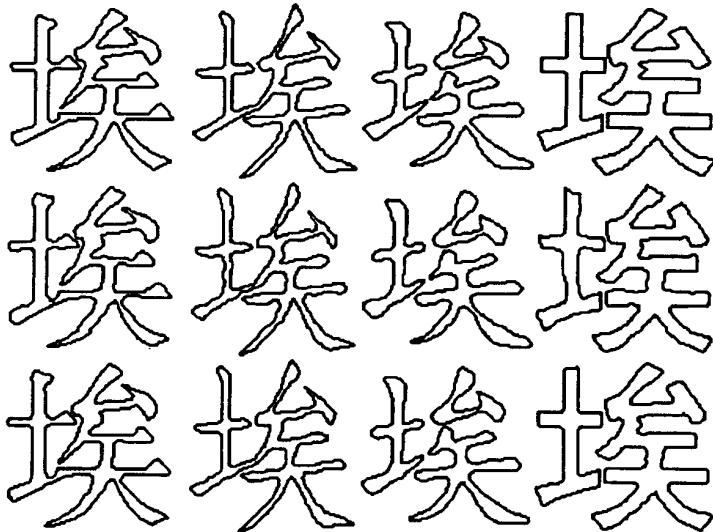


Fig. 11.3 The compressed results using the first algorithm: The first row is an original Chinese character with four fonts; The second row shows the compression result with the quantization level $L = 32$; The third one is that result with quantization level $L = 64$.

with arbitrary scale. In this way, a bitmap image is converted to one or more contours, then, the arbitrary scale enlargement is performed to these contours based on wavelet transform.

11.2.1 Algorithms

In practice application, when a curve (line) is displayed or printed on the output device, it is formed by a finite amount of unit points such as pixel or dot. For a certain curve, if an amount of unit points is fixed, the higher is the resolution of the output device, the smaller the curve is displayed or printed, vice versa. Thus, under the same resolution, for zooming in a curve, it is necessary to increase the amount of unit points by adding some pixels in the original curve. Several techniques, such as interpolation and complement can be used to perform this task.

According to wavelet transform theory, we know that the wavelet reconstruction is performed by interpolation. Is it possible to enlarge a curve by a wavelet reconstruction? The answer is definite. In order to do so, we

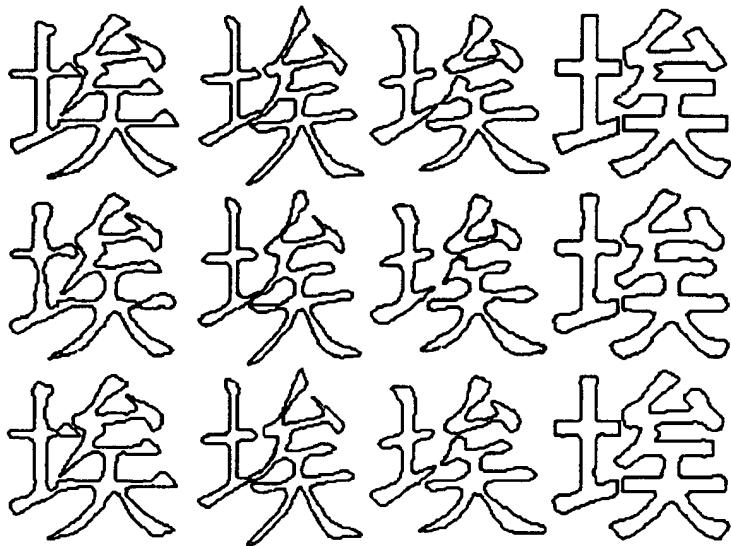


Fig. 11.4 The compressed results using the second algorithm: The first row is an original Chinese character with four fonts; The second row shows the compression result one with the error $\varepsilon = 2$; The third one is that result with the error $\varepsilon = 1$.

should keep the distance between two neighboring control points on a curve unchangeable. Otherwise, the length of the reconstructed curve will not be changed.

Let C^n denote the control points of an original curve, and C^{n+1} be that of its enlarged curve. Based on the interpolation theory, we have known that

$$C^{n+1} = P^{n+1} \cdot C^n + Q^{n+1} \cdot D^n \quad (11.5)$$

where P^{n+1} and Q^{n+1} are two filters. The particulars, D^n , contained on the original curve can be viewed as zero, therefore, the control points of the new curve is

$$C^{n+1} = P^{n+1} \cdot C^n. \quad (11.6)$$

Eq.(11.6) can be considered to be a procedure of the interpolation by wavelet reconstruction based on the cubic B-spline function. According to the wavelet reconstruction, the control points on a curve are doubled

after the interpolation. In this way, an original curve can be enlarged at any 2^j times, where j is a positive integer. It means that the scale of the enlargement is always multiple of 2. A question is that how the Chinese type size can be enlarged with arbitrary scale? We solve this question in this section, where two algorithms are proposed:

- The first one is used to perform the enlargement of a curve with arbitrary scale by a cubic B-spline wavelet transform.
- The second algorithm is employed to accomplish the arbitrary enlargement of Chinese type size using the first algorithm associated with other techniques.

(1) Algorithm 1:

Let S^j and S^{j+1} represent two enlarged curves, the sizes of which are 2^j and 2^{j+1} times of the original curve, respectively. Our object is to find a curve S^{j+t} ($0 \leq t \leq 1$) between S^j and S^{j+1} as shown in Fig. 11.5. It can

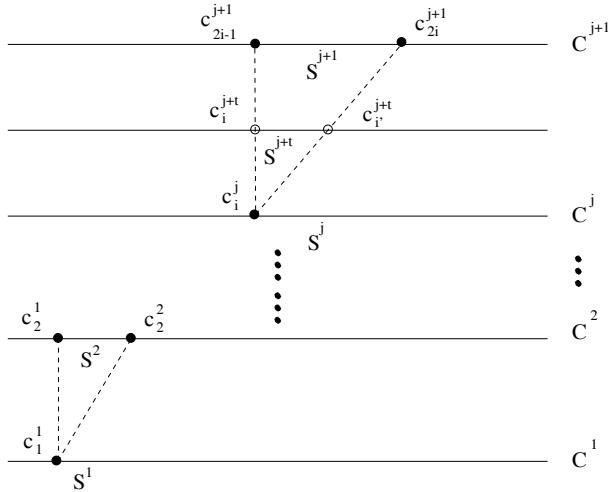


Fig. 11.5 Curve S^{j+t} is produced by the different wavelet transform layers.

be seen that, after wavelet reconstruction, a control point c_i^j on the curve S^j ($1 \leq i \leq 2^j + 1$) corresponds to two control points c_{2i-1}^{j+1} and c_{2i}^{j+1} on the curve S^{j+1} , thus, one point becomes two. Note that, in this section, the control points can be considered to be the coordinate points. For example, in Fig. 11.5, c_1^1 on S^1 corresponds to c_1^2 and c_2^2 on S^2 . If we draw two lines

from c_i^j to c_{2i-1}^{j+1} and c_{2i}^{j+1} respectively, we can obtain two cross points c_i^{j+t} and $c_{i'}^{j+t}$ on the curve S^{j+t} , which can be computed with line equations as follows:

$$\begin{aligned} c_i^{j+t} &= c_i^j + t \cdot (c_{2i-1}^{j+1} - c_i^j) \\ &= (1-t) \cdot c_i^j + t \cdot c_{2i-1}^{j+1} \end{aligned} \quad (11.7)$$

$$\begin{aligned} c_{i'}^{j+t} &= c_i^j + t \cdot (c_{2i}^{j+1} - c_i^j) \\ &= (1-t) \cdot c_i^j + t \cdot c_{2i}^{j+1} \end{aligned} \quad (11.8)$$

Note that, we should consider a special case, i.e. the distance between the control points c_i^{j+t} and $c_{i'}^{j+t}$ is less than a unit point (a pixel). In this situation, we should combine these two control points to a single one.

Consequently, the control points C^{j+t} on the curve S^{j+t} can be obtained by the control points C^j and C^{j+1} on the curves S^j and S^{j+1} , and can be written by

$$\begin{aligned} C^{j+t} &= C^j + (C^{j+1} - C^j) \cdot t \\ &= (1-t) \cdot C^j + t \cdot C^{j+1}. \end{aligned} \quad (11.9)$$

It is clear that two special cases also should be considered, namely: (1) $S^{j+t} \rightarrow S^j$ when $t \rightarrow 0$; (2) by contrast, $S^{j+t} \rightarrow S^{j+1}$ when $t \rightarrow 1$.

Finally, we have the following algorithm for the enlargement of a curve with arbitrary scale.

Algorithm 1 (Enlargement of Curves)

Step 1: For a given scale d , to compute j and t by $2^j \leq d \leq 2^{j+1}$ and $t = (d - 2^j)/2^j$.

Step 2: According to formula (11.6), to enlarge the original curve to 2^j and 2^{j+1} times with the wavelet reconstruction based on the cubic B-spline function.

Step 3: To obtain C^{j+t} by formula (11.9).

(2) Algorithm 2:

A Chinese character is composed of a set of contour (closed curves). Therefore, the enlargement of the entire character can be performed by treating each closed curves using Algorithm 1 separately. In practice, it is found that the jaggy and staircase phenomenon emerge on the contour when a curve is directly enlarged without any preprocessing. An example

of the character with such jaggy and staircase phenomenon can be fund in Fig. 11.6(e). To avoid this phenomenon, it is necessary to remove the jaggy and staircase before we enlarge a curve. We have known that the decomposition transform functions as a filter, which can eliminate the jaggy and staircase phenomenon. For this reason, the wavelet decomposition with the cubic B-spline function is employed to perform this task.

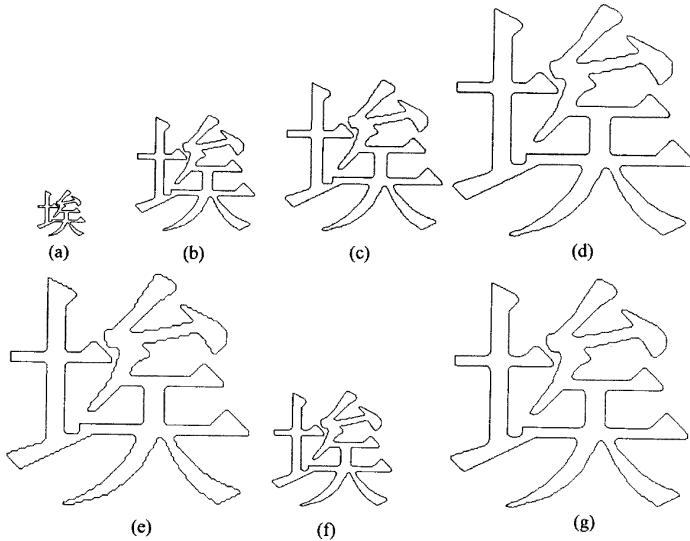


Fig. 11.6 Results and comparison of enlarged fonts. (a) original font, (b) enlarged two times, (c) enlarged three times, (d) enlarged four times, (e) directly enlarged four times without step 5 in Algorithm 2, (f) enlarged two times by traditional interpolation and (g) enlarged four times by traditional interpolation.

As a conclusion, we have the following algorithm for enlargement of a Chinese type size with arbitrary scale based on the cubic B-spline wavelet transform.

Algorithm 2 (Enlargement of Chinese type Size)

Step 1: To divide a Chinese character into a set of connected domains in terms of some regular techniques of image processing.

Step 2: To extract the edge of each connected domain by the edge detection algorithm, which have been presented in the previous chapter

in this book.

Step 3: To trace the contour of each edge, as a result, a set of closed curves (contours) can be obtained for the Chinese character.

Step 4: To perform the pre-processing of each curve with wavelet decomposition, firstly, to decompose C^n into C^{n-1} by

$$C^n = D^{n-1} \oplus C^{n-1}.$$

Let $D^{n-1}=0$, which means we remove the particulars (high-frequency components). Then, the preprocessing can be implemented by reconstructing C^n with formula (11.6).

Step 5: To enlarge each curve with a given scale by Algorithm 1, which has been described in this section.

Step 6: After enlarging the contours, to fill the domains, which are enclosed by the contours.

Step 7: To repeat Steps 4 - 6, until all contours have been processed.

11.2.2 Experiments

In our experiments, 100×100 bitmap images of Chinese characters are selected to be processed using both Algorithm 2 and a traditional algorithm, which is an interpolation algorithm with Bezier curve equation, respectively. The original image of a character is scanned into a computer system, and its contour can be extracted by the edge detection technique and contour tracing algorithm. The contour points of an original curve are chosen as its control points C^n . The results of our experiments are illustrated in Fig. 11.6, in that (a) is an original font; (b)~(d) are the results, where the original font is enlarged by two ~ four times with Algorithm 2 respectively; (e) is the font which is directly enlarged from the original one by four times with the Algorithm 2 without the preprocessing (Step 4); (f)~(g) are the resulting fonts, where the original font is enlarged by two and four times respectively with the traditional interpolation. Clearly, the curves enlarged by Algorithm 2 is smoother than the curves enlarged by the traditional interpolation. On the other hand, the distortion produced by Algorithm 2 is less than that produced by the traditional method, especially for the corner or short line in the characters.

Fig. 11.7 shows other experiments, where the handwritten Chinese characters are enlarged by the algorithm described above in this section.

In this section, we apply the cubic B-spline wavelet transform to im-

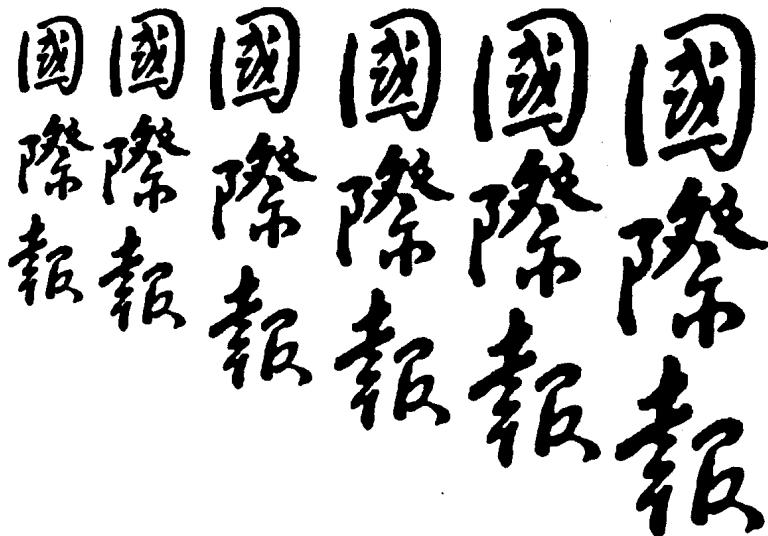


Fig. 11.7 Chinese handwritings are enlarged with arbitrary scales by the algorithms described in this section.

plement the enlargement of Chinese characters with arbitrary scales. Its advantages can be concluded below:

(1) The algorithm is very simple and easy to be implemented. (2) The computational complexity of the algorithm is $O(N)$. (3) The distortion of amplified font is very light because it is only related to that of the original image itself.

11.3 Generation of Chinese Type Style Based on Wavelet Transform

With the increasing requirement of the Chinese computing processed by the computer including the Chinese character recognition and Chinese press, the generation of type style has become another important application of the Chinese character processing systems. In this section, a new approach, which applies the cubic B-spline wavelet transform to the generation of Chinese type style, will be discussed.

Usually, a new type style can be derived by modifying the structural fea-

tures of an existing font. For instance, changing the length and the width of its strokes can produce a new type style. The structural features of a font correspond to the details at various wavelet decomposition layers, when each contour of a character is decomposed by a wavelet transform. With the cubic B-spline wavelet transform, which is viewed as a "mathematical microscope", the details of a font at different layers can be effectively extracted. For a given font, modifying its structural features can be performed by changing its details in some layers, or composing its details with that of others. In this way, three steps are involved:

- the re-sample of the spline curves of the Chinese character,
- the wavelet transform of each curve,
- the modification or composition of the original curves to produce a new type style.

Two approaches are proposed in this section:

- The first one is to generate a new font by modifying the structural features of an original one.
- The second one is to generate a new font by composing the structural features of several existing ones.

11.3.1 *Modification*

Several algorithms, which can perform the modifications of the type styles, will be discussed in this sub-section.

Algorithm 1

When a cubic B-spline curve is utilized to fit the contour of a font with M coordinate points, the number of the control points should be $M + 2$. Further, if $M \neq 2^n + 1$, the number of the control points of a cubic B-spline curve needs to be extended to $2^n + 3$. On the other hand, when two or more fonts are composed to form a new one, it is necessary to ensure that they have the same topological structure, e.g., their length, direction, start and end point. In order to extend the control points with the required length and hold the structure unchangeable, it is necessary to re-sample the B-spline curve as follows:

Step-1 Given M coordinate points of a curve, to find a cubic B-spline

curve $F^n(u)$, which passes through these coordinate points,

$$F^n(u) = (f_1^n(u), f_2^n(u), \dots, f_M^n(u)),$$

where $f_i^n(u)$ is the i^{th} B-spline sub-curve, it can be written as:

$$f_i^n(u) = \sum_{j=1}^4 c_{i+j-2}^n N_{j,3}(u) \quad i = 1, 2, \dots, M, u \in [0, 1] \quad (11.10)$$

where $N_{j,3}(u)$ are the base functions of the cubic B-spline. c_i^n are the control points of the cubic B-spline curve, corresponding to the coordinate points of the curve.

Step-2 To select the decomposition levels n , in accordance with the length of the curve, L , satisfying $M \leq L$ and $L = 2^n - 1$.

Step-3 To re-sample $F^n(u)$ with a new sample interval of $\Delta t = \frac{M}{2^n - 1}$ producing a set of new coordinate points $Q'_i(i = 1, 2, \dots, M')$, $M' = L + 1$.

Step-4 Owing to formula (11.10), we can obtain the extended control points $c_{i'-1}^n(i' = 1, 2, \dots, M')$ with the length of $2^n + 3$. we have

$$F^n(u) = (f_1^n(u), f_2^n(u), \dots, f_{M'}^n(u)).$$

The wavelet base we used here is the cubic B-spline wavelet, which can interpolate a function with equal interval. Further, since the contour of the character is a closed curve, its start point is the same as its end point, that is, $c_0^n = c_1^n = c_{M'+2}^n = c_{M'+3}^n$. Therefore, only $2^n - 1$ points in curve $F^n(u)$ need to be re-sampled.

Based on the re-sampled B-spline curve, the wavelet transform can be employed to modify the Chinese type styles. We have three stages to do so:

Step-1 To apply wavelet transform to the re-sampled B-spline curve, which has the length of $2^n + 3$.

Step-2 To process the details at some wavelet decomposition layers.

Step-3 To reconstruct a new type style by the wavelet reconstruction algorithm.

Two examples are shown in Fig. 11.8 and Fig. 11.9, in which four different fonts of a Chinese character are decomposed into seven layers by wavelet transform.

In Fig. 11.8, the modifications are conducted as follows:

- The detail D^{n-1} is removed from the first layer;
- The details D^{n-1} and D^{n-2} are deleted from the second layer;
- The details D^{n-1} , D^{n-2} and D^{n-3} are discarded from the third layer;
-
- All of the details $D^{n-1}, D^{n-2}, \dots, D^{n-7}$ are no longer kept at the seventh layer.

The results of the above modifications are illustrated from left to right in Fig. 11.8.

The modifications in Fig. 11.9 are carried out below:

- Only the detail D^{n-1} is removed from the first layer;
- Only the detail D^{n-2} is deleted from the second layer;
- Only the detail D^{n-3} is discarded from the third layer;
-
- Only the detail D^{n-7} is no longer kept at the seventh layer.

The results of the above modification are illustrated from left to right in Fig. 11.9. From these, it is easy to see that the details contained at different layers affect the features of the type style variously. The deeper is the layer, the greater the effect of the details will be.

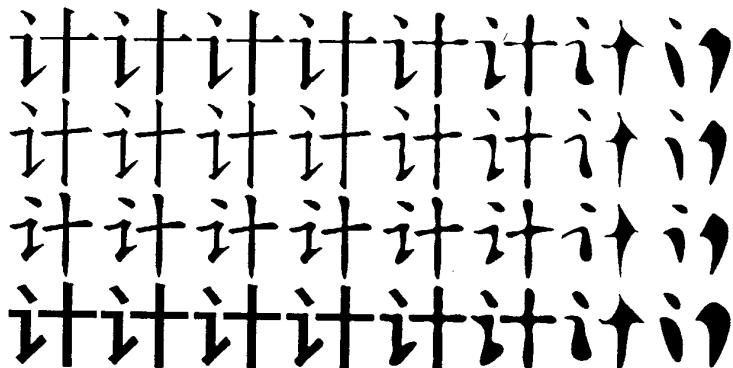


Fig. 11.8 Example 1.

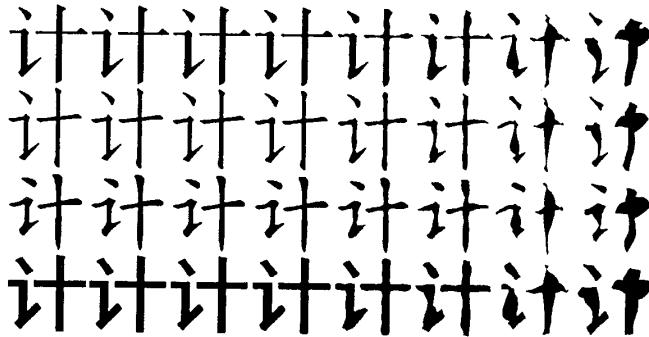


Fig. 11.9 Example 2.

Algorithm 2 (Smoothing of Curves)

Suppose a objective curve $r(t)$ contains $m = 2^j + 3$ control points: $C^j = (c_1^j, c_2^j, \dots, c_m^j)$. According to the least-squared error, an approximated curve with $m' = 2^{j'} + 3$ control points can be obtained ($j' < j$ is a non-negative integer). From the previous section, we have known

$$C^{n-1} = A^n \cdot C^n. \quad (11.11)$$

Therefore, the control points C' of the approximated curve can be expressed as

$$C' = A^{j'+1} A^{j'+2} \cdots A^j C^j. \quad (11.12)$$

A remarkable property of the multiresolution curve is its discrete nature, i.e., we can use K control points to construct approximated curve efficiently at the j^{th} layer. In this way, K can be any of the integers 4, 5, 7, 11 or $2^j + 3$, and j can be any integer. In practice, we can also define another non-integer-layer curve $\gamma^{j+\mu}(t)$, $\mu \in R$ and $0 \leq \mu \leq 1$, which can be achieved by two curves, $\gamma^j(t)$ and $\gamma^{j+1}(t)$, at the neighboring integer layers, as below:

$$\begin{aligned} \gamma^{j+\mu}(t) &= (1 - \mu)\gamma^j(t) + \mu\gamma^{j+1}(t) \\ &= (1 - \mu)\Phi^j(t)C^j + \mu\Phi^{j+1}(t)C^{j+1} \end{aligned} \quad (11.13)$$

This non-integer-layer curve can benefit to smooth the curve at any continuous scales. We can continuously edit any segments on this curve, which

vary curve from the smoothest form (with only four control points) to its highest resolution (with $m > 4$ control points).

Suppose curve $\gamma(t)$ is to be smoothed, the concrete procedure to smooth it is presented below:

Step 1: To decompose the curve $\gamma(t)$ into a sequence of basic control points (including four basic points) and a group of details using wavelet transform.

Step 2: To choose the corresponding details of wavelet transform coefficients, so that the requirement of the smoothness is satisfied (Note that, the less is the value of j , the better the smoothness is). To fix some chosen details from one layer or more layers to be zero, generating new details D'^j .

Step 3: To obtain a new curve $\gamma'(t)$ by wavelet reconstruction using changed details D'^j and C^0 .

Algorithm 3 (Edition of the Shape and Details of the Curve)

For this algorithm, we will only give some conclusion. The detailed discussion and inference can be found in [stollnitz et al., 1996]. A curve, which has C^J control points, can be described by wavelet coefficients C^0 and D^0, D^1, \dots, D^{J-1} , or the control points C^0, C^1, \dots, C^{J-1} at the different layers. Two methods can be used to edit such a curve:

- In the first method, the global shape of the curve is changed, while the details of it are kept.
- The second method is just opposite the above method. The details of the curve are only altered, while the elementary shape of curve is nearly the same as the original one.

We will discuss these methods as follows.

Method 1:

The basic idea of this method is that: firstly some of C^j at some layers are modified, then the original details $D^0, D^1, \dots, D^{J-1}, (0 < j < J)$ are added to the modified C^j , finally the global shape of the whole curve can be changed.

We shall describe the modified control points by writing: $\hat{C}^j = C^j + \Delta C^j$, where C^j and \hat{C}^j denote the original points and modified ones respectively. The ΔC^j indicates the difference between them. At the same

time, some modification appears in C^J through reconstruction, i.e.,

$$\hat{C}^J = C^J + \Delta C^J. \quad (11.14)$$

Here the j value mainly affects the above modification. Precisely, the less is j value, the wider the influenced range of the control points is, furthermore, the stronger the effect to the whole curve is. Otherwise, the closer to J the j value is, the more narrow the influenced range of the control points is, at the same time, the less the effect to the whole curve is. This method can be extended to the non-integer layers. Suppose a non-integer-layer curve $\gamma^{j+\mu}$ satisfies (11.13), and it contains control points $C^{j+\mu}$, we have

$$\gamma^{j+\mu}(t) = \Phi^{j+1}(t) \cdot C^{j+\mu}. \quad (11.15)$$

If we modify a particular point $c_i^{j+\mu}$ in curve $C^{j+\mu}$, the positions of the neighboring points of $c_i^{j+\mu}$ will be changed. The size of the influenced range and degree is inverse proportion to μ , namely:

- If μ closes to zero, then all control points at the $j + \mu$ layer will be simultaneously moved. In other words, every points at the j layer will be edited.
- If μ approximates 1, the neighboring points will not be moved. Only a single point at the $j + 1$ layer will be edited.

Let the $\Delta C^{j+\mu}$ be the values of the modifications at the $j + \mu$ non-integer layer. Here the $\Delta c_i^{j+\mu}$, which is one of the i^{th} point in $\Delta C^{j+\mu}$, is selected by the user. The $\Delta C^{j+\mu}$ can be considered as two parts, namely, ΔC^j and ΔD^j at the j^{th} layer. We can define: $\Delta D^j = B^{j+1} \cdot \Delta C^{j+1}$. Therefore, the modified $\Delta C^{j+\mu}$ of the whole curve is represented as follows:

$$\Delta C^J = P^J P^{J-1} \dots P^{j+2} (P^{j+1} \Delta C^j + Q^{j+1} \Delta D^j). \quad (11.16)$$

Furthermore, from (11.13), (11.15) and $\Phi^{j-1}(u) = \Phi^j(u) \cdot P^j$.

We obtain

$$\Delta C^{j+\mu} = (1 - \mu) P^{j+1} \cdot \Delta C^j + \mu \Delta C^{j+1}. \quad (11.17)$$

Considering the filter equation

$$C^j = P^j \cdot C^{j-1} + Q^j \cdot D^{j-1},$$

where P^j and Q^j are metrics of the filters. By replacing ΔC^{j+1} with both ΔC^j and ΔD^j , we can rewrite (11.17) as

$$\Delta C^{j+\mu} = P^{j+1} \cdot \Delta C^j + \mu Q^{j+1} \Delta D^j. \quad (11.18)$$

From (11.18), it is clear that the modified control points at the non-integer layer can be described by two components: (1) the modified control points at the lower layer, and (2) the modified details at the lower layer. The entire modified curve can be determined by both (11.16) and (11.14). If $\Delta C^{j+\mu}$ is known, it is quite difficult to calculate directly ΔC^j and ΔD^j from (11.18). In practice, we can denote: $\Delta C^{j+\mu} := (0, \dots, \Delta c_i^{j+\mu}, 0, \dots, 0)^T$, and define

$$\begin{aligned}\Delta C^j &= (1 - \mu) A^{j+1} \cdot \Delta C^{j+\mu} \\ \Delta D^j &= \mu B^{j+1} \cdot \Delta C^{j+\mu},\end{aligned} \quad (11.19)$$

The main editing procedure can be implemented as follows:

Step 1: To determine editing vector $C^{j+\mu} = (0, \dots, \Delta c_i^{j+\mu}, 0, \dots, 0)^T$.

Step 2: To compute corresponding ΔC^j and ΔD^j , according to (11.19).

Step 3: To obtain the bias of the curve, according to (11.16).

Step 4: To compose the new curve, according to (11.14).

Method 2:

This method is just the opposite process to that described above. In Method 1, the global shape of the curve is altered, but the detail of it can be kept. However, in this method, the details of the curve are changed, and the basic shape of it is nearly the same as the original curve. The basic idea of this method is that, we only modify the details $D^j, D^{j+1}, \dots, D^{J-1}$, ($0 < j < J$), and preserve the low-resolution components of the curve, C^0, C^1, \dots, C^j , ($0 < j < J$). Here the original details $D^j, D^{j+1}, \dots, D^{J-1}$, ($0 < j < J$) are replaced by a group of new details $\hat{D}^j, \hat{D}^{j+1}, \dots, \hat{D}^{J-1}$, ($0 < j < J$). This editing method needs to establish a library, which stores various content about curve details. They are samples of the standard curves such as folding line, spiral and snake line, etc. When the object curve is to be edited, we can perform the edition by the following steps:

Step 1: To choose some standard samples from the library, in accordance with the requirement of the edition.

Step 2: To decompose these curves by the wavelet transform.

Step 3: To extract the details, D_n^k , from one or several layers. They can represent the details of the shape efficiently, and we have $\hat{D}_i^j = \zeta(D_k^k, l, k)$.

Step 4: To add these details to the original low-frequency coefficients, to obtain the designed curve, i.e., $C^J = P^J \cdot C^{J-1} + Q^J \cdot \hat{D}^{J-1}$

180 × 80 bitmap images of Chinese characters are chosen as the basic samples in this experiment. The coordinate x and y in the contour of a character is considered to be the control points, i.e., C_x^n and C_y^n . According to the above method, we can process Chinese character by the local edition. Figs. 11.10 and 11.11 show the experimental results.



Fig. 11.10 Results of modifying the typeface of Chinese character “ji”. (a) original character, where 1, 2 and 3 describe three closed curves, which can be transformed with wavelet transform, (b),(c),(d) the result of the local edition.



Fig. 11.11 Results of modifying the typeface of Chinese character “shu”. (a) original character, where 1, 2 and 3 describe three closed curves, which can be transformed with wavelet transform, (b),(c),(d) the result of the local edition.

In Figs. 11.10 and 11.11, each character consists of three contours, which are labeled by the numbers 1, 2 and 3 as shown in the figure. The result of the local edition is displayed in Fig. 11.10 using the following procedure:

- For label one, the point c_i^3 at the third layer is tensioned in both x and y directions;
- For label two, the point c_i^4 at the forth layer is tensioned;
- For label three, the point c_3^7 at the seventh layer is tensioned.

Due to the extent of the modification, there is a little difference among (b),(c) and (d) in Fig. 11.10.

For the Chinese character in Fig. 11.11, the following modification is done:

- For label one, the control point c_i^8 at the eighth layer is tensioned towards the directions of down-left and up-right. So an italic type style is obtained. The result show in Fig. 11.11(b).
- For label two, the high-frequency component (details)will be removed. T he result is shown in Fig. 11.11(c).
- The character is modified in accordance with the above two operations. Fig. 11.11(d) illuminates the result.

11.3.2 Composition

Another way of creating new fonts is to compose the details of two or more existing fonts with different ratios. The new type style, which contains the mixture of details, can be restored by wavelet reconstruction. The composition algorithm based on the cubic B-spline wavelet transform is presented as follows:

Step-1 To divide each Chinese character into many separate connected domains(strokes). Then, extract the edge of each connected domain to produce a contour, and thereafter process each contour by **Step 2~Step 6**.

Step-2 To represent a contour by a cubic B-spline curve with 2^n+3 control points by re-sample, where n is the number of the decomposition levels.

Step-3 To apply wavelet decomposition to control points C^n with the cubic B-spline wavelet transform, so that the detail information D^0, D^1, \dots, D^{n-1} at different layers and the control points C^0 at the lowest layer can be obtained.

Step-4 To compose the details of m ($m \geq 1$) fonts at certain layers, and the detail information can be found as follows:

$$D_{New}^j = \sum_{i=1}^m r_i D_i^j,$$

where D_i^j ($i = 1, 2, \dots, m$) are the font details at the j th layer, and

$r_i (i = 1, 2, \dots, m)$ are ratio factors, $0 \leq r_i \leq 1 (i = 1, 2, \dots, m)$.

- Step-5** To reconstruct a new set of control points C_{New}^n with new details D_{New}^j and a set of control points C^j at the j th layer by wavelet reconstruction. The length of C_{New}^n is $L = 2^n + 3$.
- Step-6** To re-sample the B-spline curve $F_{New}^n(u)$ formed by control points C_{New}^n . It is said to convert the B-spline curve $F_{New}^n(u)$ to a curve (contour) with the length of M again. Here, the sample interval is $\Delta't = \frac{2^n - 1}{N}$.

- Step-7** To fill in all the connected domains surrounded by the contours. The new Chinese type style is completed.

Two examples can be found in Figs. 11.12 and 11.13. Four Chinese type styles are presented in these examples, namely, Song (the first row in Fig. 11.12), Hei (the last row in Fig. 11.12), Fang Song (the first row in Fig. 11.13) and Kai (the last row in Fig. 11.13). The size of the images is 200×200 .

In Fig. 11.12:

- Characters in groups (A) and (B) are obtained by removing the details of Song and Hei at the 1 ~ 4th layers, respectively.
- Ones in (C) and (D) are the resulting fonts, which are composed by the type style of Song and Hei.

In Fig. 11.13:

- Characters in the group (A) are the composition results of three fonts, namely, Fang Song, Kai and Song, with ratios $r_1 = 0.2$, $r_2 = 0.4$, $r_3 = 0.4$, and the control points C^{n-4} at the lowest decomposition layer is in the style of Fang Song.
- Ones in the group (B) are the composition results of the four fonts with ratios $r_1 = 0.1$ (Fang Song), $r_2 = 0.3$ (Kai), $r_3 = 0.3$ (Song), $r_4 = 0.3$ (Hei), and C^{n-4} is in Fang Song.
- The group (C) are the composition results of the four fonts with ratios $r_1 = 0.1$ (Kai), $r_2 = 0.3$ (Fang Song), $r_3 = 0.3$ (Song), $r_4 = 0.3$ (Hei), and C^{n-4} is in the font of Kai.
- The type styles in the group (D) are the composition results of three fonts with ratios $r_1 = 0.2$ (Kai), $r_2 = 0.4$ (Fang Song), $r_3 = 0.4$ (Song), C^{n-4} is in the type style of Kai.

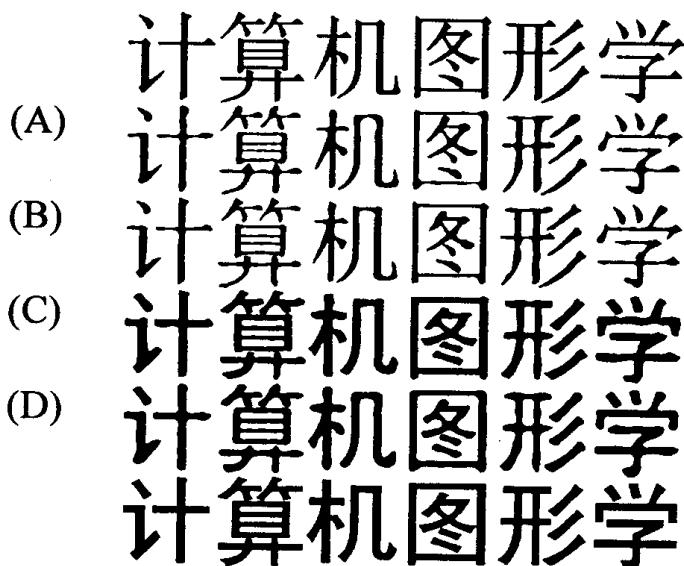


Fig. 11.12 Results of combining two typefaces. The first row and the last row are original typefaces.

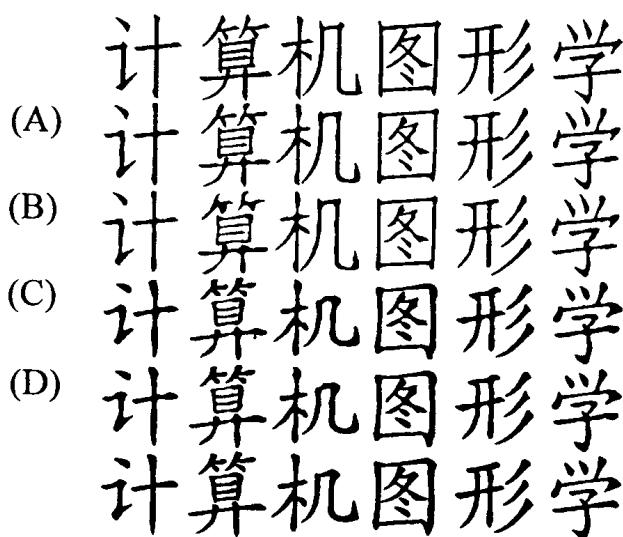


Fig. 11.13 Results of combining several typefaces. The first row and the last row are original typefaces.

This page intentionally left blank

Chapter 12

Classifier Design Based on Orthogonal Wavelet Series

Conceptually speaking, the methodology of statistical pattern recognition draws on classic Bayesian statistical decision theory. It essentially deals with two categories of problem; one is concerned with the identification and extraction of significant features for describing features, and another with the design of classifiers for discriminating and recognizing patterns. In this chapter, we shall provide an in-depth examination of the classifier design problem. First, we shall present an overview of the fundamentals in pattern classifier design. In so doing, our emphasis will be on *minimum average-loss classifier design* and *minimum error-probability classifier design*. Next, we shall specifically describe and discuss the use of orthogonal wavelet series in classifier design.

12.1 Fundamentals

Without loss of generality, we shall only consider two-class pattern recognition problems. That is to say, the complete set of patterns to be classified, Ω , is composed of patterns of two classes, Ω_1 and Ω_2 . For instance, Ω may contain samples of human cells, and Ω_1 and Ω_2 may correspond to non-cancer and cancer cells, respectively. In other words, $\Omega_1 \cup \Omega_2 = \Omega$, and $\Omega_1 \cap \Omega_2 = \emptyset$. If based on some medical statistics, we observe that the priori probabilities of not having and having cancers in a certain region are P_1 and P_2 , respectively, then we also know that the probabilities of Ω_1 and Ω_2 occurrences much be P_1 and P_2 , and furthermore, $P_1 + P_2 = 1$. In order to classify these samples (e.g., to identify cancer cells), it is a common approach that we select a set of features to form a feature vector, X , which

may be symbolically expressed as follows:

$$X = (X_1, X_2, \dots, X_n) \quad (12.1)$$

where X denotes an n -dimensional random vector that defines the following mapping:

$$(\Omega, \mathcal{F}, P) \xrightarrow{X(\omega)} (\mathbb{R}^n, \mathcal{B}^n, P_{X^{-1}}) \quad (12.2)$$

where Ω and \mathbb{R}^n correspond to the domain and range of X . \mathbb{R}^n is an n -dimensional Euclidean space. \mathcal{F} denotes a σ -field constructed by the subsets of Ω . P denotes a probability measure defined over \mathcal{F} and determined by priori probabilities. \mathcal{B}^n denotes a Borel σ -field in \mathbb{R}^n . $P_{X^{-1}}$ is a probability measure derived from random variable X in feature vector space $(\mathbb{R}^n, \mathcal{B}^n)$; whose probability density function is denoted as $p(x)$, $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$. Suppose that $p(x | j)$, the conditional probability density function of each pattern Ω_j , is given. Thus, we can have the following expression:

$$p(x) = P_1 p(x | 1) + P_2 p(x | 2). \quad (12.3)$$

In what follows, we shall introduce the notion of decision function.

Definition 12.1 Function $d(\omega)$ is called a decision function if and only if it satisfies the following conditions: $d(\omega)$ is defined in domain Ω , and its corresponding values are given in set $\{1, 2\}$. In other words,

$$(\Omega, \mathcal{F}, P) \xrightarrow{d(\omega)} \{1, 2\} \quad (12.4)$$

In addition, $d^{-1}(\{i\}) \in \mathcal{F}$, $i = 1, 2$, where $d^{-1}(\cdot)$ denotes the inverse of $d(\cdot)$ and $\{i\}$ denotes a single-number set.

From Definition 12.1, we note that $d(\omega)$ is a discrete random variable. When $d(\omega) = i$, we say that ω belongs to the i th pattern, Ω_i .

Notice that here we make a decision on whether or not a pattern belongs to a specific class depends on the value of the feature vector. In other words, we have to consider a Borel measurable function $c(x)$ in feature vector space \mathbb{R}^n , whose value is given by $\{1, 2\}$:

$$(\mathbb{R}^n, \mathcal{B}^n) \xrightarrow{c(x)} \{1, 2\}. \quad (12.5)$$

The above-mentioned decision function $d(\omega)$ is merely the composite of random feature vector X and the Borel measurable function $c(x)$, i.e.,

$$d(\omega) = c(X(\omega)). \quad (12.6)$$

When x , the observed sample value of the feature vector for pattern ω , satisfies $c(x) = i$, we can arrive at the conclusion that pattern $\omega \in \Omega_i$. In this respect, we usually also refer to medium function $c(x)$ as a decision function.

It can be noted that since classifications based on $c(x)$ are in essence statistical decisions, they are inevitably subject to statistical errors. In real-life applications, misclassifications can cause damages of varying degrees. For instance, in character recognition, it is sometimes possible to misclassify letter c into letter d and vice versa. In both cases, the damages caused may not seem to be as serious as misclassifying non-cancer cells into cancer ones in cancer diagnosis. The misclassification of cells could make patients devastated, who may spend fortunes on their medication. The situation can be even worse if we misclassify cancer cells into non-cancer ones. In such a case, we may delay necessary medical treatments for the patients. And as a result of that, the patients may lose their lives. Owing to the above considerations, it is important to explicitly define a loss function in order reflect the degree of damage caused by misclassification. Before we formally define a loss function, let us first introduce the notion of a class function as follows:

Definition 12.2 $J(\omega)$ is called a class function if and only if it satisfies the following:

$$(\Omega, \mathcal{F}, P) \xrightarrow{J(\omega)} \{1, 2\} \quad (12.7)$$

where

$$J(\omega) = \begin{cases} 1 & \text{if } \omega \in \Omega_1 \\ 2 & \text{if } \omega \in \Omega_2 \end{cases}$$

$J(\omega)$ is a discrete random variable.

We can now give the formal definition of a loss function, based on those of decision function $d(\omega)$ and class function $J(\omega)$.

Definition 12.3 Let bivariate function $L(i, j)$ be defined as follows:

$$L(i, j) = c_{ij}, \quad i, j \in \{1, 2\}, \quad c_{ij} \in (-\infty, +\infty)$$

If we compose $L(i, j)$ with the random vector made of decision function $d(\omega)$ and class function $J(\omega)$, i.e., $(d(\omega), J(\omega))$, we will have composite random variable $L(d(\omega), J(\omega))$. This variable is referred to as a lose function. c_{ij} is referred to as the lose caused by misclassifying pattern Ω_j into pattern Ω_i .

Since $d(\omega)$ is the composite of medium function $c(x)$ and feature vector $x(\omega)$, we can write the following:

$$L(d(\omega), J(\omega)) = L(c(X(\omega)), J(\omega)). \quad (12.8)$$

In this way, we can also view $L(c(X), J)$ as the composite of function $L(c(\cdot), \cdot)$ and random vector (X, J) . The joint distribution of (x, J) can be given as follows:

$$p(x, j) = P_j p(x | j) \quad j = 1, 2. \quad (12.9)$$

Based on the above interpretation of lose function, it becomes quite convenient to calculate an average lose.

Definition 12.4 The mathematical expectation of $L(c(x), J)$ is referred to as an average lose, which can be written as:

$$\begin{aligned} R &= E[L(c(X), J)] \\ &= \sum_{j=1}^2 \int_{\mathbb{R}^n} L(c(x), J) p(x, j) dx \\ &= \sum_{j=1}^2 \int_{\mathbb{R}^n} L(c(x), J) P_j p(x | j) dx \end{aligned}$$

From the above definition, it can be noted that given the priori probability of a pattern, P_j , and the conditional probability, $p(x | j)$, selecting different decision function $c(x)$ can result in different average lose. Practically speaking, we hope to have the average lose as little as possible. Therefore, we are particularly interested in the problem of how to find $c_0(x)$ such that the average lose is the minimum. In what follows, we shall attempt to provide a detailed solution to this problem.

For ease of description, for the ease of description, we shall consider only two-class pattern recognition problems, and assume that priori probability P_i , conditional probability density function $p(x | j)$, and the value of a lose function, c_{ij} , are known.

12.2 Minimum Average Lose Classifier Design

In statistical pattern recognition, since the feature vector for a pattern is a random variable, there is always a probability of committing errors no matter which decision scheme we choose to apply in classifications. To put it more accurately, the problem that we shall focus on here is how to compare the strengths and weakness of various decision methods from the point of view of certain statistical criteria, and thereafter under such criteria to find the best solution to our problem.

In order to make our discussions more concise, apart from the assumption that P_j , $p(x | j)$, and c_{ij} are known, we will further assume that the damage caused by misclassifications is greater than that by correct classifications, that is, $c_{12} > c_{22}$ and $c_{21} > c_{11}$. Under these assumptions, we shall attempt to find the optimal decision function, $c_0(x)$, such that the criterion of minimizing $R = E[L(c(X), J)]$ is satisfied, i.e.,

$$\begin{aligned} R_0 &= E[L(c_0(X), J)] \\ &= \min_{c(x)} E[L(c(X), J)] \end{aligned}$$

where decision function $c_0(x)$ is referred to as the minimum average lose decision function.

Since decision function $c(x)$ is essentially a Borel measurable function whose value is given by $\{1, 2\}$, that is:

$$(\mathbb{R}^n, \mathcal{B}^n) \xrightarrow{c(x)} \{1, 2\} \quad (12.10)$$

Now let $B_i = c^{-1}(\{i\})$, $i = 1, 2$ where $c^{-1}(\cdot)$ denotes the inverse of $c(\cdot)$. Thus, we can have $B_1 \cap B_2 = \emptyset$, $B_1 \cup B_2 = \mathbb{R}^n$. As a result, we can reduce the problem of finding decision function $c(x)$ to that of decomposing the complete pattern feature vector space, \mathbb{R}^n , into two non-intersecting Borel measurable sets, B_1 and B_2 . B_1 and B_2 are called decision regions; when $x \in B_i$, we define $c(x) = i$. Once decision region B_i is determined, the average lose can be rewritten as follows:

$$R = \sum_{j=1}^2 \sum_{i=1}^2 \int_{B_i} L(c(x), J) P_j p(x | j) dx$$

$$\begin{aligned}
&= \sum_{j=1}^2 \sum_{i=1}^2 \int_{B_i} c_{ij} P_j p(x | j) dx \\
&= c_{11} P_1 \int_{B_1} p(x | 1) dx + c_{12} P_2 \int_{B_1} p(x | 2) dx \\
&\quad + c_{21} P_1 \int_{B_2} p(x | 1) dx + c_{22} P_2 \int_{B_2} p(x | 2) dx
\end{aligned}$$

Since

$$\int_{B_2} p(x | j) dx = 1 - \int_{B_1} p(x | j) dx \quad j = 1, 2, \quad (12.11)$$

it is obvious that

$$\begin{aligned}
R &= c_{21} P_1 + c_{22} P_2 \\
&\quad + \int_{B_1} \{ [P_2(c_{12} - c_{22})p(x | 2)] - [P_1(c_{21} - c_{11})p(x | 1)] \} dx
\end{aligned}$$

Note that in the above expression, $c_{21} P_1 + c_{22} P_2$ is a constant, and P_j , $p(x | j)$, and c_{ij} are given. The only thing changeable in the integral term is the region of integral, B_1 . In other words, we can only change B_1 in order to change average lose R . At the same time, we may also observe that two integrands $P_2(c_{12} - c_{22})p(x | 2)$ and $P_1(c_{21} - c_{11})p(x | 1)$ are both non-negative. If we are to change B_1 in order to find the minimum average lose, we must realize that B_1 will include only those feature vectors x at which integrands $P_2(c_{12} - c_{22})p(x | 2)$ is smaller than $P_1(c_{21} - c_{11})p(x | 1)$. That is to say, the selected B_1 and the corresponding B_2 ($B_2 = \mathbb{R}^n - B_1$) will allow to establish the following relationships:

$$\begin{aligned}
P_2(c_{12} - c_{22})p(x | 2) &< P_1(c_{21} - c_{11})p(x | 1) \quad \text{when } x \in B_1 \\
P_2(c_{12} - c_{22})p(x | 2) &\geq P_1(c_{21} - c_{11})p(x | 1) \quad \text{when } x \in B_2
\end{aligned}$$

Based on the above, we can readily have the following decision rules: Let x is the feature vector of a certain sample, ω , thus:

$$\begin{aligned}
\text{if } \frac{p(x|2)}{p(x|1)} &< \frac{P_1(c_{21}-c_{11})}{P_2(c_{12}-c_{22})} \quad \text{then } x \in B_1, \text{ hence } \omega \in \Omega_1 \\
\text{if } \frac{p(x|2)}{p(x|1)} &\geq \frac{P_1(c_{21}-c_{11})}{P_2(c_{12}-c_{22})} \quad \text{then } x \in B_2, \text{ hence } \omega \in \Omega_2
\end{aligned} \quad (12.12)$$

In Rule 12.12, $\frac{p(x|2)}{p(x|1)}$ an $\frac{P_1(c_{21}-c_{11})}{P_2(c_{12}-c_{22})}$ are referred to as likelihood ratio and decision threshold, respectively. The selected decision regions, B_1 and B_2 ,

together with the above decision rule will enable us to design a minimum average loss classifier.

12.3 Minimum Error-Probability Classifier Design

Recall that in the above discussion on minimum average loss classifiers, we assume that the values of loss function $c(x)$ satisfy the following:

$$c_{11} = c_{22} = 0, \quad c_{21} = c_{12} = 1$$

From the decision rule of minimum average loss classifiers, we know that if x denotes the value of a feature vector for some pattern ω , then:

$$\begin{aligned} \text{if } \frac{p(x|2)}{p(x|1)} &< \frac{P_1}{P_2} \text{ then } x \in B_1, \text{ hence } \omega \in \Omega_1 \\ \text{if } \frac{p(x|2)}{p(x|1)} &\geq \frac{P_1}{P_2} \text{ then } x \in B_2, \text{ hence } \omega \in \Omega_2 \end{aligned} \quad (12.13)$$

The above decision rule can be rewritten as follows:

$$\begin{aligned} \text{if } P_2 p(x|2) &< P_1 p(x|1) \text{ then } x \in B_1, \text{ hence } \omega \in \Omega_1 \\ \text{if } P_2 p(x|2) &\geq P_1 p(x|1) \text{ then } x \in B_2, \text{ hence } \omega \in \Omega_2 \end{aligned} \quad (12.14)$$

As pointed out earlier, in statistical pattern recognition feature vector x of a certain sample is a random vector, and hence decision function $d(c(X))$ is also a random function. Therefore, it cannot be guaranteed that all classifications based on such a decision function are correct. To evaluate the effectiveness and features of a certain decision rule, we need to consider the probability of misclassifying one pattern into another. In the problems involving two classes of patterns, the feature vectors of patterns Ω_1 and Ω_2 are to be classified into regions B_1 and B_2 , respectively. In such a case, two types of error may be committed; namely, (1) misclassification of samples from pattern Ω_1 as pattern Ω_2 , and (2) misclassification of samples from pattern Ω_2 as pattern Ω_1 . The probability of total error will be equal to the sum of the probabilities of the two errors, that is:

$$R = \int_{B_2} P_1 p(x|1) dx + \int_{B_1} P_2 p(x|2) dx \quad (12.15)$$

In order to give a graphical illustration of the above-mentioned probabilities, let us suppose that feature vector x of a certain sample is a one-dimensional random variable satisfying a normal distribution. Figure 12.3 shows the illustration of the error probabilities in the two-pattern problems.

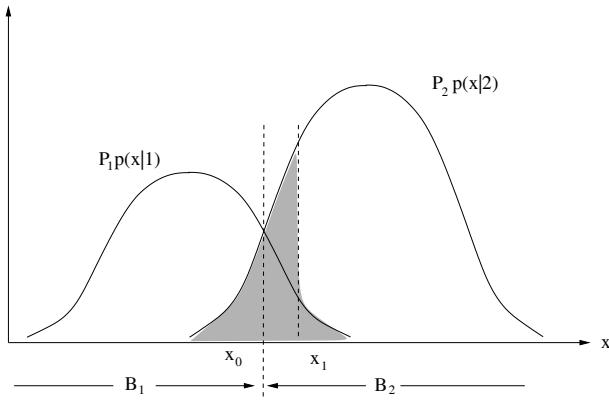


Fig. 12.1 The error probabilities in the two-pattern problems.

From Figure 12.1, we readily note that dividing point x_0 between decision regions B_1 and B_2 satisfies the following:

$$P_1p(x_0 | 1) = P_2p(x_0 | 2). \quad (12.16)$$

Also, we know that if the dividing point (i.e., decision threshold) is shifted to an arbitrary point, x_1 , the total error probability will consequently be increased. This can be confirmed from the following calculations:

$$\begin{aligned} & \int_{B_1} P_2p(x | 2)dx + \int_{B_2} P_1p(x | 1)dx \\ &= \int_{-\infty}^{x_0} P_2p(x | 2)dx + \int_{x_0}^{+\infty} P_1p(x | 1)dx \\ &\leq \int_{-\infty}^{x_1} P_2p(x | 2)dx + \int_{x_1}^{+\infty} P_1p(x | 1)dx \end{aligned} \quad (12.17)$$

Therefore, decision rule 12.14 has the minimum error probability. It is because of this reason that people often refer to the classifiers built using decision rule minimum error probability classifiers.

Based on the above discussion, we may notice that a minimum error probability classifier is a special case of a minimum average lose classifier. To be more specific, a minimum average lose classifier is called a minimum error probability classifier when lose function $c(x)$ satisfies $c_{11} = c_{22} = 0$ and $c_{12} = c_{21} = 1$.

In addition, it should be pointed out that in the above discussion on classifier design, we have assumed that priori probability P_i and conditional probability density function $p(x | j)$ are both given. The two probabilities are used in decision rules 12.13 and 12.14 for the minimum average lose and

the minimum error probability classifiers. However, in real-life applications, it may not be the case that both P_i and $p(x | j)$ are known beforehand. In this book, we shall not deal with the cases of unknown priori probability. Interested readers may find discussions about such cases from other sources. As far as the problems of unknown $p(x | j)$ are concerned, we shall discuss how to estimate $p(x | j)$ by applying a method of orthogonal series approximation from the theory of statistical non-parametric estimation. In particular, we shall discuss how to use orthogonal wavelet series in estimating conditional probability density function $p(x | j)$.

12.4 Probability Density Estimation Based on Orthogonal Wavelet Series

In the preceding sections, we have assumed that conditional probability density function $p(x | j)$ is known. From such an assumption, we have shown how to design classifiers. In the case where $p(x | j)$ is unknown, we have to estimate $p(x | j)$ based on a set of sample feature vectors, X_1, X_2, \dots, X_N . Since the estimation of $p(x | j)$ is the same as the estimation of general probability density function $p(x)$. In order to simplify the notations, in what follows we shall deal only with the non-parametric estimation of general density function $p(x)$.

12.4.1 Kernel Estimation of a Density Function

In density function estimation, the easiest as well as most commonly-used method is the Histogram method. In this method, we use a series of points, $\dots < a_{-1} < a_0 < a_1 < \dots$ to subdivide a real domain into a set of disjoint intervals (a_i, a_{i+1}) . At each interval, a probability value can be estimated based on the following calculation:

$$\#(\{j; 1 \leq j \leq N, a_i \leq X_j < a_{i+1}\}) / N \quad (12.18)$$

where $\#(A)$ returns the numbers of elements in set A . Thus, density function $p(x)$ in $[a_i, a_{i+1})$ can be estimated as follows:

$$\#(\{j; 1 \leq j \leq N, a_i \leq X_j < a_{i+1}\}) / N(a_{i+1} - a_i) \quad (12.19)$$

Using the above-mentioned method can be dated back to as early as the seventh century. In the middle of the twentieth century, statisticians made several significant progresses in the area of non-parametric estimation of density functions. The most pioneering contributions during that time proposed and developed an important density estimator, known as *kernel estimator*. In this way, for each x , we can construct a small region $[x - h_N, x + h_N]$, and then use the following to provide an estimate:

$$\tilde{p}_N(x) = \# (\{j; 1 \leq j \leq N, x - h_N \leq X_j < x + h_N\}) / 2Nh_N \quad (12.20)$$

where $\tilde{p}_N(x)$ denotes an estimate. h_N is a pre-defined positive constant related to N .

The kernel estimation method can be viewed as an improvement over the earlier mentioned histogram method. Now let us define a uniform probability density function, $\tilde{K}(x)$, over $[-1, 1]$, as follows:

$$\begin{aligned} \tilde{K}(x) &= \frac{1}{2}I_{[-1,1]}(x) \\ &= \begin{cases} \frac{1}{2} & -1 \leq x < 1 \\ 0 & \text{otherwise} \end{cases} \end{aligned} \quad (12.21)$$

Using $\tilde{K}(x)$, we can rewrite Eq. 12.20 as follows:

$$\tilde{p}_N(x) = \frac{1}{Nh_N} \sum_{i=1}^N \tilde{K}\left(\frac{x - X_i}{h_N}\right) \quad (12.22)$$

As a matter of fact, $\tilde{K}(x)$ can be changed to any density functions or even other general functions. Thus, we can readily give the following definition:

Definition 12.5 Let $\tilde{K}(y)$ be a Borel measurable function in a 1-D Euclidean space. Thus, we call

$$\tilde{p}_N(x) = \frac{1}{Nh_N} \sum_{i=1}^N \tilde{K}\left(\frac{x - x_i}{h_N}\right) \quad (12.23)$$

a kernel estimator of probability density function $p(x)$ for pattern feature vector X . $\tilde{K}(y)$ is a kernel function and h_N is a window width.

From Eq. 12.23, we can observe the following geometric interpretation for the estimate of $p(x)$, $\tilde{p}_N(x)$: For each sample X_i , we construct step function $\frac{1}{h_N}\tilde{K}\left(\frac{x-X_i}{h_N}\right)$ of window width $2h_N$. The average for N such step

functions will become the estimate of density function $p(x)$. Note that $\lim_{N \rightarrow \infty} h_N = 0$; that is, when the number of samples, N , increases, the step function is gradually becoming an impulse function.

Since the pioneering work on kernel estimation, statisticians have further studied the characteristics of large sample size in density function kernel estimation, and examined the asymptotically unbiasedness, mean square consistency, asymptotical properties of mean square error, and uniformly convergence in probability. Many interesting theoretical results have been obtained from such efforts. Nevertheless, when they apply the large-sample-size kernel estimation in solving practical problems, they immediately face a difficulty, that is, how to determine window width h_N given the actual number of samples, so that the estimation error of $\tilde{p}_N(x)$ with respect to $p(x)$ is smaller than a threshold.

12.4.2 Orthogonal Series Probability Density Estimators

As an alternative to the above-mentioned density function kernel estimation, we can also use the orthogonal series of a function in the $L^2(\mathbb{R}^n)$ space as an asymptotical estimator for density function $p(x)$.

Since most probability density functions are square integrable, i.e., $p(x) \in L^2(\mathbb{R}^n)$, it is possible to expand $p(x)$ using the orthogonal basis of $L^2(\mathbb{R}^n)$. For the sake of brevity, here we shall consider square integrable function space $L^2(\mathbb{R})$ on one-dimensional real space \mathbb{R} .

Let $\{\phi_j(x)\}$ be an orthonormal basis (ONB) in $L^2(\mathbb{R})$ and $\hat{p}(x)$ denote an estimator of density function $p(x)$. Since $p(x) \in L^2(\mathbb{R})$, we have:

$$p(x) = \sum_{j=-\infty}^{+\infty} c_j \phi_j(x) \quad (12.24)$$

Generally speaking, we can use

$$\hat{p}(x) = \sum_{j=1}^m c_j \phi_j(x) \quad (12.25)$$

that is, part of the orthogonal series summation for $p(x)$, $\sum_{j=1}^m c_j \phi_j(x)$, to approximate $p(x)$. In such a case, the mean square error of estimation

becomes

$$\begin{aligned}\gamma_e &= \int_{-\infty}^{+\infty} |p(x) - \hat{p}(x)|^2 dx \\ &= \int_{-\infty}^{+\infty} |p(x) - \sum_{j=1}^m c_j \phi_j(x)|^2 dx\end{aligned}\quad (12.26)$$

The necessary condition for minimizing the above mean square error can be stated as follows:

$$\frac{\partial \gamma_e}{\partial c_k} = 0, \quad k = 1, 2, \dots, m \quad (12.27)$$

That is,

$$\int_{-\infty}^{+\infty} 2(p(x) - \sum_{j=1}^m c_j \phi_j(x)) \phi_k(x) dx = 0 \quad (12.28)$$

Hence,

$$\sum_{j=1}^m c_j \int_{-\infty}^{+\infty} \phi_j(x) \phi_k(x) dx = \int_{-\infty}^{+\infty} \phi_k(x) p(x) dx \quad (12.29)$$

Since $\{\phi_j(x)\}$ is a orthogonal system, the above expression is essentially the following:

$$c_k = \int_{-\infty}^{+\infty} \phi_k(x) p(x) dx \quad (12.30)$$

The right hand side of Eq. 12.30, as the mathematical expectation of random variable $\phi(x)$, can be estimated using the mean value of N samples. Hence, we can write:

$$\hat{c}_k = \frac{1}{N} \sum_{i=1}^N \phi_k(X_i) \quad (12.31)$$

Therefore,

$$\hat{p}(x) = \sum_{j=1}^m \hat{c}_j \phi_j(x) \quad (12.32)$$

There exist many orthonormal bases in $L^2(\mathbb{R})$. The commonly-used ones include Hermite orthogonal system, Laguerre orthogonal system, and Legendre orthogonal system, etc.. No matter which system is chosen in asymptotical series expression $\sum_{j=1}^m \hat{c}_j \phi_j(x)$ to approximate density function $p(x)$, we are inevitably facing the next difficult problem. As we know,

the quality of estimator $\hat{p}(x)$ is closely related to the term number, m , in the base function of $\phi_j(x)$. Our problem here is how to determine m according to the number of samples, N , such that $\hat{p}(x) = \sum_{j=1}^m \hat{c}_j \phi_j(x)$ can best represent $p(x)$ – if so, any pattern classifier based on $\hat{p}(x)$ will be able to function effectively. Theoretically speaking, there is no explicit decision rule for determining such an m . The best way to do so is through empirical experimentation.

12.4.3 Orthogonal Wavelet Series Density Estimators

In order to derive, from a theoretical point of view, a theorem about the large sample size in the case of orthogonal wavelet series density estimation, we shall first of all introduce notions of slowly increasing generalized functions space and Sobolev space.

Definition 12.6 Rapidly decreasing function space S is composed of functions, $C^\infty(\mathbb{R}^n)$, that satisfy the following condition:

$$\sup_{\mathbb{R}^n} |x^\alpha D^\beta \theta(x)| < \infty, \quad \forall \alpha, \beta \in N^n \quad (12.33)$$

$C^\infty(\mathbb{R}^n)$ denotes infinitely differentiable real functions in \mathbb{R}^n .

From the above definition, it can be noted that any function in space S is a C^∞ function that will approach to 0 when $|x| \rightarrow \infty$, at a speed faster than any power of $\frac{1}{|x|}$. Because of this reason, S is normally referred to as rapidly decreasing function space. In S , it is possible to several countable semi norm.

$$\gamma_{\alpha, \beta}(\theta) = \sup_{\mathbb{R}^n} |x^\alpha D^\beta(\theta)|, \quad \alpha, \beta \in N^n \quad (12.34)$$

Thus space S becomes a linear topological space, which means when $\theta_v \rightarrow \theta$, for any exponents α, β , we have:

$$\lim_{v \rightarrow \infty} x^\alpha D^\beta (\theta_v(x) - \theta(x)) = 0, \quad uniformly \quad x \in \mathbb{R}^n \quad (12.35)$$

Definition 12.7 The space composed of all continuous linear functionals on rapidly decreasing function space S is called slowly increasing generalized functions space or tempered distributions space, denoted by S' .

In addition, we use S_γ to denote the space composed of $C^\infty(\mathbb{R}^n)$ functions that satisfy the following conditions:

$$\sup_{\mathbb{R}^n} |x^\alpha D^\beta \theta(x)| < \infty, \quad \forall \alpha, \beta \in N^n, \quad |\beta| \leq \gamma \quad (12.36)$$

Definition 12.8 Sobolev space $H^S(\mathbb{R}^n)$ is defined as follows:

$$H^S(\mathbb{R}^n) = \{u(x); u(x) \in S', (1 + |\omega|^2)^{S/2} \hat{u}(\omega) \in L^2\} \quad (12.37)$$

In Sobolev space, with Hermite inner product:

$$(u, v)_S = \frac{1}{(2\pi)^n} \int (1 + |\omega|^2)^S \hat{u}(\omega) \hat{v}(\omega) d\omega \quad (12.38)$$

it is possible to show that this inner product can make $H^S(\mathbb{R}^n)$ become Hilbert space.

In what follows, we shall introduce the notion of reproducing kernel Hilbert space. The multiresolution analysis $\{V_m\}$ in wavelet analysis theory is in fact a sequence in reproducing kernel Hilbert space.

Definition 12.9 Let a bivariate function $\mathcal{L}(x, y) : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ be symmetric and nonnegative. It is known that there exists a unique Hilbert space, $H(\mathbb{R})$, such that $\forall x \in \mathbb{R}$, $\mathcal{L}(x, \cdot) \in H(\mathbb{R})$. Furthermore, $\forall g(y) \in H(\mathbb{R})$, the following holds:

$$(g(\cdot), \mathcal{L}(x, \cdot))_{H(\mathbb{R})} = g(x) \quad (12.39)$$

The space, $H(\mathbb{R})$ is called a reproducing kernel Hilbert space (RKHS), and bivariate function $\mathcal{L}(x, y)$ is called a reproducing kernel (RK) for $H(\mathbb{R})$.

In fact, each space, V_m , in multiresolution analysis is a reproducing kernel Hilbert space, and the reproducing kernel, $\mathcal{L}(x, y)$, of V_0 can be written as follows:

$$\mathcal{L}(x, y) = \sum_{-\infty}^{+\infty} \phi(x - n) \phi(y - n) \quad (12.40)$$

where $\phi(x)$ is a scaling function. The reproducing kernel, $\mathcal{L}_m(x, y)$, of V_m is given in the following:

$$\mathcal{L}_m(x, y) = 2^m \mathcal{L}(2^m x, 2^m y) \quad (12.41)$$

In order to study the asymptotical properties of reproducing kernel $\mathcal{L}_m(x, y)$, we need to further introduce the property, Z_λ , of scaling function $\phi(x)$.

Definition 12.10 Let scaling function $\phi(x) \in S_r$. $\phi(x)$ is said to satisfy property Z_λ if and only if it satisfies the following:

$$(i) \quad \hat{\phi}(\omega) = 1 + O(|\omega|^\lambda) \quad \text{as } \omega \rightarrow 0 \quad (12.42)$$

$$(ii) \quad Z\phi(x, \omega) = e^{-i\omega x}(1 + O(|\omega|^\lambda)) \quad \text{uniformly as } \omega \rightarrow 0 \quad (12.43)$$

where $Z\phi(x, \omega) \stackrel{\Delta}{=} \sum_{k=-\infty}^{+\infty} e^{-i\omega k} \phi(x - k)$ is called the Zak transform of scaling function $\phi(x)$.

Having introduced the Z_λ property for scaling function $\phi(x)$, in what follows we shall state a related theorem without giving its proof.

Theorem 12.1 Let scaling function $\phi(x) \in S_r$, and for a certain $\lambda > 0$ satisfying the Z_λ , $\mathcal{L}_m(x, y)$ is a reproducing kernel for space V_m . Thus, we have:

$$\|\mathcal{L}_m(x, \cdot) - \delta(x - \cdot)\|_{-\alpha} = O(2^{-m\lambda}) \text{ uniformly for } y \in \mathbb{R} \quad (12.44)$$

where $\|\cdot\|_{-\alpha}$ is a Sobolev norm, and $\alpha > \lambda + \frac{1}{2}$

With the above preparation, we can now address the issue of how to derive a probability density estimate based on orthogonal wavelet series.

As we have mentioned earlier, common density function $p(x) \in L^2(\mathbb{R})$. From the multiresolution theory in wavelet analysis, it is known that:

$$L^2(\mathbb{R}) = \overline{\bigcup_m V_m} \quad (12.45)$$

Let $p_m(x)$ denote the orthogonal project of $p(x)$ in space V_m . Thus,

$$(L^2) \lim_{m \rightarrow \infty} p_m(x) = p(x) \quad (12.46)$$

where

$$p_m(x) = \sum_{n=-\infty}^{+\infty} a_{mn} 2^{m/2} \phi(2^m x - n) \quad (12.47)$$

The minimum mean square error estimator of $p_m(x)$ will be written as follows:

$$\hat{p}_m(x) = \sum_{n=-\infty}^{+\infty} \hat{a}_{mn} 2^{m/2} \phi(2^m x - n) \quad (12.48)$$

where

$$\hat{a}_{mn} = \frac{1}{N} \sum_{i=1}^N 2^{m/2} \phi(2^m X_i - n) \quad (12.49)$$

Thus, we can have the following:

$$\begin{aligned} \hat{p}_m(x) &= \sum_{n=-\infty}^{+\infty} \left[\frac{1}{N} \sum_{i=1}^N 2^{m/2} \phi(2^m X_i - n) \right] 2^{m/2} \phi(2^m x - n) \\ &= \frac{1}{N} \sum_{i=1}^N \sum_{n=-\infty}^{+\infty} 2^m \phi(2^m x - n) \phi(2^m X_i - n) \\ &= \frac{1}{N} \sum_{i=1}^N \mathcal{L}_m(x, X_i) \end{aligned} \quad (12.50)$$

Comparing 12.50 with 12.22, we note that orthogonal wavelet series density estimator $\hat{p}_m(x)$ and kernel estimator $\tilde{p}_N(x)$ are quite similar. From the geometrical point of view, $\mathcal{L}_m(x, X_i)$ is an impulse function scaled from $\mathcal{L}(x, X_i)$. The mean of N impulse functions corresponds to the density estimator, $\hat{p}_m(x)$.

The theorem given below further indicates that when scaling function $\phi(x)$ and unknown density function $p(x)$ satisfy certain specific properties, orthogonal wavelet series density estimator $\hat{p}_m(x)$ will converge to $p(x)$.

Theorem 12.2 *Let scaling function $\phi(x) \in S_r$, and for a certain $\lambda \geq 1$, it satisfies property Z_λ . Let X be a continuous bounded density function random variable, and X_1, X_2, \dots, X_N be N independent identically distributed samples of X . Thus, if $p(x) \in H^\alpha$, $\alpha > \lambda + \frac{1}{2}$, $m \approx \lg N / (2\lambda + 1) \lg 2$, then:*

$$E|\hat{p}_m(x) - p(x)|^2 \leq O(2^{-2m\lambda}) \quad (12.51)$$

Proof: First, we have

$$\begin{aligned}
 & E|\hat{p}_m(x) - p(x)|^2 \\
 = & E|(\hat{p}_m(x) - p_m(x)) + (p_m(x) - p(x))|^2 \\
 = & E|\hat{p}_m(x) - p_m(x)|^2 + E|p_m(x) - p(x)|^2 \\
 & + 2E[(\hat{p}_m(x) - p_m(x))(p_m(x) - p(x))]
 \end{aligned}$$

Note that

$$\begin{aligned}
 E\hat{p}_m(x) &= E\left(\frac{1}{N} \sum_{i=1}^N \mathcal{L}_m(x, X_i)\right) \\
 &= \frac{1}{N} \sum_{i=1}^N E\mathcal{L}_m(x, X_i) \\
 &= \int \mathcal{L}_m(x, y)p(y)dy \\
 &= p_m(x)
 \end{aligned} \tag{12.52}$$

Hence,

$$E[(\hat{p}_m(x) - p_m(x))(p_m(x) - p(x))] = 0 \tag{12.53}$$

Therefore, we can have:

$$\begin{aligned}
 & E|\hat{p}_m(x) - p(x)|^2 \\
 = & E|\hat{p}_m(x) - p_m(x)|^2 + |p_m(x) - p(x)|^2
 \end{aligned} \tag{12.54}$$

where

$$\begin{aligned}
 & E|\hat{p}_m(x) - p_m(x)|^2 \\
 = & E\left|\frac{1}{N} \sum_{i=1}^N [\mathcal{L}_m(x, X_i) - p_m(x)]\right|^2 \\
 = & \frac{1}{N} \left[\int \mathcal{L}_m^2(x, y)p(y)dy - p_m^2(x) \right] \\
 \leq & \frac{1}{N} \int \mathcal{L}_m^2(x, y)p(y)dy \\
 \leq & \frac{1}{N} \|p(\cdot)\|_\infty \mathcal{L}_m(x, x) \\
 = & \frac{2^m}{N} \|p(\cdot)\|_\infty \mathcal{L}_m(2^m x, 2^m x) \\
 = & O\left(\frac{2^m}{N}\right)
 \end{aligned} \tag{12.55}$$

where $\|p(\cdot)\|_\infty$ is a constant, $\mathcal{L}_m(2^m x, 2^m x)$ is an impulse function that has the same magnitude as $\mathcal{L}(x, x)$. If we let $m = O(\lg N)$, then when $m \rightarrow \infty$,

$\mathcal{L}(2^m x, 2^m x)$ will gradually become a point impulse. Furthermore, since $\lim_{m \rightarrow \infty} \frac{2^m}{N} = 0$, we have $\lim_{m \rightarrow \infty} E|\hat{p}_m(x) - p_m(x)|^2 = 0$. On the other hand,

$$\begin{aligned} & |p_m(x) - p(x)| \\ = & \left| \int \mathcal{L}_m(x, y)p(y)dy - p(x) \right| \\ = & \left| \int [\mathcal{L}_m(x, y) - \delta(x - y)]p(y)dy \right| \\ \leq & \|\mathcal{L}_m(x, \cdot) - \delta(x - \cdot)\|_{-\alpha} \|p\|_\alpha \end{aligned} \quad (12.56)$$

Inequality 12.56 is a Schwarz inequality in the Sobolev space. From Theorem 12.1, we know that:

$$\|\mathcal{L}_m(x, \cdot) - \delta(x - \cdot)\|_{-\alpha} = O(2^{-m\lambda}) \quad (12.57)$$

Therefore,

$$|p_m(x) - p(x)|^2 \leq O(2^{-2m\lambda}) \quad (12.58)$$

If we let $N = 2^{m(2\lambda+1)}$, we can write $m = \lg N / (2\lambda + 1) \lg 2 = O(\lg N)$. Since $\frac{2^m}{N} = 2^{-2m\lambda}$, we have $\lim_{m \rightarrow \infty} \frac{2^m}{N} = \lim_{m \rightarrow \infty} 2^{-2m\lambda} = 0$.

Hence, based on the above derivations, we can arrive at the following expression:

$$E|\hat{p}_m(x) - p(x)|^2 \leq O(2^{-2m\lambda}) \quad (12.59)$$

This concludes our proof.

Let us now recall the orthogonal wavelet series density estimator, $\hat{p}_m(x)$,

$$\hat{p}_m(x) = \frac{1}{N} \sum_{I=1}^N \mathcal{L}_m(x, X_i) \quad (12.60)$$

Based on Theorem 12.2, we know that if the number of samples X_i for feature vector X is given, as denoted by N , m can be rewritten as $m \approx \lg N / (2\lambda + 1) \lg 2$. Once m is known, the window width of impulse function $\mathcal{L}_m(x, X_i)$ can accordingly be determined. In other words, Theorem 12.2 provides a criterion for determining the window width of impulse function $\mathcal{L}_m(x, X_i)$ from the number of samples, N , which is exactly what the early-mentioned kernel estimator is lacking of.

From the above discussions, we can note that the orthogonal wavelet series estimator differs from the kernel estimator and the traditional orthogonal series density estimator. Its basic idea shares some similarities to that of the traditional orthogonal series density estimator. However, it also satisfies several key properties of kernel estimator and exhibits some additional features. Generally speaking, the orthogonal wavelet series density estimator represents a new non-parametric way of estimating density functions, which has a great potential for practical applications. For instance, in pattern classifier design, sometimes, the probability density function, $p(x)$, of a certain feature vector may not be available. In such a case, we can readily replace $p(x)$ with $\hat{p}_m(x)$ using the above-described orthogonal wavelet series density estimator, and thus, effectively design the classifiers.

This page intentionally left blank

List of Symbols

- \mathbb{R} : the set of all the real numbers;
- \mathbb{N} : the set of all the natural numbers;
- \mathbb{Z} : the set of all the integers;
- \mathbb{Z}^+ : the set of all the negative integers;
- \mathbb{R}^d : d-dimensional Euclidean space;
- \mathbb{T} : the quotient group $\mathbb{R}/2\pi\mathbb{Z}$;
- $Re(z)$: the real part x of complex number $z = x + iy$;
- $Im(z)$: the imaginary part y of complex number $z = x + iy$;
- For sets A and B , $A \setminus B := \{x|x \in A \text{ and } x \notin B\}$;
- a.e.: almost everywhere;
- Let M be a matrix, its transposed matrix is denoted by M^t ; If M is invertible, its inverse matrix is denoted by M^{-1} and its determinant is denoted by $det(M)$;
- L^p , $L^p(\mathbb{R})$ ($1 \leq p < \infty$): the space of all the p -power integrable functions, i.e.:

$$L^p := L^p(\mathbb{R}) := \left\{ f \mid \int_{\mathbb{R}} |f(x)|^p dx < \infty \right\}.$$

- L^p , $L^p(\mathbb{T})$ ($1 \leq p < \infty$): the space of all the p -power integrable, 2π -periodic functions, i.e.:

$$L^p := L^p(\mathbb{T}) := \left\{ f \mid \int_{\mathbb{T}} |f(x)|^p dx < \infty \right\}.$$

- l^p , $l^p(\mathbb{Z})$ ($1 \leq p < \infty$): the space of all the p -power sumable

sequences, i.e.:

$$l^p := l^p(\mathbb{Z}) := \left\{ \{c_k\} \mid \sum_{k \in \mathbb{Z}} |c_k|^p < \infty \right\}.$$

- The Fourier transform of $f \in L^1(\mathbb{R})$ is defined by

$$\hat{f}(\xi) := \int_{\mathbb{R}} f(x) e^{-ix \cdot \xi} dx,$$

and the Inverse Fourier is defined by

$$\check{f} = \frac{1}{2\pi} \int_{\mathbb{R}} f(x) e^{ix \cdot \xi} dx.$$

- The inner product of $f, g \in L^2(\mathbb{R})$ is defined by

$$\langle f, g \rangle := \int_{\mathbb{R}} f(x) \bar{g}(x) dx,$$

and the following equality holds always:

$$\langle f, g \rangle = \frac{1}{2\pi} \langle \hat{f}, \hat{g} \rangle.$$

- The Fourier coefficients of $f \in L^1(\mathbb{T})$ are defined by

$$c_k(f) := \int_{\mathbb{T}} f(x) e^{-ik \cdot x} dx.$$

It is always true that $f(x) \sim \sum_{k \in \mathbb{Z}} c_k(f) e^{ik \cdot x}$.

- Let $\phi(x)$ is a function defined on \mathbb{R} , denote

$$\phi_{j,k}(x) := 2^{j/2} \phi(2^j x - k) \quad (\forall j \in \mathbb{Z}, k \in \mathbb{Z}).$$

- The support of a complex sequence $\{c_k\}_{k \in \mathbb{Z}}$ is defined by

$$\text{supp}\{c_k\}_{k \in \mathbb{Z}} := \{k \in \mathbb{Z} \mid c_k \neq 0\}.$$

Bibliography

- Aghajan, H. K. and Kailath, T. (1994). SLIDE: Subspace-based line detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(11):1057–1073.
- Auslander, L., Kailath, T., and Mitter, S., editors (1990). *Signal Processing I: Signal Processing Theory*. Springer-Verlag, New York.
- Beylkin, G., Coifman, R., Daubechies, I., Mallat, S., Meyer, Y., Raphael, L., and Ruskai, B. (1991). *Wavelets and their Applications*. MA, MA: Jones and Bartlett.
- Blum, H. (1967). *A Transformation for Extracting New Descriptors of Shape*, W. Wathen-Dunn, Eds., *Models for the Perception of Speech and Visual Form*. The MIT Press, Massachusetts.
- Bow, S. T. (1992). *Pattern Recognition and Image Preprocessing*. Marcel-Dekker, Marcel-Dekker: New York.
- Brady, M. (1983). *Criteria for Representation of Shape*, J. Beck and B. Hope and A. Rosenfeld, Eds., *Human and Machine Vision*. Academic Press, New York.
- Brady, M. and Asada, H. (1993). Smoothed local symmetries and their implementation. *Int'l J. Robotics Res.*, 15:973–981.
- Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6):679–698.
- Chambolle, A., DeVore, R. A., Lee, N. Y., and Lucier, B. J. (1998). Nonlinear wavelet image processing: Variational problems, compression, and noise removal through wavelet shrinkage. *IEEE Transactions on Image Processing*, 7(3):319–335.
- Chang, H. S. and Yan, H. (1999). Analysis of stroke structures of handwritten chinese characters. *IEEE Transactions on Systems, Man, and Cybernetics (B)*, 29:47–61.
- Chao-hao, G. (1992). *Mathematics Dictionary*. Shanghai Dictionary Press, Shanghai, China.

- Chen, C., Lee, J., and Sun, Y. (1995). Wavelet transformation for gray-level corner detection. *Pattern Recognition*, 28(6):853–861.
- Chen, G. and Yang, Y. H. H. (1995). Edge detection by regularized cubic B-spline fitting. *IEEE Transactions on Systems, Man, and Cybernetics*, 25(4):636–643.
- Chuang, G. C. H. and Kuo, C. C. J. (1996). Wavelet descriptor of planar curves: Theory and applications. *IEEE Transactions on Image Processing*, 5(1):56–70.
- Chui, C. K. (1992). *An Introduction to Wavelets*. Academic Press, Boston.
- Combettes, P. L. (1998). Convex multiresolution analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(12):1308–1318.
- Combettes, P. L. and Pesquet, J. C. (2004). Wavelet-constrained image restoration. *International Journal of Wavelets, Multiresolution and Information Processing*, 2(4):371–389.
- Connell, J. H. and Brady, M. (1987). Generating and generalizing models of visual objects. *Artifial Intelligence*, 31:159–183.
- Daubechies, I. (1988). Orthonormal bases of compactly supported wavelets. *Comm. on Pure & Applied Mathematics*, 41(7):909–996.
- Daubechies, I. (1990). Wavelet transform, time-frequency localization and signal analysis. *IEEE Transactions on Information Theory*, 36:961–1005.
- Daubechies, I. (1992). *Ten Lectures on Wavelets*, volume 61. CBMS - Conference Lecture Notes, SIAM Philadelphia.
- Daugman, J. (2003). Demodulation by complex-valued wavelets for stochastic pattern recognition. *International Journal of Wavelets, Multiresolution and Information Processing*, 1(1):1–317.
- de Wouwer, G., Scheunders, P., Livens, S., and Dyck, D. V. (1999a). Wavelet correlation signatures for color texture characterization. *Pattern Recognition*, 32:443–451.
- de Wouwer, G. V., Schenuders, P., and Dyck, D. V. (1999b). Statistical texture characterization from discrete wavelet representation. *IEEE Transactions on Image Processing*, 8:592–598.
- Deng, W. and Lyengar, S. S. (1996). A new probability relaxation scheme and its application to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(4):432–443.
- Edgar, G. A. (1990). *Measure, Topology, and Fractal Geometry*. Springer-Verlag, New York.
- El-Khamy, S. E., Hadhoud, M. M., Dessouky, M. I., Salam, B. M., and El-Samie, F. E. A. (2006). Wavelet fusion: A tool to break the limits on lmmse image super-resolution. *International Journal of Wavelets, Multiresolution and Information Processing*, 4(1):105–118.
- Falconer, K. L. (1985). *The Geometry of Fractal Sets*. Cambridge University Press.
- Falconer, K. L. (1990). *Fractal Geometry: Mathematical Foundation and Applications*. New York: Wiley.

- Gabor, D. (1946). Theory of communication. *Journal Inst. Elect. Engr.*, 93:429–457.
- Ge, Y. and Fitzpatrick, J. M. (1996). On the generation of skeletons from discrete euclidean distance maps. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18:1055–1066.
- Gilbarg, D. and Trudinger, N. S. (1977). *Elliptic Partial Differential Equations of Second Order*. Berlin ; New York : Springer.
- Grossmann, A. and Morlet, J. (1984). Decomposition of Hardy function into square integrable wavelets of constant shape. *SIAM J. Math. Anal.*, 15:723–736.
- Grossmann, A. and Morlet, J. (1985). Decomposition of functions into wavelets of constant shape, and related transforms. In Streit, L., editor, *Lecture on Recent Results*, Mathematics and Physics. World Scientific publishing, Singapore.
- Grossmann, A., Morlet, J., and Paul, T. (1985). Transforms associated to square integrable group representations I. General results. *J. Math. Phys.*, 26:2473–2479.
- Haley, G. M. and Manjunath, B. S. (1999). Rotation-invariant texture classification using a complete space-frequency model. *IEEE Transactions on Image Processing*, 8(2):255–269.
- Hsieh, J.-W., Liao, H.-M., Ko, M.-T., and Fan, K.-C. (1995). Wavelet-based shape form shading. *Graphical Models and Image Processing*, 57(4):343–362.
- IEEE (1992). Special issue on wavelet transforms and multiresolution signal analysis. *IEEE Transactions on Information Theory*, 38(2):529–924.
- IEEE (1993). Special issue on wavelets and signal processing. *IEEE Transactions on Signal Processing*, 41(12):3213–3600.
- Jain, P. and Merchant, S. N. (2004). Wavelet-based multiresolution histogram for fast image retrieval. *International Journal of Wavelets, Multiresolution and Information Processing*, 2(1):59–73.
- Janssen, R. D. T. (1997). *Interpretation of Maps: from Bottom-up to Model-based*, H. Bunke and P. S. P. Wang, Eds., *Handbook of Character Recognition and Document Image Analysis*. World Scientific, Singapore.
- Kouzani, A. Z. and Ong, S. H. (2003). Lighting-effects classification in facial images using wavelet packets transform. *International Journal of Wavelets, Multiresolution and Information Processing*, 1(2):199–215.
- Kovesi, P. (1995). Image features from phase congruency. Technical report, University of Western Australia, Robotics and Vision Group, Australia.
- Kovesi, P. (1997). Symmetry and asymmetry from local phase. In *Tenth Australian Joint Conference on Artificial Intelligence*, pages 2–4.
- Ksantini, R., Ziou, D., Dubeau, F., and Harinarayan, P. (2006). Image retrieval based on region separation and multiresolution analysis. *International Journal of Wavelets, Multiresolution and Information Processing*, 4(1):147–175.
- Kubo, M., Aghbari, Z., and Makinouchi, A. (2003). Content-based image retrieval technique using wavelet-based shift and brightness invariant edge feature.

- International Journal of Wavelets, Multiresolution and Information Processing*, 1(2):163–178.
- Kumar, S. and Kumar, D. K. (2005). Visual hand gestures classification using wavelet transforms and moment based features. *International Journal of Wavelets, Multiresolution and Information Processing*, 3(1):79–101.
- Kumar, S., Kumar, D. K., Sharma, A., and McLachlan, N. (2003). Visual hand gestures classification using wavelet transforms. *International Journal of Wavelets, Multiresolution and Information Processing*, 1(4):373–392.
- Kunte, R. S. and Samuel, R. D. S. (2007). Wavelet descriptors for recognition of basic symbols in printed kannada text. *International Journal of Wavelets, Multiresolution and Information Processing*, 5(2):351–367.
- Lai, J. H., Yuen, P. C., and Feng, G. C. (1999). Spectroface: A Fourier-based approach for human face recognition. *Proc. of ICMI '99*, pages VI115–120.
- Lam, L., Lee, S. W., and Suen, C. Y. (1992). Thinning methodologies - a comprehensive survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14:869–885.
- Law, T., Iton, H., and Seki, H. (1996). Image filtering, edge detection, and edge tracing using fuzzy reasoning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(5):481–491.
- Lee, S. W., Kim, C. H., Ma, H., and Tang, Y. Y. (1996). Multiresolution recognition of unconstrained handwritten numerals with wavelet transform and multilayer cluster neural network. *Pattern Recognition*, 29:1953–1961.
- Leyton, M. (1988). *Symmetry, Causality, Mind*. The MIT Press, Massachusetts.
- Li, H. (2006). Wavelet-based weighted average and human vision system image fusion. *International Journal of Wavelets, Multiresolution and Information Processing*, 4(1):97–103.
- Li, S. (2008). Multisensor remote sensing image fusion using stationary wavelet transform: Effects of basis and decomposition level. *International Journal of Wavelets, Multiresolution and Information Processing*, 6(1):37–50.
- Liang, K. H., Chang, F., Tan, T. M., and Hwang, W. L. (1999). Multiresolution Hadamard representation and its application to document image analysis. In *Proc. of The Second Int. Conf. on Multimodel Interface (ICMI'99)*, pages V1–V6, Hong Kong.
- Liang, K. H. and Tjahjadi, T. (2006). Adaptive scale fixing for multiscale texture segmentation. *IEEE Transactions on Image Processing*, 15(1):249–256.
- Liao, Z. and Tang, Y. Y. (2005). Signal denoising using wavelets and block hidden markov model. *International Journal of Pattern Recognition and Artificial Intelligence*, 19(5):681–700.
- Long, R. (1995). *High-Dimensional Wavelet Analysis*. World Books Co. Pte. Ltd, Beijing.
- Ma, H., Xi, D. H., Mao, X. G., Tang, Y. Y., and Suen, C. Y. (1995). Document analysis by fractal signatures. In Spitz, A. L. and Dengel, A., editors, *Document Analysis Systems*. World Scientific Publishing Co. Pte, Ltd., Singapore.

- Mallat, S. (1989a). Multiresolution approximations and wavelet orthonormal bases of $L^2(R)$. *Trans. Amer. Math. Soc.*, 315:69–87.
- Mallat, S. (1989b). A theory of multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11:674–693.
- Mallat, S. (1998). *Wavelet Tour of Signal Processing*. Academic Press, San Diego, USA.
- Mallat, S. and Hwang, W. L. (1992). Singularity detection and processing with wavelets. *IEEE Transactions on Information Theory*, 38:617–643.
- Mallat, S. and Zhong, S. (1992). Characterization of signals from multiscale edges. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(7):710–732.
- Mallat, S. G. (1989c). Multifrequency channel decompositions of images and wavelet models. *IEEE Transactions on Acoust. Speech Signal Process.*, 37(12):2091–2110.
- Mandelbrot, B. B. (1982). *The Fractal Geometry of Nature*. New York: Freeman.
- Marr, D. and Hildreth, E. C. (1980). Theory of edge detection. In *Proc. of Roy. Soc.*, pages 187–217, London.
- Matalas, L., Benjamin, R., and Kitney, R. (1997). An edge detection technique using the facet model and parameterized relaxation labeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(4):328–341.
- Meyer, Y. (1990). *Ondelettes et Operateurs*, volume I, II. Hermann, Paris.
- Moghaddam, H. A., Khajoe, T., Rouhi, A. H., and Tarzjan, M. S. (2005). Wavelet correlogram: A new approach for image indexing and retrieval. *Pattern Recognition*, 38(12):2506–2518.
- Morlet, J., Arens, G., Fourgeau, E., and Giard, D. (1982a). Wave propagation and sampling theory-Part 1: Complex signal and scattering in multilayered media. *Geophysics*, 47:203–221.
- Morlet, J., Arens, G., Fourgeau, E., and Giard, D. (1982b). Wave propagation and sampling theory-Part 2: Sampling theory and complex waves. *Geophysics*, 47:222–236.
- Muneeswaran, K., Ganesan, L., Arumugam, S., and Harinarayan, P. (2005). A novel approach combining gabor wavelet transforms and moments for texture segmentation. *International Journal of Wavelets, Multiresolution and Information Processing*, 3(4):559–572.
- Munkres, J. R. (1975). *Topology, A First Course*. New Jersey: Prentice-Hall Inc.
- Murtagh, F. and Starck, J. L. (1998). Pattern clustering based on noise modeling in wavelet space. *Pattern Recognition*, 31:847–855.
- Nastar, C. and Ayache, N. (1996). Frequency-based non-rigid motion analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(11).
- Ogniewicz, R. L. and Kubler, O. (1995). Hierarchic voronoi skeletons. *Pattern Recognition*, 28:343–359.
- O'Toole, A., Abdi, H., Deffenbacher, K., and Valentin, D. (1993). Low-dimensional representation of faces in higher dimensions of the face space.

- J. Opt. Soc. Am. A.*, 10(3):405–411.
- Pavlidis, T. (1986). Vectorizer and feature extractor for document recognition. *Computer Vision, Graphics, Image Processing*, 35:111–127.
- Procter, P. (1993). *Longman English-Chinese Dictionary of Contemporary English*. Hong Kong: Longman Asia Limited.
- Rom, H. and Medioni, G. (1984). Hierarchical decomposition and axial shape description. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 3:36–61.
- Rudin, W. (1973). *Functional Analysis*. McGraw-Hill, Inc., New York.
- Rudin, W. (1974). *Real and Complex Analysis*. McGraw-Hill Book Company, 2nd edition.
- Saint-Marc, P., Rom, H., and Medioni, G. (1993). B-spline contour representation and symmetry detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15:1191–1197.
- Shankar, B. U., Meher, S. K., and Chosh, A. (2007). Neuro-wavelet classifier for multispectral remote sensing images. *International Journal of Wavelets, Multiresolution and Information Processing*, 5(4):589–611.
- Sharnia, A., Kumart, D. K., and Kumar, S. (2004). Wavelet directional histograms of the spatio-temporal templates of human gestures. *International Journal of Wavelets, Multiresolution and Information Processing*, 2(3):283–298.
- Shen, D. and Ip, H. H. S. (1999). Discriminative wavelet shape descriptors for recognition of 2-D pattern. *Pattern Recognition*, 32:151–165.
- Shensa, M. J. (1992). The discrete wavelets: Wedding the atrous and Mallat algorithms. *IEEE Transactions on Signal Processing*, 40:2464–2482.
- Sirovich, L. and Kirby, M. (1987). Low-dimensional procedure for the characterization of human faces. *J. of Opt. Soc. Amer. A.*, 4(3):519–524.
- Smeulders, A. W. M., Worring, M., Santini, S., Gupta, A., and Jain, R. (2000). Content-based image retrieval at the end of early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22:1349–13805.
- Smith, R. W. (1987). Computer processing of line images: A survey. *Pattern Recognition*, 20:7–15.
- Special-Issue-Digital-Library (1996). *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18.
- SPIE (1994). Special issue on wavelet applications. In Szu, H. H., editor, *Proc. of SPIE 2242*.
- Starck, J. L., Bijaoui, A., and Murtagh, F. (1995). Multiresolution support applied to image filtering and deconvolution. *Graphical Models Image Processing*, 57:420–431.
- stollnitz, E. J., Derose, T. D., and Salesin, D. H. (1996). *Wavelets for Computer Graphics: Theory and Applications*. Morgan Kaufmann Publishers.
- Swets, D. L. and Weng, J. (1996). Using discriminant eigenfeatures for image retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(8):831–836.

- Tang, Y. Y., Cheng, H. D., and Suen, C. Y. (1991). Transformation-ring-projection (TRP) algorithm and its VLSI implementation. *International Journal of Pattern Recognition and Artificial Intelligence*, 5(1 and 2):25–56.
- Tang, Y. Y., Li, B., Ma, H., and Liu, J. (1998a). Ring-projection-wavelet-fractal signatures: A novel approach to feature extraction. *IEEE Transactions on Circuits and Systems II*.
- Tang, Y. Y., Lihua, Q. S., Yang, and Feng, L. (1998b). Two-dimensional overlap-save method in handwriting recognition. In *Proc. of 6th International Workshop on Frontiers in Handwriting Recognition(IWFHR'98)*, pages 627–633, Taejon, Korea.
- Tang, Y. Y., Liu, J., Ma, H., and B.Li (1996a). Two-dimensional wavelet transform in document analysis. In *Proc. of 1st Int. Conf. on Multimodal Interface*, pages 274–279, Beijing, China.
- Tang, Y. Y., Liu, J., Ma, H., and Li, B. F. (1999). Wavelet orthonormal decomposition for extracting features in pattern recognition. *International Journal of Pattern Recognition and Artificial Intelligence*, 13(6):803–831.
- Tang, Y. Y. and Ma, H. (2000). Classifier design based on orthogonal wavelet series. Technical report, Hong Kong Baptist.
- Tang, Y. Y., Ma, H., B.Li, and Liu, J. (1996b). Character recognition based on Doubuchies wavelet. In *Proc. of 1st Int. Conf. on Multimodal Interface*, pages 215–220, Beijing, China.
- Tang, Y. Y., Ma, H., Liu, J., Li, B., and Xi, D. (1997a). Multiresolution analysis in extraction of reference lines from documents with graylevel background. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19:921–926.
- Tang, Y. Y., Ma, H., Xi, D., Cheng, Y., and Suen, C. Y. (1995a). Extraction of reference lines from document with grey-level background using sub-image of wavelets. In *Proc. of 3-nd Int. Conf. on Document Analysis and Recognition*, pages 571–574, Montreal, Canada.
- Tang, Y. Y., Ma, H., Xi, D., Cheng, Y., and Suen, C. Y. (1995b). A new approach to document analysis based on modified fractal signature. In *Proc. of 3-rd Int. Conf. on Document Analysis and Recognition*, pages 567–570, Montreal, Canada.
- Tang, Y. Y., Ma, H., Xi, D., Mao, X., and Suen, C. Y. (1997b). Modified fractal signature (MFS): A new approach to document analysis for automatic knowledge acquisition. *IEEE Transactions on Knowledge and Data Engineering*, 9(5):747–762.
- Tang, Y. Y., Suen, C. Y., and Yan, C. D. (1994). Document processing for automatic knowledge acquisition. *IEEE Transactions on Knowledge and Data Engineering*, 6(1):3–21.
- Tang, Y. Y., Yan, C. D., Cheriet, M., and Suen, C. Y. (1995c). Financial document processing based on staff line and description language. *IEEE Transactions on Systems, Man, and Cybernetics*, 25(5):738–754.

- Tang, Y. Y., Yang, L., and Liu, J. (2000). Characterization of dirac-structure edges with wavelet transform. *IEEE Transactions on Systems, Man, and Cybernetics (B)*, 30(1):93–109.
- Tang, Y. Y., Yang, L. H., and Feng, L. (1998c). Characterization and detection of edges by Lipschitz exponent and MASW wavelet transform. In *Proc. of the 14th Int. Conf. on Pattern Recognition*, pages 1572–1574, Brisbane, Australia.
- Tang, Y. Y., Yang, L. H., and Feng, L. (1998d). Contour detection of handwriting by modular-angle-separated wavelets. In *Proc. of the 6-th Inter. Workshop on Frontiers of Handwriting Recognition (IWFHR-VI)*, pages 357–366, Taejon, Korea.
- Tang, Y. Y., Yang, L. H., Feng, L., and Liu, J. (1998e). Scale-independent wavelet algorithm for detecting step-structure edges. Technical report, Hong Kong Baptist University University.
- Tang, Y. Y., Yang, L. H., and Liu, J. (1997c). Wavelet-based edge detection in chinese document. In *Proc. of the 17th Int. Conf. on Computer Processing of Oriental Languages*, volume 1, pages 333–336.
- Tieng, Q. M. and You, X. G. (2003). Skeletonization of ribbon-like shapes based on a new wavelet function. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25:1118–1133.
- Thune, M., Olstad, B., and Thune, N. (1997). Edge detection in noisy data using finite mixture distribute analysis. *Pattern Recognition*, 30(5):685–699.
- Tieng, Q. M. and Boles, W. W. (1997a). Recognition of 2D object contours using the wavelet transform zero-crossing representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19:910–916.
- Tieng, Q. M. and Boles, W. W. (1997b). Wavelet-based affine invariant representation: A tool for recognizing planar objects in 3D space. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19:846–857.
- Tombre, K. (1998). *Analysis of Engineering Drawings: State of the Art and Challenges*, K. Tombre and A. K. Chhabra, Eds., *Graphics Recognition: Algorithm and Systems*, volume 1389 of *Lecture Notes in Computer Science*. Springer-Verlag, Berlin.
- Tou, J. T. and Gonzalez, R. C. (1974). *Pttern Recognition Principles*. Addison-Wesley, London.
- Turk, M. and Pentland, A. (1991). Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1):71–86.
- Unser, M., Aldroubi, A., and Eden, M. (1992). On the asymptotic convergence of B-spline wavelets to Gabor functions. *IEEE Transactions on Information Theory*, 38:864–872.
- Unser, M., Aldroubi, A., and Eden, M. (1993). A family of polynomial spline wavelets transforms. *Signal Processing*, 30:141–162.
- Venkatesh, S. and Owens, R. (1990). On the classification of image features. *Pattern Recognition Letters*, 11:339–349.
- Wunsch, P. and Laine, A. F. (1995). Wavelet descriptors for multiresolution

- recognition of handprinted characters. *Pattern Recognition*, 28(8):1237–1249.
- Yang, F., Wang, Z., and Yu, Y. L. (1998). Chinese typeface generation and composition using B-Spline wavelet transform. In *Proc. of SPIE, Wavelet Applications V Orlando, Florida, 1998*, pages 616–620.
- Yang, L., Suen, C. Y., and Tang, Y. Y. (2003a). A width-invariant property of curves based on wavelet transform with a novel wavelet function. *IEEE Transactions on Systems, Man, and Cybernetics (B)*, 33(3):541–548.
- Yang, L. H., Bui, T. D., and Suen, C. Y. (2003b). Image recognition based on nonlinear wavelet approximation. *International Journal of Wavelets, Multiresolution and Information Processing*, 1(2):151–161.
- Yang, L. H., You, X., Haralick, R. M., Phillips, I. T., and Tang, Y. (2001). Characterization of dirac edge with new wavelet transform. In *Proc. of 2th Int. Conf. Wavelets and its Application*, volume 1, pages 872–878, Hong Kong.
- Yoon, S. H., Kim, J. H., Alexander, W. E., Park, S. M., and Sohn, K. H. (1998). An optimum solution for scale-invariant object recognition based on the multi-resolution approximation. *Pattern Recognition*, 31:889–908.
- You, X., Chen, Q., Fang, B., and Tang, Y. Y. (2006). Thinning character using modulus minima of wavelet transform. *International Journal of Pattern Recognition and Artificial Intelligence*, 20(3):361–376.
- You, X. and Tang, Y. Y. (2007). Wavelet-based approach to character skeleton. *IEEE Transactions on Image Processing*, 16:1220–1231.
- Young, R. K. (1993). *Wavelet Theory and its Applications*. Kluwer Academic Publishers, Boston.
- Yuen, P. C., Dai, D. Q., and Feng, G. C. (1998). Wavelet-based PCA for human face recognition. *Proceeding of IEEE Southwest Symposium on Image Analysis and Interpretation*, pages 223–228.
- Zhang, G. (1986). *Lecture of Functional Analysis*, volume 1. Beijing University Press.
- Zou, J. J. and Yan, H. (2001). Skeletonization of ribbon-like shapes based on regularity and singularity analyses. *IEEE Transactions on Systems, Man, and Cybernetics (B)*, 31(3):401–407.

This page intentionally left blank

Index

- C^∞ function
 - fast decreasing, 112
- admissibility
 - admissible wavelet, 75
 - condition, 75
- algorithm, 116, 239, 241, 252, 272, 274, 301, 305, 306, 314, 330, 331, 383, 386, 388, 391, 392, 394, 404, 406, 409, 411–413, 416, 419, 420
- B-spline, 143, 399
- bank cheque, 381
- basic condition, 128
 - basic condition I, 120, 128
 - basic condition II, 120, 128
 - basic condition III, 128
- basic smooth condition, 143, 144
- basic wavelet, 75, 83, 85, 88, 98
- basic window
 - center of, 78
 - radius of, 78
- biorthonormality, 109
- cardinal splines, 181
- character processing, 399
- characteristic function, 137
- class function, 431
- classification, 19, 31
- classifier, 31, 32
- classifier design, 429
 - minimum average loss, 433
 - minimum error-probability, 435
- clustering, 19, 33
- continuous multiresolution approximation (CMA), 52, 53
- decision function, 430
- dilation, 76
- document
 - analysis, 19, 37–39, 381
 - form, 38, 387
 - processing, 38, 381
- downsample, 151
- edge
 - detection, 45, 201
 - Dirac, 226, 229, 236, 251, 252, 255, 257, 258, 268, 272, 274, 275, 279, 284, 285, 287, 294, 301
 - pixel, 206
 - roof, 226, 229, 236, 251
 - signal, 202
 - step, 45, 226, 229, 236, 251, 275
- estimator
 - kernel, 438
- Euler-Frobenius polynomials, 184, 193
- even function, 285, 317

- filter function, 108
- Fourier transform, 5, 8, 16, 75
- frequency window, 78
- Gaussian, 80, 284, 296
- generalized function
 - slowly increasing, 112
 - support of, 117
- handprinted character, 48, 56
- Heisenberg, 80
- image
 - fusion, 20, 70
 - indexing, 20, 67
 - retrieval, 20, 67
- Indian language, 59
- invariant
 - grey-level, 251, 259, 265, 279, 286
 - moment, 50
 - representation, 20, 49
 - rotation, 49, 50, 62, 345, 349
 - scale-, 52
 - slope, 251, 259, 263, 279, 287
 - width, 280, 287, 294
- limit function, 120, 128
- linear phase, 93
 - generalized, 93
- Lipschitz exponent, 95, 225, 227
- Lipschitz regularity, 94, 226
 - uniformly, 94, 95, 225
- local maximum modulus, 258, 262, 264, 273, 274, 276, 285, 294, 299, 305, 307, 329
- Mallat algorithm, 148, 383, 387
 - decomposition, 149, 150
 - discrete form, 150
 - reconstruction, 149, 151
- mask, 108
- matrix, 135
 - eigenpolynomial of, 135
 - eigenvalue of, 135
- eigenvector of, 135, 141
- maximum moduli, 322, 325–327, 329–334, 336, 339
- MMAWT, 331, 332
- modulus-angle-separated wavelets (MASW), 233–235, 237, 242, 245, 252
- motion history image (MHI), 28, 29
- multiresolution analysis, 69, 105, 107, 383
- biorthonormal, 109, 127
- orthonormal, 107, 109, 132, 133
- multiresolution feature, 56, 59
- multiresolution Hadamard representation (MHR), 39–41
- new wavelet, 284, 314, 316
- odd function, 285, 316
- orthonormality, 109
- pattern recognition, 7, 18, 201, 202, 222, 252, 311, 343–345, 356, 358, 365, 381, 399, 429, 432, 435
- quadratic spline, 317
- recognition
 - character, 20, 47, 56, 59
 - face, 18, 22, 24, 27
 - hand gesture, 18, 28–30
 - handwritten, 20
 - iris, 18, 21
- reconstruction, 85
- reference line, 381
- Ribbon-like shape, 311, 316–318, 331, 332
- Riesz basis, 106, 107
 - lower bound of, 107
 - upper bound of, 107
- ring-projection, 49, 344–346, 349, 350, 356, 367, 368
- scaling function, 107

- singularity, 20, 43
- skeleton, 311, 318, 322, 325, 327, 331
- CYM, 336, 337, 341
- ZSM, 336, 337, 341
- smoothness function, 285
- symmetry, 311, 314, 315, 318, 319
 - PISA, 311
 - SAT, 311
 - SLS, 311
- Tang-Yang wavelet, 284, 314–316
- texture, 62
 - analysis, 20, 61
 - classification, 20, 61, 62
 - color, 62
 - rotation-invariant, 62
 - segmentation, 66
 - statistical, 62, 64
- threshold, 204
- time window, 78
- time-frequency window, 78
 - area of, 80
- transform
 - Haar, 14, 16
- transition operator, 120
 - restriction of, 120
 - eigenvalue of, 123, 126
 - eigenvector of, 123, 126
 - spectral radius of, 120, 121, 123
- translation, 76
- two-scale equation, 106, 108
 - solution of, 112
- two-scale relation, 127
- uncertainty principle, 80
- upsampled, 151
- vanishing moment, 95, 98
- wavelet, 2
 - complex-valued, 22
 - descriptor, 20, 46–48, 56, 59
 - directional histograms, 30
- Gaussian, 99, 295
- Haar, 13, 14, 56, 101
- Mexico hat-like, 99
- nonlinear, 24
- quadratic splin, 251
- quadratic spline, 101, 214, 256–259, 263, 264, 267, 268, 273, 284, 295, 300
- spline, 100
- wavelet basis, 143
 - Battle-Lemaré-spline, 162
 - biorthonaoral, 144
 - cardinal spline, 181
 - compactly supported spline, 196
 - Daubechies, 166, 171
 - Haar, 156
 - Littlewood-Paley, 158
 - Meyer, 160
 - orthonormal, 154–156
- wavelet transform, 16, 22, 24, 27–29, 202, 205, 208, 209, 222, 228, 238, 239, 242, 244, 245, 251–253, 258, 263, 264, 271, 273, 276, 279–281, 285, 294, 299, 301, 305, 314–318, 322, 325, 327–329, 334, 415, 417, 420, 422, 424
 - calculation, 211
 - one-dimensional, 213
 - two-dimensional, 215
 - continuous, 75
 - integral, 75
 - inverse, 85, 88
 - modulus of, 203
 - local maximal, 203, 205
 - packets, 27
 - scale, 91, 253, 282, 283
 - width light-dependent, 251, 259, 267, 279
 - window function, 77
 - zero-crossing, 205