INFORMATION TO USERS

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

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

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

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

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

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

Wavelet Multiresolution and Applications

by

JIANXIN ZHU

A dissertation submitted to the Graduate Faculty in Engineering in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

1996

This manuscript has been read and accepted for the Graduate Faculty in Engineering in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Jan 16, 1996
Date

Hugh Barbe
Chair of Examining Committee

1/16/96
Date

Howard G. Lowen
Executive Officer

Prof. Samir Ahmed

Prof. Mitra Bash

Prof. Leonid Roytman

Prof. Michael Colef

Supervisory Committee

The City University of New York

Abstract

Wavelet Multiresolution and Applications

by

JIANXIN ZHU

Adviser: Professor Joseph Barba

Multiresolution analysis represents a function $f(x)$ at various levels of resolution. It is more efficient to analyze the image at resolution R_j, and then process the additional details available at the resolution R_{j+1}. Wavelet representation provides a coarse approximation of the signal plus the detail signals at the successive resolutions R_j for $1 \leq j \leq N$. In this dissertation, wavelet functions, wavelet bases and methods to construct wavelet bases are presented. The major contribution of this research is in the application of wavelet for voice pitch detection, cell contour extraction, and texture segmentation. In wavelet multiresolution voice application, detail signals on different scales are used to detect the voice signal pitch period. In cell image application, difference signals are used to extract cell contours. Edge enhancement method provide another way to extract cell contours. Different contour extraction algorithms are developed and the experimental results are demonstrated. Wavelet energy representation based on multichannel decomposition is defined. A pyramid tree structure and a quadtree structure are used to construct energy vectors which are used as the texture feature in segmentation. Test images, with structured/unstructured and rotated textures are applied to our segmentation algorithm. Experimental results are shown to demonstrate that the algorithms are very successful.

Contents

List of Tables

List of Figures

1 Introduction

1.1 Multiresolution Approximations

In an imaging system objects are displayed at different sizes and different gray levels depending on the distance of the object to the focal plane of the video camera and the lighting condition. In computer vision it is difficult to analyze the information content of an image directly from the image pixels. One image processing strategy is to process the image at different resolution scales, beginning with a coarse signal and increasing the resolution by adding more detail information until the signal is processed at a fine scale. If the processing method at each different resolution scale is identical, a fast, efficient algorithm can be developed. Multiresolution signal processing provides a simple hierarchical framework for interpreting the signal information. To realize multiresoultion signal processing, a scale space is created where each scale is at a different resolution R_j, $1 \leq j \leq N$, $R_j < R_{j+1}$. An image is decomposed into resolution levels, from coarse, R_1, to fine, R_N. The image information at resolution R_{j+1} contains the information at R_j, plus the additional detail signal. Because of the redundant information between levels, the total image information at each resolution level need not be analyzed each time; only the additional detail available at resolution R_{j+1}, after analyzing the signal at resolution R_j, need be processed. This approach simulates the human visual system, which starts with a coarse resolution, and adds more detail as we focus on particular objects. Multiresolution signal processing is very useful for pattern recognition, textual discrimination, image understanding, image reconstruction, stereo matching and template matching [1]. It has already been widely

studied in both one-dimensional and two-dimensional signal processing.

A simple multiresolution approximation is the Laplacian pyramid structure, introduced by Burt [2] and Crowley [3], using the Laplacian of the Gaussian. The approximation of the image on different levels is presented in Figure 1. The difference signal is the difference between two consecutive levels:

$$D_{2^j} f(x) = A_{2^{j+1}} f(x) - A_{2^j} f(x) \tag{1}$$

In the Laplacian pyramid, the signals on D_{2^j} do not correspond to the difference of information between $A_{2^{j+1}} f$ and $A_{2^j} f$ due to the correlation between the detail signals $D_{2^j} f$ at different resolutions. It is difficult to determine whether a similarity between the image details at different resolutions is due to a property of the image itself or to the intrinsic redundancy of the representation [4]. The details at each resolution 2^j are calculated by filtering the original image with the difference of two low-pass filters, then subsampling by 2^j. The approximation of the Laplacian of the Gaussian is given by the difference of the low-pass filters. The drawback of this algorithm is that the details at different resolutions are correlated. The number of samples representing the signal is increased by a factor of 2 in one dimension and by a factor of 4/3 in two dimensions.

The ideal multiresolution approximation at resolution 2^j should be most similar to the original signal at resolution 2^j, and should also contain all the approximation information at resolution 2^{j+1}. There should be no correlation between the difference signals and the approximation. The approximation operation is similar at all resolutions and is the same sampling length at all levels. Some information of $f(x)$ is lost when computing the approximation of $f(x)$ at resolution 2^j, but the approximation of $f(x)$ will converge to the original

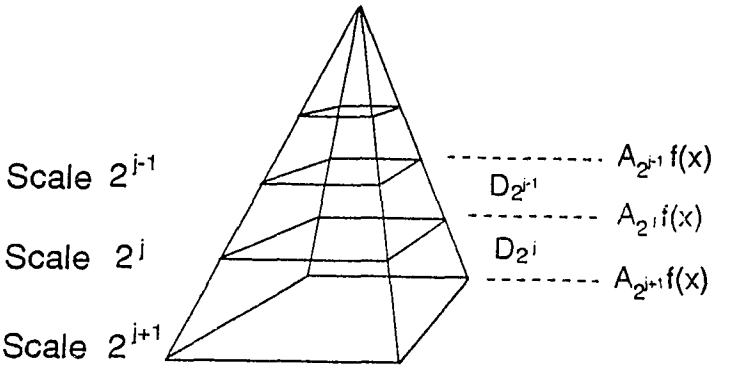


Figure 1: The approximation of the image and the difference signals of multiresolution signal processing

signal as the resolution increase to $+\infty$.

Wavelet transforms provide a new approach to realizing multiresolution pyramid expansion. It decomposes the signal into a coarse approximation, by lowpass filtering and subsampling, with a prediction error which is the difference between the original signal and the prediction based on the coarse version. The wavelet multiresolution approximation introduced by Mallat [1] [5] suppresses the correlation between the detail signals and represents the difference in information between $A_{2^{j+1}}f(x)$ and $A_{2^j}f(x)$. The wavelet multiresolution approximation, at resolution 2^j, projects $f(x)$ onto an orthogonal basis. The orthogonal basis is constructed by dilating and translating a particular function $\phi(x)$, called the scaling function.

Both the Fourier transform and the wavelet transform can be used for signal analysis. The short time Fourier transform retains many of the characteristics of the usual Fourier transform, such as a constant window size at all frequencies. However, the wavelet can focus on short duration phenomenons as the scale becomes smaller. This behavior allows a local characterization of functions which the Fourier transform does not. These properties of wavelet transforms led us in our research to use wavelet multiresolution pyramid expansion

to analyze the difference signals on different scales and to develop new approaches to explain the real physical meaning in signal processing.

1.2 About this dissertation

This dissertation is organized into 7 chapters including this introduction. A brief summary for each chapter is presented in the following:

The first part of this dissertation presents the concepts and theory of wavelet transform. Chapter 2 includes a brief review of the Fourier transform, short time Fourier transform, and the Gabor transform. The limitation of these transforms is their lack of local information. A Gabor window is used to explain the window function which is constant at all frequencies. Chapter 3 focuses on wavelet functions and wavelet bases. The definitions of wavelet function and scaling function are presented, and the properties of wavelet transform are described. Orthogonal wavelet, biorthogonal wavelet and the methods to construct wavelet bases are introduced. Some wavelet bases coefficients are listed and their figures in the time domain and frequency domain are given. Signal decomposition and reconstruction schemes are introduced. Wavelet window explanation is shown to illustrate that the width of the window is wider in low frequency and narrower in high frequency.

The second part of this dissertation is devoted to multiresolution applications. In chapter 4, voice pitch detection using wavelet transform is introduced. The first derivative of a smooth function can be used as a wavelet function to construct a wavelet basis. The difference signals on each scale can be used to explain the physical meaning of original signals. Voice pitch can be detected by using the local maximum of wavelet decomposition.

The 1-D application can be expanded to 2-D images. In chapter 5, cell contour extraction using the same strategy is introduced. Results are given and the strengths and weaknesses of this method are discussed. The multiresulotion edge representation allows us to develop a new approach to overcome the short point of using the local maximum to locate the cell edge. In chapter 6, a new multiresolution enhancement method for extracting the cell contours is developed. This method extracts the blurred cell contours successfully. The last chapter is devoted to texture segmentation. Multiresolution energy representation based on orthogonal wavelet transform is defined and the energy vector is used as a feature of texture. Test images with structured/unstructured texture and rotated texture are used in our algorithm. The results of segmentation demonstrate that the algorithm is successful.

1.3 Notation used in this dissertation

1. Z: Set of integer, Z = { ...,-1,0,1,...};

   R: Set of Real number ;

   N: Set of integer ≥ 0

2. Hilbert space $L^2(\mathbf{R})$, is the vector space of measurable, square-integrabel one-dimensional function $f(x)$.

3. $L^2(\mathbf{R}^2)$ is the vector space of measurable, square-integrable two dimensional function $f(x, y)$.

4. Inner production:

For $f(x) \in L^2(\mathbf{R})$ and $g(x) \in L^2(\mathbf{R})$,

$$< f(x), g(x) >= \int_{-\infty}^{+\infty} f(x)g(x)dx$$

For $f(x,y) \in L^2(\mathbf{R^2})$ and $g(x,y) \in L^2(\mathbf{R^2})$,

$$< f(x,y), g(x,y) >= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x,y)g(x,y)dxdy$$

5. The norm of $f(x)$ in $L^2(\mathbf{R})$

$$||f||^2 = \int_{-\infty}^{+\infty} |f(x)|^2 dx$$

6. Convolution:

For $f(x) \in L^2(\mathbf{R})$ and $g(x) \in L^2(\mathbf{R})$,

$$f * g(x) = (f(u) * g(u))(x) = \int_{-\infty}^{+\infty} f(u)g(x-u)du$$

7. Fourier Transform:

For $f(x) \in L^2(\mathbf{R})$, $\hat{f}(\omega)$ is the Fourier transform of $f(x)$,

$$\hat{f}(\omega) = \int_{-\infty}^{+\infty} f(x)e^{-i\omega x}dx$$

For $f(x,y) \in L^2(\mathbf{R^2})$, $\hat{f}(\omega_x, \omega_y)$ is the Fourier transform of $f(x,y)$

$$\hat{f}(\omega_x, \omega_y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x,y)e^{-i(\omega_x x + \omega_y y)}dxdy$$

8. The Parseval Identity

For all $f, g \in L^2(\mathbf{R})$,

$$< f, g >= \frac{1}{2\pi} < \hat{f}, \hat{g} >$$

9. Orthonormal Basis:

$$\int \psi_{j,k}(x)\psi_{j',k'}(x) = \begin{cases} 1 & \text{if } j = j', k = k' \\ 0 & \text{otherwise} \end{cases}$$

10. Down sampling:

A discrete signal $x(n)$ is down sampled by M, output $y(n)$ is every other M samples of $x(n)$.

$$y(n) = x(n) \downarrow M$$

11. Up sampling:

Insert $M - 1$ zeros between consecutive samples of $x(n)$.

$$y(n) = x(n) \uparrow M$$

12. Dilation and Translation: The dilation of $\psi(x)$ with a factor s and translation parameter b is

$$\psi_{s,b}(x) = \sqrt{s}\psi(s(x - b))$$

2 General Review of the Fourier Transform, STFT, and the Gabor Transform

2.1 Time Domain and Frequency Domain

Natural signals can be detected and converted into electrical signals by appropriate sensors, and the characteristics of the natural signals can be analyzed by processing the electrical signals. VLSI processors, powerful computers, fast supporting programs make accurate processing possible. Fast algorithms have been developed to reduce the complexity of the computation, shorten the computation time, and make signal processing more reliable and affordable. Parallel algorithm and multi-processors enhance signal processing performance. Any signal, detected and converted, based on time variation, is called signal in time domain which either continues or discrete. Signals represented or described by the frequency variation is called signal in frequency domain. These are the two general categories. Many signal analyses are developed based on these two representations, for example, the Fourier transform and the Harthey transform. These transforms play an important role not only in applied mathematics but also in engineering. They have significant physical interpretations both in the time and frequency domains and these methods have been successfully applied for many practical problems.

The Fourier transform of signal $f(x)$ is defined as:

$$\hat{f}(\omega) = \int_{-\infty}^{+\infty} f(t)e^{-i\omega t}dt \tag{2}$$

The Fourier transform coefficients are computed as inner products of the signal $f(t)$ with sinusoidal basis functions of infinite duration. Many algorithms have been developed which

make the computation of the Fourier transform more efficient [6],[7],[8]. Fourier analysis works well when $f(t)$ is a stationary signal. However, if $f(t)$ is nonstationary, its Fourier transform is spread out over the entire frequency domain, effectively masking its nonstationary character. For example, if $f(t) = \delta(t - t_0)$, the signal is non-zero at $t = t_0$, but its Fourier transform affects the entire spectral domain. Two approaches are used to overcome this drawback. One is to view the local signal through a window function $g(t)$, in which the signal is considered stationary. The short time Fourier transform uses the *local frequency* instead of the entire frequency domain. The second approach is to modify the sinusoidal basis function used in the Fourier transform, making this basis function more concentrated in time and less concentrated in the frequency domain.

2.2 Time-Frequency Joint Representation

The Gabor transform, the windowed Fourier Transform, and wavelet transform provide a time-frequency picture with good localization properties in both domains. Signals in the time domain are mapped into a joint Time-Frequency domain by these transforms. In time-frequency analysis, several important points must be considered, such as size of time-frequency window, computational complexity and efficiency, and implementation simplicity.

Gabor introduced a window function $g(t-b)$ in the time domain [9], where the parameter b is used to relocate the window to cover the entire time domain. This method is used to narrow the time variation by suppressing the signal outside the window and analyzing the *local spectrum* of the signal $f(t)$ in the time window period. A Gaussian function is commonly used in the Gabor transform since the Fourier transform of a Gaussian function is still a Gaussian function. For some other applications, different functions can also be used as window functions. This windowed Fourier transform is commonly called the Short-Time

Fourier Transform (STFT).

The STFT at time t is the Fourier transform of the signal $f(t)$ multiplied by a shifted analysis window $g(t - t_0)$ centered around t_0. The window $g(t - t_0)$ is relatively short compared with the signal $f(t)$. The STFT is a *local spectrum* of the signal $f(t)$ around the *analysis time* t and is influenced by the window function $g(t)$. The STFT has no ability to handle some applications which require a large and/or small windows since the window size is fixed. For some application a large window must be chosen in order to analyze desirable properties. However small details can not be analyzed.

For $f(t) \in L^2(\mathbf{R})$, the Gabor transform of $f(t)$ is defined by:

$$(\widehat{G_b^a})f(\omega) = \int_{-\infty}^{+\infty} f(t)g_a(t-b)e^{-i\omega t}dt \tag{3}$$

This interprets the Gabor transform as windowing the signal $f(t)$ by the window function g_a. To determine the width of the window function, an RMS (Root Mean Square) notion is employed:

$$\Delta_{g_a} = \frac{1}{||g_a||^a}\left[\int_{-\infty}^{+\infty} x^2 g_a^2(x)dx\right]^{1/2} \tag{4}$$

$g_a(t)$ is an even function and its center is 0. The width of the window function g_a is $2\Delta_{g_a}$. For $a > 0$ we have $\Delta_{g_a} = \sqrt{2}$ and the width of the window g_a is $2\sqrt{a}$.

The Gaussian function is defined as:

$$g_a(x) = \frac{1}{2\sqrt{a\pi}}e^{-\frac{x^2}{4a}} \tag{5}$$

The Fourier transform of a Gaussian is also a Gaussian. The width of the Gabor window is determined by the positive constant a. By setting $g_{b,w}^a(x) = g_a(x-b)e^{-i\omega x}$, the Gabor

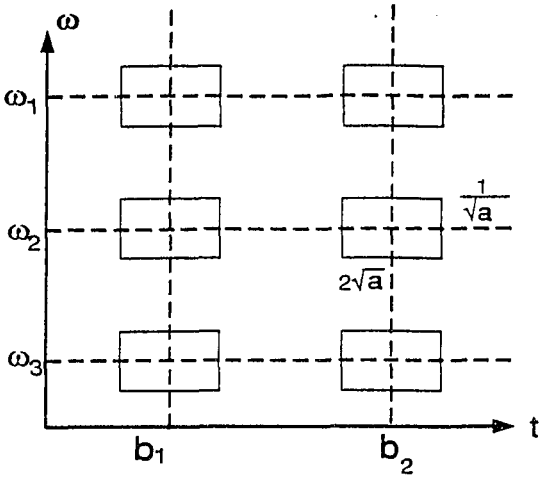


Figure 2: Gabor Transform Time-Frequency window

transform can be written as:

$$(\widehat{G_b^a f})(\omega) = \int_{-\infty}^{+\infty} f(x) g_{b,\omega}^a(x) dx = < f, g_{b,\omega}^a > \tag{6}$$

The time-frequency window uses the time window as the width and use the frequency window as the length. In the Gabor transform the width of the time-window due to $g_{b,\omega}^a$ is $2\sqrt{a}$ and the width of the frequency window of $g_{b,\omega}^a$ is $1/\sqrt{a}$, centered at ω. The rectangle time-frequency window area is:

$$\Delta(\omega)\Delta(\hat{\omega}) = [b - \sqrt{a}, b + \sqrt{a}] \times [\omega - \frac{1}{2\sqrt{a}}, \omega + \frac{1}{2\sqrt{a}}]$$

This is shown in figure 2.

The width of the Gabor transform window is unchanged regardless of its frequency or time location. However, to analyze high-frequency spectrum information, the time-interval should be relatively small; while for low-frequency spectrum, the time-interval should be relatively wide. This requires that the time-frequency analysis window be variable, automatically narrowing at high frequency and widening at low frequency.

3 Wavelet Bases and Wavelet Transform

3.1 Introduction

The wavelet transform was introduced by Grossmann and Morlet in 1984 [10]. The topic has been treated in considerable detail in the mathematics literature by Meyer [11], Stromberg [12], Daubechies [13], Mallat [1], [4], [5], Battle [14] and Lemarié [15]. Papers by Daubechies and by Mallat, particularly, have generated a great deal of activity in this area. Daubechies developed a systematic technique for generating finite-duration orthonormal wavelets, and established the connection between continuous-time orthonormal wavelets and the digital filter bank, which generated considerable interest not only in mathematics but also in signal processing [13], [16], [17], [18]. Instead of having constant time (space) and frequency domain convolutions, the wavelet transform uses short windows at high frequencies and long windows at low frequencies. The resolution of the wavelet transform varies with the scale parameter, making the wavelet transform appropriate for processing signals with features of different scales. The wavelet representation provides a multiresolution expression of a signal localization in both the time and frequency domains. Wavelet theory provides a unified framework for a number of techniques and is widely used in computer vision, subband coding, speech and image compression, texture classification and segmentation, and other signal processing fields [18] [19] [20] [21].

Any wavelet gives rise to some decomposition of the Hilbert space $L^2(\mathbf{R})$ into a direct sum of closed subspaces $W_j, j \in Z$; each subspace W_j is the closure in $L^2(\mathbf{R})$ of the linear span of the collection of functions. In real applications, a finite-energy signal is mapped into sample space; then, by applying the decomposition algorithm, the sampled signal can

be separated into wavelet components. The classic wavelets of Mayer, Battle-Lemarié, or Stromberg are based on Fourier-domain analysis. Daubechies constructed wavelets from iterated filter banks in time-domain. The regularity or smoothness of wavelets generated from iterated filters is one of the factors that should be considered. The expected wavelet should have a good cut-off frequency and be flat above cut-off frequency which corresponding to a smooth wavelet in the time domain. The more iterations performed when constructing a wavelet from filter banks, the smoother the wavelet. The smoothness, symmetry, and order of approximation should be also considered in the construction of the wavelet. B-spline function is one of the simplest functions which can be used in constructing the wavelet.

In contrast to the windowed Fourier transforms:

$$(T^{win}f)(\omega,t) = \int f(s)g(s-t)e^{-i\omega t}ds \tag{7}$$

the wavelet adapts the width of its time-slice according to the frequency components being extracted. The wavelet transform is:

$$(T^{wav}f)(a,b) = \int f(t)\psi(\frac{t-b}{a})dt \tag{8}$$

where ψ is the wavelet basis developed from the scaling function ϕ. The position of the signal window sampled by the wavelet in the time domain is controlled by parameter b, and the extent of the window is controlled by parameter a. The wavelet transform can also be written as an inner products in $L^2(\mathbf{R})$:

$$Wf(s,b) =< f(x), \psi_s(x-b) > \tag{9}$$

This corresponds to a decomposition of $f(x)$ on the family of functions $\psi_s(x-u)$. If

the family of $\psi_s(x-u)$ forms an orthogonal basis, the signal $f(x)$ is decomposed into an orthogonal basis and can be reconstructed using inverse wavelet transform. The Fourier Transform of $\psi_s(x)$ is given by:

$$\widehat{\psi_s}(\omega) = \frac{1}{\sqrt{s}}\widehat{\psi}(\frac{\omega}{s}) \tag{10}$$

The resolution of the wavelet transform varies with the scale parameter s, which is opposite to the window Fourier Transform which has a fixed resolution in the time and frequency domain.

3.2 Wavelet Basis Function and Wavelet Transform

Wavelet basis functions are obtained from a single prototype wavelet by dilations, scaling, and shifting. A function $\psi(x)$ is said to be a wavelet if and only if its Fourier transform $\hat{\psi}(\omega)$ satisfies:

$$\int_0^{+\infty} \frac{\hat{\psi}(\omega)}{\omega} d\omega = \int_{-\infty}^{0} \frac{\hat{\psi}(\omega)}{\omega} d\omega = C_\psi < \infty \tag{11}$$

This implies that $\int_{-\infty}^{+\infty} \psi(x)dx = 0$. $\psi(x)$ is the wavelet "prototype" which can be thought of as a bandpass function. The prototype $\psi(x)$ is called the basic wavelet, or simply Wavelet. The dilation of $\psi(x)$ with scale parameter a and translation parameter b is given by:

$$\psi_{a,b} = |a|^{-1/2}\psi(\frac{x-b}{a}) \tag{12}$$

The continuous wavelet transform of a function $f(x) \in L^2(\mathbf{R})$ is defined as:

$$W_\psi f(a,b) = |a|^{-(1/2)} \int_{-\infty}^{+\infty} f(x)\psi^*(\frac{x-b}{a})dx = < f, \psi^*_{a,b} > \tag{13}$$

where time x and time-scale a, b parameters vary continuously. $\psi^*_{a,b}(x)$ is the complex conjugate of $\psi_{a,b}(x)$. A wavelet series is given by:

$$f(x) = \sum_{j,k \in Z} C_{j,k} \psi_{j,k}(x) \tag{14}$$

The wavelet coefficients, analogous to the notion of the Fourier coefficient, are given by:

$$C_{j,k} =< f, \psi_{j,k} > \tag{15}$$

- The continuous time-scale parameters (a, b) in (13) can be sampled such that $a = a_0^j$ and $b = ka_0^j$. The original wavelets $\psi_{a,b}(x)$ become $\psi_{j,k}(x) = |a_0|^{-j/2}\psi(a_0^{-j}x - k), j, k \in Z$ and result in the discrete wavelet transform. Different values of j correspond to wavelets of different widths; the translation parameter b also depends on j. When j is high, the wavelet is narrow (high frequency) and translated in small steps to cover the whole time domain. When j is low, the wavelet is wider (lower frequency) and translated by larger steps. In computer vision and image processing the parameter a_0 is $a_0 = 2$. From a single prototype $\psi(x)$, by binary dilation 2^j and dyadic translation $k/2^j$, we get a wavelet family $\{\psi_{j,k}\}$:

$$\psi_{j,k}(x) = 2^{-j/2}\psi(2^{-j}x - k) \tag{16}$$

for which the $\psi_{j,k}(x)$ constitute an orthonormal basis. This result in the dyadic discrete wavelet transform(DWT):

$$W_{j,k} = CWT\{f(x); a = 2^j, b = k2^j\} = \int f(x)\psi^*_{j,k}(x)dx, \qquad j, k \in Z \tag{17}$$

which approximates $f(x)$ at resolution 2^j at location k. The dyadic discrete wavelet transform (DWT) has been popularized under the form of a signal decomposition onto a basis

of orthonormal wavelets.

A function $f(x)$ can be represented in terms of a set of wavelet coefficients, which covers all the scales and every location:

$$f(x) = \sum_{j \in Z} \sum_{k \in Z} W_{j,k} \psi_{j,k}(x) \tag{18}$$

This means that a function $f(x)$ can be decomposed onto a basis of orthonormal wavelet represented by a set of wavelet coefficients at various scale and locations. If $\hat{\psi}(\omega)$, the Fourier transform of $\psi(x)$, satisfies the condition:

$$\sum_{j=-\infty}^{+\infty} |\hat{\psi}(2^j \omega)|^2 = 1 \tag{19}$$

the function $f(x)$ can be reconstructed by superposition of all the synthesis wavelets. This ensures that the wavelet transform provides a complete representation covering all frequency axis. The similarity of the wavelet transform to the Gabor transform is that in both case the signal $f(x)$ can be represented by a set of time-varying functions; $f(x)$ is decomposed into a coarse signal and a set of detail signal on different scales. The difference is that $\psi_{a,b}$ have time-width adapted to their frequency. Figure 3 shows the shapes of wavelets $\psi_{a,b}$ when scale a and translation b changes. Figure 4 shows the Fourier transform of $\psi_{a,b}$. At high frequency, $\psi_{a,b}$ are very narrow while at low frequency $\psi_{a,b}$ are much broader. For a wavelet ψ with appropriate a_0, k, there exist $\tilde{\psi}_{j,k}$, for any function f in $L^2(\mathbf{R})$,

$$f(x) = \sum_{j \in Z, k \in Z} < f, \psi_{j,k} > \tilde{\psi}_{j,k} \tag{20}$$

Function f can be written as a superposition of $\psi_{j,k}$, and there exist $\tilde{\psi}_{j,k}$, f can be reconstructed in a numerically stable way from $< f, \tilde{\psi}_{j,k} >$.

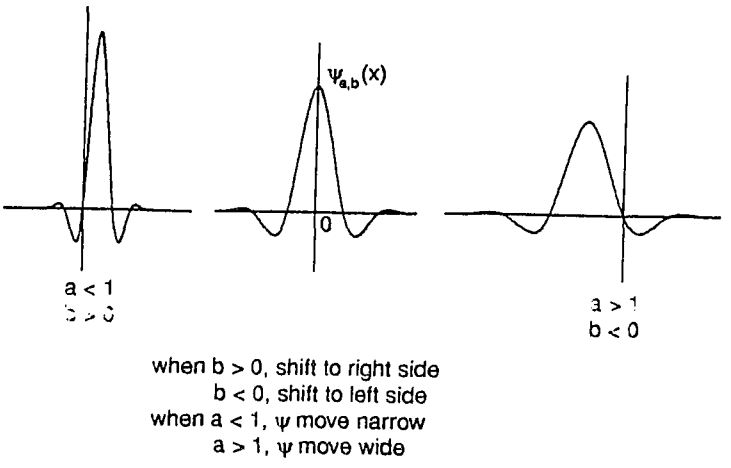


Figure 3: The shapes of $\psi_{a,b}$ when a and b changes

3.3 Wavelet Windows

We use ψ as the time domain window function and $\hat{\psi}$ as the frequency window function. The requirement for these window functions are: 1) that they both must decay sufficiently fast, and 2) that the window function must be possible to identify its center and width. If the center of the window function ψ is t^*, and the radius of ψ is Δ_ψ, then the function $\psi_{a,b}$ is a window function with its center at $b + at^*$ and the radius is Δ_ψ. Hence the wavelets transform gives local information of an analog signal $f(t)$ with a time-window:

$$[b + at^* - a\Delta_\psi, b + at^* + a\Delta_\psi] \tag{21}$$

In the frequency domain, the center and radius of the window function $\hat{\psi}$ are ω^* and $\Delta_{\hat{\psi}}$, respectively. Because $\hat{\psi}$ is the Fourier transform of time window function ψ, the radius of the Fourier transform of window is $\frac{1}{a}\Delta_{\hat{\psi}}$, the frequency window is given by:

$$[\frac{w^*}{a} - \frac{\Delta_{\hat{\psi}}}{a}, \frac{w^*}{a} + \frac{\Delta_{\hat{\psi}}}{a}] \tag{22}$$

The rectangular time-frequency window is:

$$[b + at^* - a\Delta_\psi, b + at^* + a\Delta_\psi] \times [\frac{w^*}{a} - \frac{\Delta_{\hat{\psi}}}{a}, \frac{w^*}{a} + \frac{\Delta_{\hat{\psi}}}{a}] \tag{23}$$

The ratio is:

$$\frac{[\text{Center Frequency}]}{[\text{Bandwidth}]} = \frac{w^*/a}{2\Delta_{\hat{\psi}}/a} = \frac{w^*}{2\Delta_{\hat{\psi}}} \tag{24}$$

Figure 5 gives a graphic representation of wavelet windows. The width of the window is $2\Delta_\psi$ in the time domain and $\frac{2\Delta_{\hat{\psi}}}{a}$ in the frequency domain. This window will be narrowed for detecting high-frequency phenomena, and made wider at low-frequency. Unlike the STFT where the time and frequency axes are typically uniform, in the wavelet transform the frequency window is narrow while the time window is wider or vice versa. Figure 6 shows the STFT and wavelet time-frequency windows.

3.4 Wavelet Frame

The theory of wavelet frames provides a general framework which covers the two extreme situations, redundancy and restrictions for the reconstruction scheme. Discrete wavelet transforms provide a very redundant description of the original function $f(x)$, it can only approximately represent the signal. If the redundancy is large (high oversampling) the basis functions are mildly restricted. If the redundancy is small (low sampling), the basis functions are very constrained. The question is, how do we select a level of redundancy while still obtaining a precise reconstruction of $f(x)$ from its wavelet transform?

Based on the choice of the basic wavelet $\psi(x)$ an orthonormal basis can be set up. However, not all types of $\psi(x)$ can be used to set up an orthogonal basis. There exist some wavelets $\psi(x) \in L^2(\mathbf{R})$ which can constitute orthonormal bases. These particular wavelets are called orthonormal wavelets. To reconstruct $f(x)$ from its wavelet coefficients $< f, \psi_{j,k} >$ and still have a numerically stable algorithm requires that the wavelet $\psi_{j,k}$ constitute a frame.

Let wavelet $\psi_{j,k}(x) = a_0^{-j/2}\psi(a_0^{-j}x - k)$, $j, k \in \mathbf{Z}$. A wavelet is an admissible basic wavelet if it constitute a frame with frame bounds A and B for $L^2(\mathbf{R})$. Then:

$$A \leq \int_0^{+\infty} C(a_0, k)\frac{|\hat{\psi}(\xi)|^2}{\xi}d\xi \leq B$$

and

$$A \leq \int_{-\infty}^{0} C(a_0, k)\frac{|\hat{\psi}(\xi)|^2}{\xi}d\xi \leq B$$

where $C(a_0, k) = 2n\pi/a_0 k$.

Not all choices for ψ, a_0, and k lead to frames of wavelet, even if ψ is admissible. Daubechies proved in [13] that if ψ decays reasonably fast in the time and frequency domain, and $\int \psi(x)dx = 0$, then there exists a range of a_0 and k, and the corresponding $\psi_{j,k}$ constitutes a frame. For this wavelet transform there exists positive constants A and B, with $0 < A \leq B < \infty$, such that:

$$AEx \leq \sum_{j,k \in Z} |C_{j,k}|^2 \leq BEx \tag{25}$$

A and B are called frame bounds and Ex is the energy of the signal. If the two frame bounds are equal, $A = B$, it is called a tight frame, and the bound is given by:

$$A = \frac{2n\pi}{k \ln a_0}\int_0^{+\infty}\frac{|\hat{\psi}(\xi)|^2}{\xi}d\xi = \frac{2n\pi}{k \ln a_0}\int_{-\infty}^{0}\frac{|\hat{\psi}(\xi)|^2}{\xi}d\xi \tag{26}$$

The stability condition on the basic wavelet ψ requires the ψ generates a *frame* of $L^2(\mathbf{R})$. Under this condition, any $f(x) \in L^2(\mathbf{R})$ can be recovered from its inverse wavelet transform(IWT).

$$\chi(x) = \frac{2}{(A+B)} \sum C_{j,k} f(x) \tag{27}$$

Basically, if ψ decays reasonably in the time and frequency domain, and $\int \psi(x)dx = 0$, then there exists a range of a_0 and k so the corresponding wavelet $\psi_{j,k}$ constitute a frame.

3.5 Classification of Wavelet Bases

Many families of orthogonal wavelets that have been constructed in $L^2(\mathbf{R})$. Their differences lie in their localization in physical and frequency space. Stromberg [12] constructed wavelets which have C^k continuity in 1982, where k is arbitrary but finite. These wavelets have exponential decay. Battle [14] and Lemarié [15] constructed identical families of orthonormal wavelets based on spline functions. These wavelets have exponential decay and are C^k continuous. In 1986, Mallat [5] developed the theory of multiresolution analysis, which provides a general framework for the construction of wavelets. Daubechies [13] provided the first general theory for the derivation of compactly supported wavelets in 1988. These wavelets are strictly localized in physical space while also retaining good localization in frequency space. Unlike the wavelets of Meyer, and Battle-Lemarié, these wavelets do not exhibit symmetry [22].

In 1985, Meyer [23] developed compactly supported wavelets with C^k continuity, where k is arbitrary and may be ∞. These wavelets are well localized in frequency space and physical space; decay is faster than any inverse polynomial, but they do not show exponential decay.

3.6 Orthogonal Wavelet Bases

3.6.1 Definition of Orthogonal Wavelet Bases

An orthonormal wavelet basis is defined as:

$$\int \psi_{j,k}(x)\psi_{j',k'}(x) = \begin{cases} 1 & \text{if } j = j', k = k' \\ 0 & \text{otherwise} \end{cases} \tag{28}$$

A multiresolution analysis is an increasing sequence $V_j, j \in Z$ of closed subspace of $L^2(\mathbf{R})$ with the following properties:

1) $\{0\} \subset ... \subset V_1 \subset V_0 \subset V_{-1} \subset ...$

2) $\bigcap_{j \in Z} V_j = \{0\}$

3) $\overline{\bigcup}_{j \in Z} V_j = L^2(\mathbf{R})$

4) $f(x) \in V_j \Longleftrightarrow f(2x) \in V_{j+1}$

5) $f(x) \in V_j \Longleftrightarrow f(x+n) \in V_j \qquad n \in Z$

6) There exists a scaling function $\phi \in V_0$,

its integer translates $\phi(x-k), k \in Z$ form a Riesz basis for the space V_0

$$V_0 = span \;\; \phi(x-k), k \in Z$$

$\phi(2x-k)$ form a basis for the space V_1,

$$V_1 = span \;\; \phi(2x-k), k \in Z$$

The scaling parameters do not have to be powers of 2; in computer vision and image processing, we are only interested in powers of 2.

Since the space V_0 is embedded within the space V_1, any function in V_0 can be expressed in terms of the basis function of V_1.

$$\phi(x) = \sum_{k=-\infty}^{\infty} h_n \phi(2x-k) \tag{29}$$

The scaling function ϕ and the square summable sequence h_n also have the following properties:

a) ϕ and its Fourier transform $\hat{\phi}$ have reasonable decay.

b) $\int \phi(x)dx \neq 0$

c) $h_n =< \phi, \phi_{-1,n} >, \quad \sum_n |h_n|^2 < \infty$

and

$$0 < A \leq \sum_{l \in Z} |\phi\ (\xi + 2\pi l)|^2 \leq B < \infty$$

$\phi_{0,k} = \phi(x-k), k \in Z$ form a Riesz basis for V_0, $\phi_{1,k} = \phi(2x-k), k \in Z$ form a Riesz basis for V_1. Define a function:

$$\phi_{m,k}(x) = 2^{\frac{m}{2}} \phi(2^m x - k) \tag{30}$$

$\phi_{m,k}(x), k \in Z$ forms a Riesz basis for the space V_m. The difference between subspace V_j and V_{j-1} is defined as W_{j-1}, which is the orthogonal complement of V_j in V_{j-1}.

$$V_{j-1} = V_j \oplus W_j, \qquad W_j \perp W_{j'} \qquad \text{if} \quad j \neq j' \tag{31}$$

where $\oplus$ represents a direct sum. We use $\psi(x)$ as a wavelet function and $\psi(x-k)$ forms a Riesz basis for the subspace W_0, which is the difference between subspace V_0 and V_1. Then:

$$\psi_{m,k}(x) = 2^{-\frac{m}{2}} \psi(2^{-m} x - k) \tag{32}$$

is a Riesz basis for W_m, the same as defined for V_m. In addition, if $\psi(x-k), k \in Z$ form an orthonormal set, then, $\psi_{m,k}, m, k \in Z$ form an orthonormal basis for $L^2(R)$.

All the subspace W_j are orthogonal, and the W_j spaces have the scaling property

$$f(x) \in W_j \longleftrightarrow f(2^j x) \in W_0$$

It follows that, for $j < J$:

$$V_j = V_J \oplus \bigoplus_{k=0}^{J-j-1} W_{J-k} \tag{33}$$

Because all the subspaces are orthogonal, the Hilbert space $L^2(\mathbf{R})$ can be decomposed into orthogonal subspace:

$$L^2(\mathbf{R}) = \overline{\bigcup_{j \in \mathbf{z}} V_j} = \oplus_{j \in \mathbf{Z}} W_j \tag{34}$$

Since the space W_0 is contained in the space W_{-1}, the wavelet function can also be expressed in terms of the scaling function:

$$\psi(x) = \sum_{k=-\infty}^{+\infty} g_k \phi(2x - k) \tag{35}$$

We use P_j as an orthogonal projection of f onto V_j, and Q_j as the projection of f on W_j, we have:

$$P_{m-1} f = P_m f + Q_m f \tag{36}$$

$P_m f$ is the approximation on scale m, and $Q_m f$ is the detail which can be added to the approximation on scale m in order to obtain the next finer scale approximation. This step can be iteratively processed. Every $f(x) \in L^2(\mathbf{R})$ projected onto subspace V_{j-1} can be represented as $f(x)$ projected onto subspace V_j and a set of detail information which is the difference between the version of $f(x)$ at two successive resolution levels.

$$P_{j-1} f = P_j f + \sum_{j,k \in Z} < f, \psi_{j,k} > \psi_{j,k} \tag{37}$$

The scaling function $\phi(x)$ plays a very important role. It is the starting point in creating an orthogonal wavelet basis. In general, we want to have the scaling function $\phi(x)$ as smooth as possible and well concentrated around 0 in the spatial domain. It should be more regular, and decay fast, and satisfy:

$$\left|\int_{-\infty}^{+\infty} \phi(x)dx\right| = \left|\hat{\phi}(0)\right| = 1 \tag{38}$$

A scaling function $\phi(x)$ can be written as:

$$\phi(x) = \sqrt{2}\sum_{n\in\mathbf{Z}} h(n)\phi(2x-n) \tag{39}$$

or

$$\hat{\phi}(\xi) = \frac{1}{\sqrt{2}}\sum_{n} h_n\hat{\phi}(\frac{\xi}{2})e^{-in\frac{\xi}{2}} \tag{40}$$

Let $m_0 = \frac{1}{\sqrt{2}}\sum_n h_n e^{-in\frac{\xi}{2}}$,

$$\hat{\phi}(\xi) = m_0(\frac{\xi}{2})\hat{\phi}(\frac{\xi}{2})e^{-in\frac{\xi}{2}} \tag{41}$$

m_0 is a 2π-periodic function in $L^2([0,2\pi])$, and

$$|m_0(\xi)|^2 + |m_0(\xi+\pi)|^2 = 1$$

After properly selecting the scaling function ϕ, V_0 can be constructed from $\phi(\bullet - k)$, and all V_j can be generated. V_j constitute a multiresolution analysis. The mother wavelet $\psi(x)$ associated with $\phi(x)$ can be constructed as

$$\hat{\psi}(\xi) = m_0(\frac{\xi}{2} + \pi)\hat{\phi}(\frac{\xi}{2})e^{in\frac{\xi}{2}} \tag{42}$$

or written as

$$\psi(x) = \sqrt{2}\sum_{n\in\mathbf{Z}} g(n)\phi(2x - n) \tag{43}$$

where $g(n) = (-1)^n h(1-n)$.

Let H and G represent the Fourier series of h and g. The decomposition and reconstruction requires:

$$H^* H \; + \; G^* G \; = \; 1 \tag{44}$$

where $*$ is the complex conjugate. In the decomposition the subspace should be orthogonal to remove correlations in the signal sequence:

$$H\, G^* = 0 \tag{45}$$

The filter coefficients $h(n)$ and $g(n)$ are fixed by the chosen multiresolution analysis framework. The $h(n)$ satisfy a normalization condition, $\sum_n h(n) = \sqrt{2}$, and $\sum g(n) = 0$, and:

$$\sum_{n\in\mathbf{Z}} h(n-2l)h(n-2k) \;=\; \delta_{l,k}, \tag{46}$$

$$\sum_{n\in\mathbf{Z}} g(n-2l)g(n-2k) \;=\; \delta_{l,k}, \tag{47}$$

$$\sum_{n\in\mathbf{Z}} h(n-2l)g(n-2k) \;=\; 0 \tag{48}$$

3.6.2 Construction of Orthogonal Wavelet Bases

In general, a scaling function ϕ is the solution to a dilation equation of the form:

$$\phi(x) = \sum h_n \phi(sx - k) \tag{49}$$

In image processing and computer vision, the convenient choice of the dilation factor is $s = 2$. The scaling function and its translates form an orthonormal set:

$$\int_{-\infty}^{+\infty} \phi(x)\phi(x + l)dx = \delta_{0,l}; \qquad l \in Z \tag{50}$$

where

$$\delta_{0,l} = \begin{cases} 1 & \text{if } l = 0 \\ 0 & \text{otherwise} \end{cases}$$

The scaling function should have the following properties:

1. $\int_{-\infty}^{+\infty} \phi(x)dx = 1, \qquad$ and $\qquad \sum h_n = \sqrt{2}.$

2. $\int_{-\infty}^{+\infty} \phi(x)\phi(x + l)dx = \delta_{0,l}, \qquad l \in Z.$

We are interested in finding scaling functions so that the wavelet bases of $L^2(\mathbf{R})$ can be defined in terms of these scaling functions. Let m to be a trigonometric polynomial, $m(0) = 1$. The infinite product of $m(2^{-i}\xi)$ defines an entire function $m(\xi)$ [13]:

$$m(\xi) = \prod_{i=1}^{\infty} m(2^{-i}\xi) \tag{51}$$

If m_0 satisfy:

$$|m_0(\xi)|^2 + |m_0(\xi + \pi)|^2 = 1 \tag{52}$$

m_0 is of the form:

$$m_N(\xi) = \left(\frac{1}{2}(1 + e^{i\xi})\right)^N Q(e^{i\xi}) \tag{53}$$

where $N \in \mathbf{N}$, $N \geq 1$. Q is a trigonometric polynomial satisfying:

$$|Q(e^{i\xi})|^2 = \sum_{k=0}^{N-1} \begin{pmatrix} N-1+k \\ k \end{pmatrix} \sin^{2k}(\frac{\xi}{2}) + \sin^{2N}(\frac{\xi}{2}) R(\cos(\frac{\xi}{2})) \tag{54}$$

where R is an odd polynomial. The polynomial $|Q(e^{i\xi})|^2$ can be written as a polynomial in $cos(\xi)$. Since $cos(\xi) = 1/2(1 + \sin^2 \frac{\xi}{2})$, it can also be a polynomial in $sin^2(\xi/2)$.

$$|Q(e^{i\xi})|^2 = \cos^{2N} \frac{\xi}{2} |Q(e^{i\xi})| \tag{55}$$

To construct m_0:

Step 1. Select N, where $N \in \mathbf{N}$, and $N \neq 0$, where $\mathbf{N}$ is the set of integer ≥ 0.

Step 2. Select an odd polynomial R. R cannot be completely free. Mostly set $R \equiv 0$.

Step 3. Compute roots $z_j, z_j^{-1}, \bar{z}_j, \bar{z}_j^{-1}$. If z_j is real, $z_j = r_j, z_j^{-1} = r_j^{-1}$, $j \in \mathbf{N}$ from the equation:

$$p(\xi) = \sum_{n=0}^{k-1} \begin{pmatrix} k+n-1 \\ n \end{pmatrix} (\sin^2(\frac{\xi}{2}))^n + (\sin^2(\frac{\xi}{2}))^k R(cos\xi) \tag{56}$$

Step 4. Select pairs of zeros from each quadruplet of complex zeros, select one zero from each duplex of real zeros. For the filter with minimal phase, retain all the zeros inside the unit circle. The Q_N will be:

$$Q_N(e^{i\xi}) = \sum_{n=0}^{N-1} q_N(n) e^{in\xi}, \quad \text{with } q_0 \neq 0 \tag{57}$$

The corresponding m_0 with selected N will be:

$$
\begin{aligned}
{}_N m_0(\xi) &= [\frac{1}{2}(1+e^{i\xi})]^N \sum_{n=0}^{N-1} q_N e^{in\xi} \\
&= 2^{-1/2} \sum_{n=0}^{2N-1} h_N(n) e^{in\xi} \qquad (58)
\end{aligned}
$$

Example:

1). When $N = 3$:

$$|Q_3(z)|^2 = \sum_{k=0}^{2} \left(\begin{array}{c} k+2 \\ k \end{array} \right) (\sin^2(\frac{\xi}{2}))^n$$

by setting $R \equiv 0$. Then the Q_3 is:

$$Q_3(\xi) = \frac{1}{4}[(1+\sqrt{10}+\sqrt{5+2\sqrt{10}}) + 2(1-\sqrt{10})e^{i\xi} + (1+\sqrt{10}-\sqrt{5+2\sqrt{10}})e^{2i\xi}]$$

the polynomial roots are: $(2.7127 + 1.4439i)$, $(2.7127 - 1.4439i)$, $(0.2873 + 0.1529i)$, and $(0.2873 - 0.1529i)$. Select the roots within the unit circle, which are $(0.2873 + 0.1529i)$ and $(0.2873 - 0.1529i)$. Compute the q_N and h_N based on eqations (57) and (58). One gets:

$$
\begin{aligned}
q_N &= \{0.1409, -0.7645, 0.13307\} \\
h_N &= \{0.3327, 0.8069, -0.1350, -0.0854, 0.0352\}
\end{aligned}
$$

In sec. 3.5.3, coefficients of $h_N(n)$ are listed for the cases $N = 4, 20$, and the figures of ${}_N\phi$, ${}_N\psi$ are shown.

An orthonormal basis of compactly supported wavelet can never be an even function. Only the Haar basis is the compactly supported wavelet which ϕ_{Haar} has a symmetry axis. If m_0 is trigonometric polynomial.

$$m_0(\xi) = e^{2i\lambda\xi} m_0(-\xi) \tag{59}$$

the possible symmetry point can only be $\lambda = 1/2$ in $]0, 1[$.

If a function $\phi \in L^2(\mathbf{R})$ generates a nested sequence of closed subspace V_j and is framed, it is said that the function ϕ generates a multiresolution analysis. The function ϕ is called a scaling function. Polynomial spline bases are often used as scaling function because they have a simple explicit analytic form. They are smooth and easy to manipulate. The standard basis functions, B-spline, are compactly supported and can be generated by repeated convolutions of a rectangular pulse. A fundamental result according to Stromberg is that any polynomial spline function of degree n can be represented by a linear combination of shifted B-splines [12].

$$g^n(x) = \sum_{k=-\infty}^{+\infty} c(u)\beta^n(x-k) \tag{60}$$

where β^n is B-spline polynomial in order n. β^n is recursively convoluted by B-spline in order 0.

3.6.3 Examples: ϕ and ψ for N

In this section we list some wavelet coefficents and give their figures in the time and frequence domains. The first wavelet, the Haar wavelet and scaling function are shown in Figure 7. The Haar wavelet can be viewed as the smallest degree Battle-Lemarié wavelet

($\psi_{Haar} = \psi_{B-L,0}$) or also as the first of Daubechies compactly supported wavelets ($\psi_{Haar} =_1 \psi$).

Table 1 are listed Daubechies' N=4, N=20 orthogonal wavelets coefficients, the scaling functions and wavelets in time domain and frequency domain are shown in Figures 8, 9, and 10. Figures 11 and 12 are shown the Daubechies' vanish moment and nearly symmetric wavelets [24]. Coefficients of Battle-Lemariés' scaling function and wavelets are listed in Table 2. Figures 13 and 14 are shown two of Battle-Lemarié wavelet in the time and frequency domains.

3.6.4 Orthogonal Wavelet Decomposition and Reconstruction

Decomposition Scheme:

Let ψ be any wavelet, $\psi_{j,k}$ is a wavelet basis. Let W_j denote the linear span of $\psi_{j,k}$; $k \in Z$. We use $\dot{+}$ as the direct sum and $\oplus$ as the orthogonal sum. $L^2(\mathbf{R})$ can be decomposed as a direct sum of the spaces W_j:

$$L^2(\mathbf{R}) = \dot{\sum_{j\in Z}} \mathbf{W_j} = \cdots \dot{+} \mathbf{W_{-1}} \dot{+} \mathbf{W_0} \dot{+} \mathbf{W_1} \dot{+} \cdots \tag{61}$$

in the sense that every function $f \in L^2(\mathbf{R})$ has a unique decomposition:

$$f(x) = \cdots + g_{-1}(x) + g_0(x) + g_1(x) + \cdots \tag{62}$$

where $g_j \in W_j$ for all $j \in Z$.

If ψ is an orthogonal wavelet, then the subspaces W_j of $L^2(\mathbf{R})$ are orthogonal.

$$W_j \perp W_l, \qquad j \neq l$$

Then $< g_j, g_l >= 0$, $j \neq l$ where $g_j \in W_j, g_l \in W_l$. Consequently, the direct sum becomes an orthogonal sum:

$$L^2(\mathbf{R}) = \bigoplus_{j \in Z} W_j = \cdots \oplus W_{-1} \oplus W_0 \oplus W_1 \oplus \cdots \tag{63}$$

Any $f \in L^2(\mathbf{R})$ can be decomposed into a sum of functions $g_j \in W_j$, which is not only unique but also orthogonal. V_j is generated by scaling function $\phi \in L^2(\mathbf{R})$, W_j is generated by wavelet $\psi \in L^2(R)$, and W_j are the orthogonal complement of V_j in V_{j-1}.

From the nested sequence of spline subspace V_j, the orthogonal complementary subspace W_j,

$$V_{j-1} = V_j \oplus W_j, \qquad j \in Z$$

For any $f \in L^2(\mathbf{R})$, $f(x)$ can be approximated into $f_N \in V_N$ where $N \in Z$.

$$V_N = V_{N-1} \dot{+} W_{N-1} \tag{64}$$

$$f_N = f_{N-1} \dot{+} g_{N-1} \tag{65}$$

where $f_{N-1} \in V_{N-1}$, and $g_{N-1} \in W_{N-1}$. By repeating this process, choosing M so that f_{N-M} is sufficiently "blurred" in terms of the variation, we get:

$$f_N = g_{N-1} \dot{+} g_{N-2} \dot{+} \cdots \dot{+} g_{N-M} \dot{+} f_{N-M} \qquad \text{where} \qquad f_j \in V_j, g_j \in W_j$$

If P_j denotes the orthogonal projection onto V_j, and Q_j denotes the orthogonal projection onto W_j, a decomposition can be written as:

$$P_{j-1}f = P_jf + Q_jf \tag{66}$$

The difference between two successive approximations is

$$Q_{j+1}f = P_jf - P_{j+1}f \tag{67}$$

At the first scale, when $m = 1$, $V_0 = V_1 \oplus W_1$, and $f = P_1f + Q_1f$. Each of these components can be expanded with respect to the orthonormal bases $\phi_{1,n}$ and $\psi_{1,n}$, respectively:

$$P_1f = \sum_k c_{1,k}\phi_{1,k} \tag{68}$$

$$Q_1f = \sum_k d_{1,k}\psi_{1,k} \tag{69}$$

where $c_{1,k} =< \phi_{1,k}, P_1f >=< \phi_{1,k}, f >= \sum c_{0,n} < \phi_{1,k}, \phi_{0,n} >$.

The sequence $c_{1,k}, k \in Z$, represents a smoothed version of the original data sequence c_0. $d_{1,k}, k \in Z$, represents the difference in information between c_0 and c_1.

$$\begin{aligned} < \phi_{1,k}, \phi_{0,n} > &= 2^{-1/2} \int \phi(\frac{x}{2} - k)\phi(x - n)dx \\ &= 2^{-1/2} \int \phi(\frac{x}{2})\phi(x - (n - 2k))dx \end{aligned} \tag{70}$$

Let $h(n) = 2^{-1/2} \int \phi(\frac{x}{2})\phi(x - n)dx$, then:

$$< \phi_{1,k}, \phi_{0,n} >= h(n - 2k) \tag{71}$$

We get:

$$c_{1,k} = \sum_n h(n-2k)c_{0,n} \tag{72}$$

Similarly, let $g(n) = 2^{-1/2}\int \psi(\frac{x}{2})\phi(x-n)dx$, and we can get $d_{1,k} = \sum_n g(n-2k)c_{0,n}$. Rewrite these two equations:

$$c_1 = Hc_0 \tag{73}$$

$$d_1 = Gc_0 \tag{74}$$

The iterated formula is:

$$P_{j-1}f = P_jf + Q_jf = \sum_k c_{j,k}\phi_{j,k} + \sum_k d_{j,k}\psi_{j,k} \tag{75}$$

and

$$c_j = Hc_{j-1} \tag{76}$$

$$d_j = Gc_{j-1} \tag{77}$$

The successive c_j are lower and lower resolution versions of the original c_0 , and the d_j contain the difference in information between c_{j-1} and c_j . Compared with the Laplacian pyramid scheme introduced by P.Burt and E.Adelson [2], one can make the following notes. 1) The computation of the wavelet analysis is as easy as the Laplacian pyramid scheme; the data are decomposed into a *Pyramid* of approximations, corresponding to less and less

detail. The difference between each of two successive approximations are computed. 2) Wavelet decomposition is more economical than in the Laplican Pyramid. 3) The difference is in wavelet, the blurred lower resolution c_j and difference sequence d_j are obtained via a lowpass filter and bandpass filter. When multiresolution analysis framework is chosen, the filters $h(n)$ and $g(n)$ are fixed. There is no need to recalculate the $h(n)$ and $g(n)$ on different scales. The one-dimension decomposition and reconstruction flow chart is given in Figure 15(a). Figure 15(b) shows the spectrum-spliting performance by the filter bank. There is a "finest" resolution associated with the space V_0. After a finit steps j decomposition, it leads to a "coarsest" resolution associated with V_j. V_j's are called approximation spaces and W_j's are called detail space. In Figure 16, shows a ideal division of the spectrum by the discrete-time wavelet series. The spectrums are symmetric around zero. The original signal is lowpass filtered using the ideal half-band filter. Starting from the space V_0, we have derived a lower-resolution signal by having V_0, resulting in V_1, then coarser version is obtained by using the same process. The final coarse (approximation) space V_j can be obtained, and difference space W_j have been created. A one-deminsional signals decomposition and reconstruction example is shown in Figure 17(a) and (b). In this example, a segment of signals with 512 samples is decomposed in three scales. The original signals and the low passed signals on each scales are shown on the left column. The right column shows the wavelet transform difference signals. The reconstructed signal is shown in Figure 17(c). In Figure 17(d) is the errors between the original signals with the reconstructed signals.

3.6.5 Decomposition and Reconstruction in Two-Dimensional Wavelet

To construct two dimensional wavelets, one can select wavelet bases obtained in one-dimension, and apply tensor products, or use the same analogue of the construction in

the one-dimension, starting with an arbitrary multiresolution approximation. The two dimensional wavelets are given by the following expression:

$$2^j\psi(2^jx-k,2^jy-l), \qquad j,k,l\in \mathbf{Z} \tag{78}$$

and the scaling function in two dimensional wavelet can be written as:

$$2^j\phi(2^jx-k,2^jy-l), \qquad j,k,l\in \mathbf{Z} \tag{79}$$

. A tensor products $\otimes$ is used here to expand the one dimensional wavelet into a two dimensional wavelet. Let $\vec{V}_j$ be the vector subspace of $L^2(\mathbf{R^2})$, $\vec{V}_j = V_j \otimes V_j$. Let $\vec{\Phi}(x,y) = \phi(x)\phi(y)$. An orthogonal basis of $\vec{V}_0$ is then made up of the production $\phi(x-k)\phi(y-l), (k,l)\in \mathbf{Z}^2$. Let $\vec{W}_1$ denote the orthogonal complement of $\vec{V}_1$ in $\vec{V}_0$. Then, as in one dimension, $\vec{V}_j$ is generated by the scaling function $\vec{\Phi}_j, j\in L^2(\mathbf{R^2})$:

$$\vec{V}_0 = \vec{V}_1 \oplus \vec{W}_1 \tag{80}$$

where $\vec{W}_1 = \overline{(V_1\otimes W_1)} \oplus \overline{(W_1\otimes V_1)} \oplus \overline{(W_1\otimes W_1)}$.

The family of functions $\{\vec{\Phi}_{j,k}, j,k\in Z\}$ is an orthonormal basis of $\vec{V}_j$, and $\{\vec{\Psi}_{j,k}, j,k\in Z\}$ is an orthonormal basis of $\vec{W}_j$. Hence the functions $\vec{\Psi}^{(i)}_{j,k} = 2^j\Psi^{(i)}(2^jx-k)$ are an orthonormal basis of $\vec{W}_j$, where:

$$\vec{\Psi}^{(1)}(x,y) = \psi(x)\phi(y) \tag{81}$$

$$\vec{\Psi}^{(2)}(x,y) = \phi(x)\psi(y) \tag{82}$$

$$\vec{\Psi}^{(3)}(x,y) = \phi(x)\phi(y) \tag{83}$$

are the orthogonal bases of $\overline{(V_0 \otimes W_0)}$, $\overline{(W_1 \otimes V_1)}$, and $\overline{(W_1 \otimes W_1)}$, respectively.

These functions are given by separable products of the functions ϕ and ψ. The difference of information between two successive scales is equal to the orthonormal projection $f(x, y)$ on W_j. The wavelet decomposition can be interpreted as a signal decomposition in a set of independent spatially oriented frequency channels. The filters G and H are the same as defined in one dimension. We use H_c, G_c, H_r, and G_r, c for columns, and r for rows. The c_0 is decomposed into c_1 and $d_{1,1}$, $d_{2,1}$, $d_{3,1}$, where $d_{(i),1}$ corresponds to the $\psi^{(i)}$.

$$c_1 = H_c H_r c_0 \tag{84}$$

$$d_{1,1} = H_c G_r c_0 \tag{85}$$

$$d_{2,1} = G_c H_r c_0 \tag{86}$$

$$d_{3,1} = G_c G_r c_0 \tag{87}$$

Figure 18(a) shows the two dimensional decomposition scheme.

To reconstruct f from its wavelet coefficients, the wavelet $\{\psi_{j,k}, j, k \in Z\}$ must satisfy (25). If we know c_j and d_j, the

$$P_{j-1} = P_j f + Q_j f = \sum_k c_{j,k}\phi_{j,k} + \sum_k d_{j,k}\psi_{j,k} \tag{88}$$

hence

$$\begin{aligned} c_{j-1,n} &= < \phi_{j-1,n}, P_{j-1} f > \\ &= \sum_k c_{j,k} < \phi_{j-1,n}, \phi_{j,k} > + \sum_k d_{j,k} < \phi_{j-1,n}, \psi_{j,k} > \end{aligned}$$

$$= \sum_k h(n-2k)c_{j,k} + \sum_k g(n-2k)d_{j,k} \tag{89}$$

or

$$c_{j-1} = H^* c_j + G^* d_j$$

where * denote the conjugate. The two-dimensional reconstruction scheme is shown in Figure 18(b). Figure 19 shows an image and the first two stages in the decomposition of the image. At each stage, the result is an averaged image and three detail component for horizontal, vertical, and diagonal details. As an application example, image compression using wavelet transform accompanied with vector quantization scheme is given in Figure 20. Different bit rate for different scales and details image can be applied (shown in Figure 20(a)), and multiresolution codebook is used for each subimage. Each sub-codebook has a low distortion level and contains few words, which clearly facilitates the search for the best coding vector (shown in Figure 20(b)).

3.6.6 Wavelet with Compact Support

Daubechies used iterated filter banks to construct wavelets in the time domain. The easiest way to ensure compact support for the wavelet ψ is to choose the scaling function ϕ with compact support. When the scaling function $\phi(x)$ is not 0 in $[-b, b]$, where b is a constant, the wavelet has compact support. The wavelet coefficient $h(n)$ has finite length, that is, $h(n) = 0$ for $n < K_-$ and $n > K_+$. As regularity r increases, the supports of ϕ and ψ also increase. Daubechies [13] shows that if there existd a constant C for each $r \geq 1$, the supports of ϕ and ψ are contained in $[-Cr, Cr]$.

3.7 Biorthogonal Wavelet Bases

3.7.1 Definition of Biorthogonal Wavelet Bases

In orthogonal wavelet transform, the scaling function ϕ in V_0 constructs an orthogonal basis, and the wavelet is generated in terms of the scaling function.

$$\phi(x) = 2\sum_{n\in Z} h_n \phi(2x-n) \tag{90}$$

$$\psi(x) = 2\sum_{n\in Z} g_n \phi(2x-n), \qquad \text{where} \qquad g_n = (-1)^n h(1-n) \tag{91}$$

The Fourier Transform of these two formulas are:

$$\hat{\phi}(2\omega) = m_0(\omega)\hat{2}\phi(2x-n), \qquad \text{where} \qquad m_0(\omega) = \sum_n h_n e^{-inw} \tag{92}$$

$$\hat{\psi}(2\omega) = \overline{m_0(\omega+\pi)}\hat{\phi}(\omega)e^{-iw} \tag{93}$$

where m_0 is a 2π periodic function, and satisfies:

$$|m_0(\omega)|^2 + |m_0(\omega+\pi)|^2 = 1 \tag{94}$$

$$m_0(0) = 1 \qquad \text{and} \qquad m_0(\pi) = 0 \tag{95}$$

A smooth scaling function and wavelet are interested in many applications, in which the scaling function can be written in the form of

$$\hat{\phi}(\omega) = \prod_{k=1}^{+\infty} m_0(2^{-k}\omega) \tag{96}$$

In some signal processing which starts from a low-scale signal, such as data compression, which is in the aspect of the low scale components to reconstruct the signal, the smoothness of ϕ and ψ are more useful. The limitation of the above formula is the infinite iteration number if $m_0(\omega)$ is selected from the set of trigonometric polynomials. Since they correspond to the finite impulse response (FIR) filter, they may get a high regularity of ϕ and ψ with careful selection. But orthogonal wavelet filters h and g can not be both FIR and linear phase, i.e. with real and symmetrical coefficient, except for the Haar Basis. But the Haar Basis is not continuous and is not very useful. But one still can decompose and reconstruct the signal using a pair of dual filters $m_0, \tilde{m}_0$, which satisfy:

$$m_0(\xi)\overline{\tilde{m}_0(\xi)} + m_0(\xi+\pi)\overline{\tilde{m}_0(\xi+\pi)} = 1 \tag{97}$$

For any function f in $L^2(\mathbf{R})$, the decomposition of f

$$\hat{f} = \sum_{j,k\in Z} < f, \tilde{\psi}_{j,k} > \psi_{j,k} = \sum_{j,k\in Z} < f, \psi_{j,k} > \tilde{\psi}_{j,k} \tag{98}$$

A dual scaling function and dual wavelet are defined by:

$$\hat{\phi}(\omega) = \prod_{k=1}^{+\infty} m_0(2^{-k}\omega) \tag{99}$$

$$\hat{\tilde{\phi}}(\omega) = \prod_{k=1}^{+\infty} \tilde{m}_0(2^{-k}\omega) \tag{100}$$

$$\hat{\phi}(2\omega) = m_1(\omega)\hat{\phi}(\omega) \tag{101}$$

$$\hat{\tilde{\phi}}(2\omega) = \tilde{m}_1(\omega)\hat{\tilde{\phi}}(\omega) \tag{102}$$

where $m_1 = \overline{\tilde{m}_0(\omega+\pi)}e^{-iw}$, and $\tilde{m}_1 = \overline{m_0(\omega+\pi)}e^{-iw}$. The dual trigonometric polynomials m_0 and $\tilde{m}_0$ satisfy:

$$m_0(0) = \tilde{m}_0(0) = 1, \qquad \text{and} \qquad m_0(\pi) = \tilde{m}_0(\pi) = 0$$

3.7.2 Construction of Biorthogonal Wavelet Bases

In biorthogonal wavelet bases, the dual filters associated with ψ and $\tilde{\psi}$ generate a pair of biorthogonal Riesz bases. Also, they should meet the conditions for the stability of the Fast Wavelet Transform. Define functions such as:

$$m_0(\xi) = 2^{-1/2}\sum_n h_n e^{in\xi} \tag{103}$$

$$\tilde{m}_0(\xi) = 2^{-1/2}\sum_n \tilde{h}_n e^{in\xi} \tag{104}$$

$$\hat{\phi}(\xi) = (2\pi)^{-1/2}\prod_{j=1}^{\infty} m_0(2^{-j}\xi) \tag{105}$$

$$\hat{\tilde{\phi}}(\xi) = (2\pi)^{-1/2}\prod_{j=1}^{\infty} \tilde{m}_0(2^{-j}\xi) \tag{106}$$

The infinite products can only converge if:

$$m_0(0) = \tilde{m}_0(0) = 1$$

or

$$\sum h_n = \sum \tilde{h}_n = \sqrt{2}$$

The dual scaling function can be written as:

$$\hat{\phi}(\xi) = m_0(\xi/2)\hat{\phi}(\xi/2) \tag{107}$$

$$\hat{\tilde{\phi}}(\xi) = \tilde{m}_0(\xi/2)\hat{\tilde{\phi}}(\xi/2) \tag{108}$$

or:

$$\phi(x) = \sqrt{2}\sum h(n)\phi(2x-n) \tag{109}$$

$$\tilde{\phi}(x) = \sqrt{2}\sum \tilde{h}(n)\phi(2x-n) \tag{110}$$

$$\psi(x) = \sqrt{2}\sum g(n)\phi(2x-n) \tag{111}$$

$$\tilde{\psi}(x) = \sqrt{2}\sum \tilde{g}(n)\tilde{\phi}(2x-n) \tag{112}$$

where $g(n) = (-1)^n h(1-n)$, and $\tilde{g}(n) = (-1)^n \tilde{h}(1-n)$. $\psi_{j,k}$ and $\tilde{\psi}_{j,k}$ constitute two dual Riesz bases:

$$< \psi_{j,k}, \tilde{\psi}_{j',k'} >= \delta_{j,j'}\delta_{k,k'}, \qquad \text{and} \qquad \int \phi(x)\tilde{\phi}(x-k)dx = \delta_{0,k}$$

When $m_0(-\xi) = m_0(\xi)$, or m_0 is a polynomial in $\cos\xi$, the filter associated with m_0 is linear phase:

$$m_0 = e^{-i\frac{\xi}{2}}\cos\frac{\xi}{2} \tag{113}$$

In the biorthogonal case, one can select m_0 to create a linear phase filter, or a symmetric scaling function ϕ. For real h_n, the filter associated with m_0 can be written as a polynomial in $\cos\xi$, $m_0 = p(\cos\xi)$, or:

$$m_0(-\xi) = m_0(\xi)$$

When it is symmetric around $x = 1/2$, the trigonometric polynomial can be written as:

$$m_0(-\xi) = e^{-i\frac{\xi}{2}} \cos(\frac{\xi}{2}) m_0(\xi)$$

Given m_0, we need to determine the $\tilde{m}_0(\xi)$ which satisfy:

$$m_0(\xi)\overline{\tilde{m}_0(\xi)} + m_0(\xi+\pi)\overline{\tilde{m}_0(\xi+\pi)} = 1 \tag{114}$$

when $m_0(\xi)$ is even:

$$m_0(\xi) = \cos^{2l}(\frac{\xi}{2}) p_0(\cos\xi) \tag{115}$$

when $m_0(\xi)$ is odd:

$$m_0(\xi) = e^{-i\frac{\xi}{2}} \cos^{2l+1}(\frac{\xi}{2}) p_0(\cos\xi) \tag{116}$$

where p_0 is a polynomial and $p_0 \neq 0, l \in N$

If there exist $\tilde{m}_0$ satisfying (114), then $\tilde{m}_0$ can have the same form as m_0.

when $\tilde{m}_0(\xi)$ is even:

$$\tilde{m}_0(\xi) = \cos^{2l}(\frac{\xi}{2}) \tilde{p}_0(\cos\xi) \tag{117}$$

when $\tilde{m}_0(\xi)$ is odd:

$$\tilde{m}_0(\xi) = e^{-i\frac{\xi}{2}} \cos^{2l+1}(\frac{\xi}{2}) \tilde{p}_0(\cos\xi) \tag{118}$$

Substitute m_0 and $\tilde{m}_0$ into (114):

$$\cos^{2k}(\frac{\xi}{2}) p_0(\cos\xi)\tilde{p}_0(\cos\xi) + \sin^{2k}(\frac{\xi}{2}) p_0(-\cos\xi)\tilde{p}_0(-\cos\xi) = 1 \tag{119}$$

Let $P(\sin^2(\frac{\xi}{2})) = p_0(\cos\xi)\tilde{p_0}(\cos\xi)$, for $(\frac{1-\cos\xi}{2}) = \sin^2(\frac{\xi}{2})$, $p_0(-\cos\xi)\tilde{p_0}(-\cos\xi)$ can be written as a polynomial in $\cos^2(\frac{\xi}{2})$, $P(\cos^2(\frac{\xi}{2}))$. Based on Bezout's theorem [16]:

$$(1-x)^k p(x) + x^k p(1-x) = 1 \tag{120}$$

One can get:

$$p(x) = (1-x)^{-k} - x^k(1-x)^{-k}p(1-x) \tag{121}$$

If the right side is expanded into a Taylor series, the general form is:

$$p(x) = \sum_{n=0}^{k-1} \begin{pmatrix} k+n-1 \\ n \end{pmatrix} x^n + x^k R(1-x) \tag{122}$$

where R is an odd polynomial.

$$p_0(cos\xi)\tilde{p_0}(cos\xi) = \sum_{n=0}^{k-1} \begin{pmatrix} k+n-1 \\ n \end{pmatrix} (\sin^2(\frac{\xi}{2}))^n + (\sin^2(\frac{\xi}{2}))^k R(cos\xi) \tag{123}$$

where $k = l + \tilde{l}$ when m_0 is even and $k = l + \tilde{l} + 1$ when m_0 is odd, and R is an odd polynomial.

To construct B-spline function ${}_N\phi$ in the order of N, the Fourier transform of ${}_N\phi$ is given by:

$${}_N\hat{\phi}(\xi) = e^{-ik\frac{\xi}{2}} \left(\frac{\sin(\frac{\xi}{2})}{\frac{\xi}{2}}\right)^N = e^{i\xi\lfloor N/2 \rfloor} \left(\frac{1-e^{-i\xi}}{i\xi}\right)^N \tag{124}$$

where $k = 0$ if N is even, $k = 1$ if N is odd.

The ${}_N m_0$ are given by the formula:

$${}_N m_0(\xi) = \left(\frac{1+e^{-i\xi}}{2}\right)^N e^{i\xi\lfloor N/2 \rfloor} = \sum_{n=-\lfloor N/2 \rfloor}^{N-\lfloor N/2 \rfloor} 2^{-N} \left(\frac{N}{n+\lfloor N/2 \rfloor}\right) e^{-in\xi} \tag{125}$$

The ${}_{N,\tilde{N}}\tilde{m}_0$ are given by the formula:

$$ {}_{N,\tilde{N}}\tilde{m}_0 = e^{ik\frac{\xi}{2}} \left(\cos(\frac{\xi}{2})\right)^{\tilde{N}} \left[\sum_{n=0}^{k-1} \begin{pmatrix} k+n-1 \\ n \end{pmatrix} (\sin(\frac{\xi}{2}))^{2n} + (\sin(\frac{\xi}{2}))^{2k} R(cos\xi)\right] \quad (126) $$

where $\tilde{N} \geq 1, N + \tilde{N} = 2k$ is even, R is an odd polynomial, and $k = 1$ if N is odd, $k = 0$ if N is even. Generally, R is set to zero for restricting the complex computation.

3.7.3 Biorthogonal Wavelet Bases Examples: ϕ, $\tilde{\phi}$, ψ and $\tilde{\psi}$ for N and $\tilde{N}$

In Table 3, we lave listed the Daubiches biothogonal coefficients of ${}_N m_0$ and ${}_{N\tilde{N}}\tilde{m}_0$ for the first few values of N, $\tilde{N}$ [16]. Graphs of the corresponding ${}_N\phi$, ${}_{N\tilde{N}}\tilde{\phi}$, ${}_N\psi$, and ${}_{N\tilde{N}}\tilde{\psi}$ for $N = 2$, $\tilde{N} = 2, 8$, and for $N = 3$, $\tilde{N} = 7, 11$ are given in Figures 21 to 25. From some point on, increasing $\tilde{N}$ (for fixed N) does not alter the shape of ${}_{N\tilde{N}}\tilde{\psi}$ very much. The ${}_{N\tilde{N}}\phi$ is symmetric around 0 for N even, and around $1/2$ for N odd. The symmetry axis for both ${}_{N\tilde{N}}\psi$ and ${}_{N\tilde{N}}\tilde{\psi}$ lies at $x = 1/2$ in all cases; for N even, they are symmetric, for N odd, antisymmetric. The different regularity properties of ${}_{N\tilde{N}}\psi$ and ${}_{N\tilde{N}}\tilde{\psi}$ may be useful in some application. A better compression rate may be achieved by using the pair of biorthogonal wavelet filters in decomposition and reconstruction than using the same filters in orthogonal wavelet

3.8 Computation Cell and Algorithm

Wavelet decomposition and reconstruction operations are similar at all resolutions and are the same sampling length at all levels. To implement the wavelet transform does not require the explicit form of $\phi(x)$ and $\psi(x)$, and only require $h(n)$ and $g(n)$, a lowpass filter and a bandpass filter. The wavelet coefficient on different scales can be computed by a recursive algorithm, so it is possible to develop a fast wavelet algorithm. Mallet's

algorithm is very efficient and economic. It preserves the number of non-zero entries. The decomposition and reconstruction algorithms are both tree structures. The discrete wavelet transform computation cell is given in Figure 25. The DWT algorithm is a fast algorithm because of the decomposition of the computation into elementary cells and the subsampling operations which occur at each stage. The operations required by one elementary cell at the jth octave can be counted as follows: two filters, h and g, with equal length L. Filters h and g are directly applied to the considered octave, and then the convolution results for the next cell are decimated. It requires $2L$ multiplications and $2(L$ -1) additions for every set of two inputs. The complexity per input point for each elementary cell is L multiplications and (L - 1) additions. If the cell at the jth octave has input subsampled by 2^{j-1}, the total cell elementary in a filter bank implementation of the DWT on J octaves is $(1 + 1/2 + 1/4 + \cdots + 1/2^{J-1}) = 2(1 - 2^{-J})$. The total complexity is $2L(1 - 2^{-J})$ multiplications and $2(L-1)(1-2^{-J})$ additions [25].

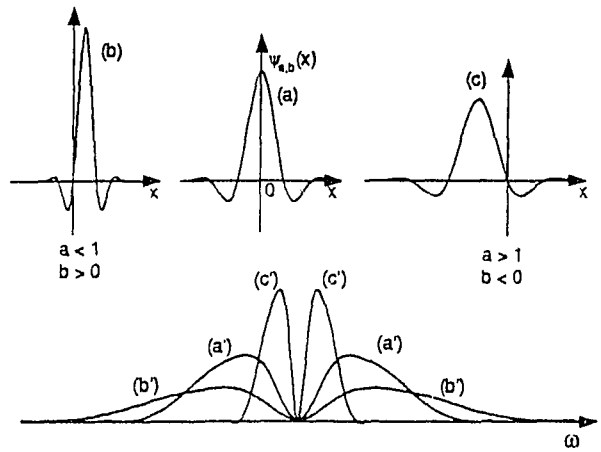


Figure 4. The Wavelet and Fourier Transform of Wavelets

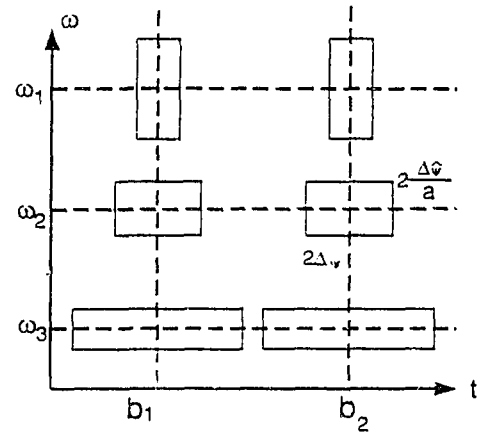


Figure 5. The Wavelet Transform Time-Frequency window

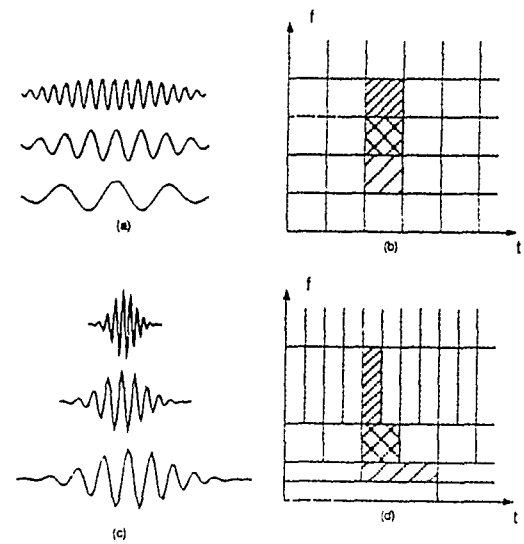


Figure 6. The short time Fourier Transform and Wavelet Transform.
(a).Shifts of a Gaussian window.
(b) Tiling of the time-frequency plane.
(c).Shifts and scales of the prototype bandpass wavelet.
(d) Tiling of the time-frequency plane.

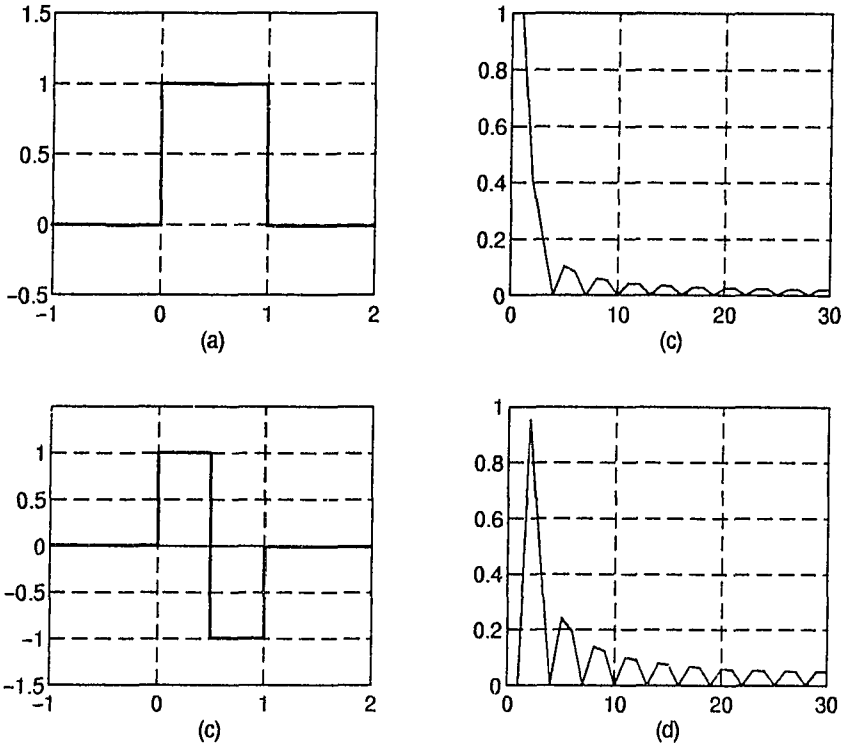


Figure 7. the Haar scaling function $\phi(x)$, (a) in time domain, (b) in frequency domain, the Haar wavelet $\psi(x)$, (c) in time domain, (d) in frequency domain

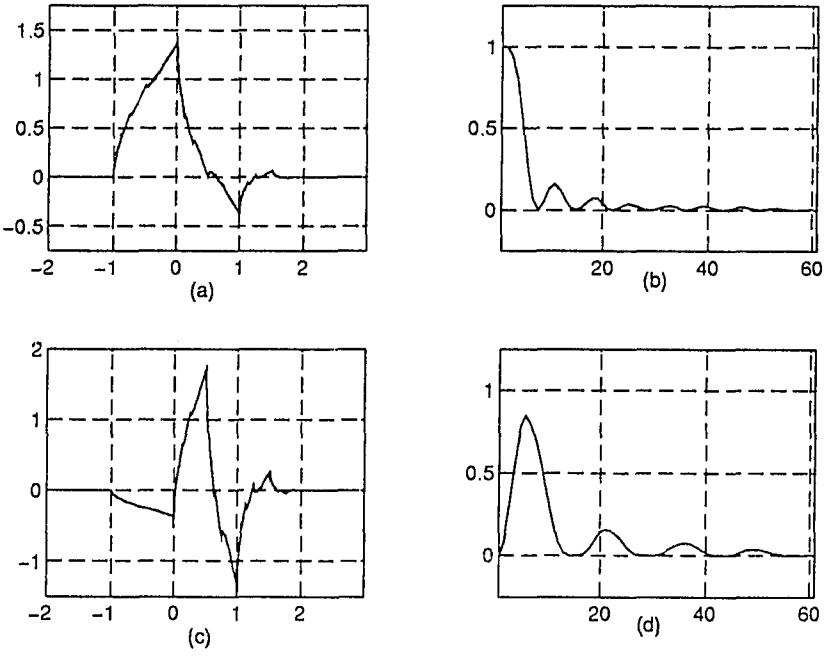


Figure 8. Daubechies compact supported wavelet, N=4.
(a) Scaling function ϕ, (b) The modulus of Fourier transform, $|\hat{\phi}|$.
(c) Wavelet function ψ. (d) The modulus of Fourier transform, $|\hat{\psi}|$.

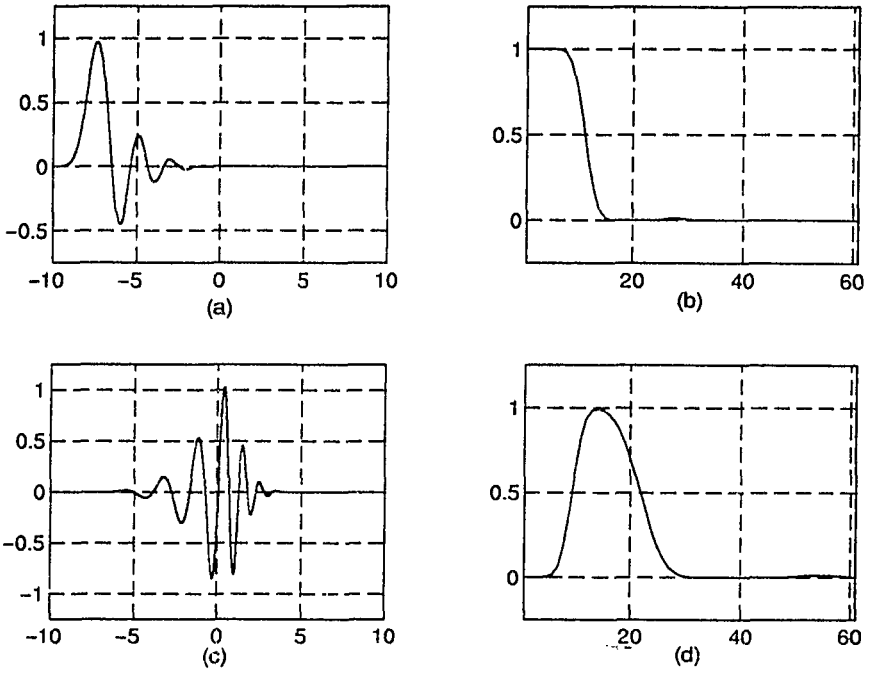


Figure 9. Daubechies compact supported wavelet, N=20.
(a) Scaling function ϕ, (b) The modulus of Fourier transform, $|\hat{\phi}|$.
(c) Wavelet function ψ. (d) The modulus of Fourier transform, $|\hat{\psi}|$.

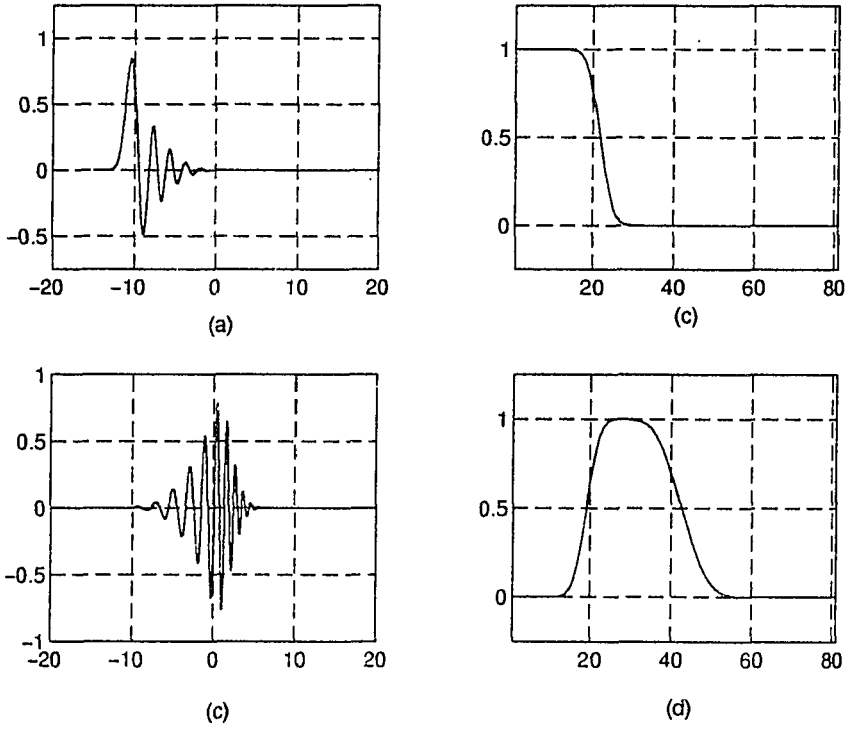


Figure 10. Daubechies compact supported wavelet, N=40.
(a) Scaling function ϕ, (b) The modulus of Fourier transform, $|\hat{\phi}|$.
(c) Wavelet function ψ. (d) The modulus of Fourier transform, $|\hat{\psi}|$.

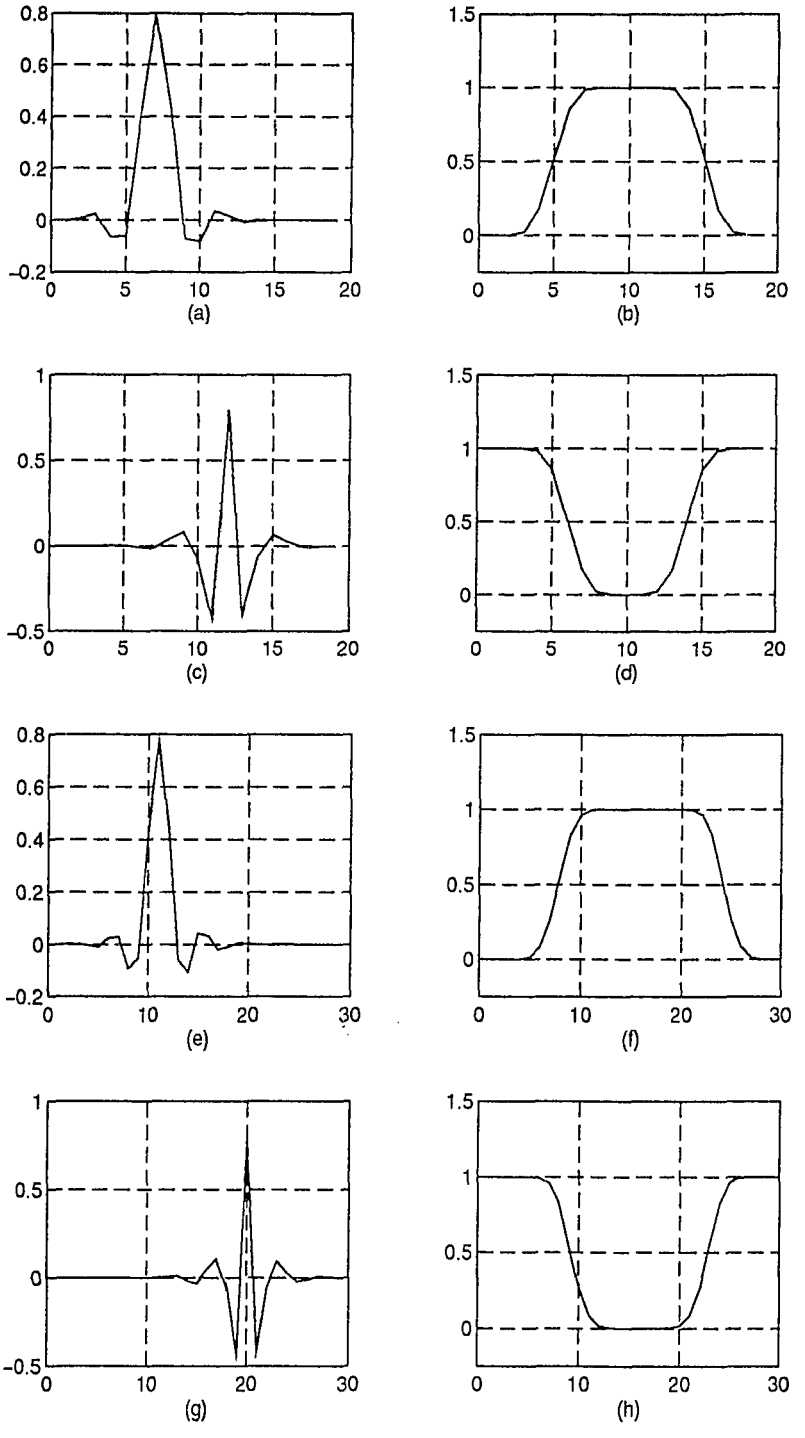


Figure 11. Daubechies orthogonal vanishing moments (a),(b),(c), and (d):vanishing moment N=3, h and g in the time and the frequency domains. (e), (f), (g), and (h): vanishing moment N=5, h and g in the time and frequency domains.

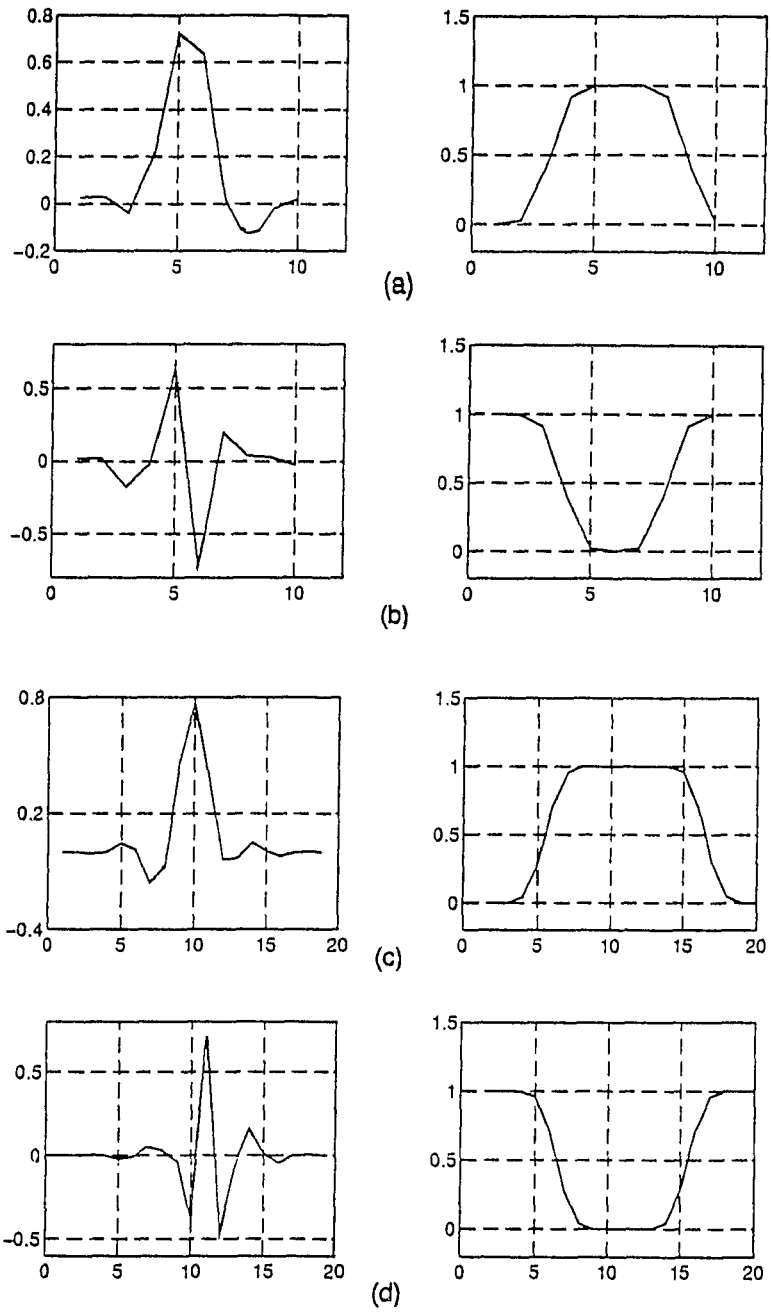


Figure 12. Orthogonal Basis, nearly symmetric,(a) N = 5, Lowpass filter in Time Domain, and in Frequency Domain, (b). N=5, Bandpass filter in Time Domain and in Frequency Domain. (c) N=10, Lowpass filter in Time Domain and in Frequency Domain, (d) N=10 Bandpass filter in Time Domain and Frequency Domain.

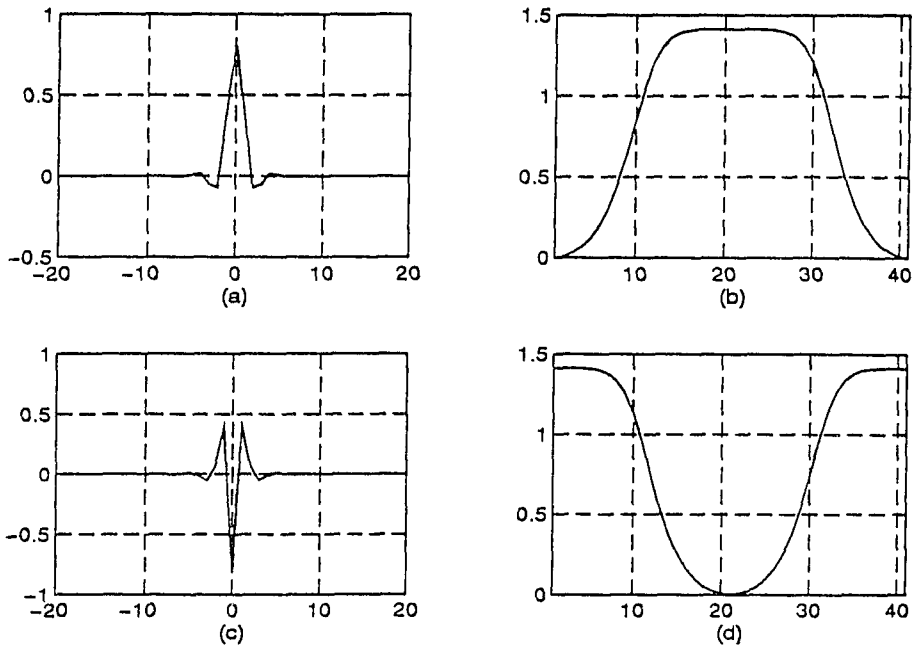


Figure 13. B-L wavelet in time domain and in frequency doman

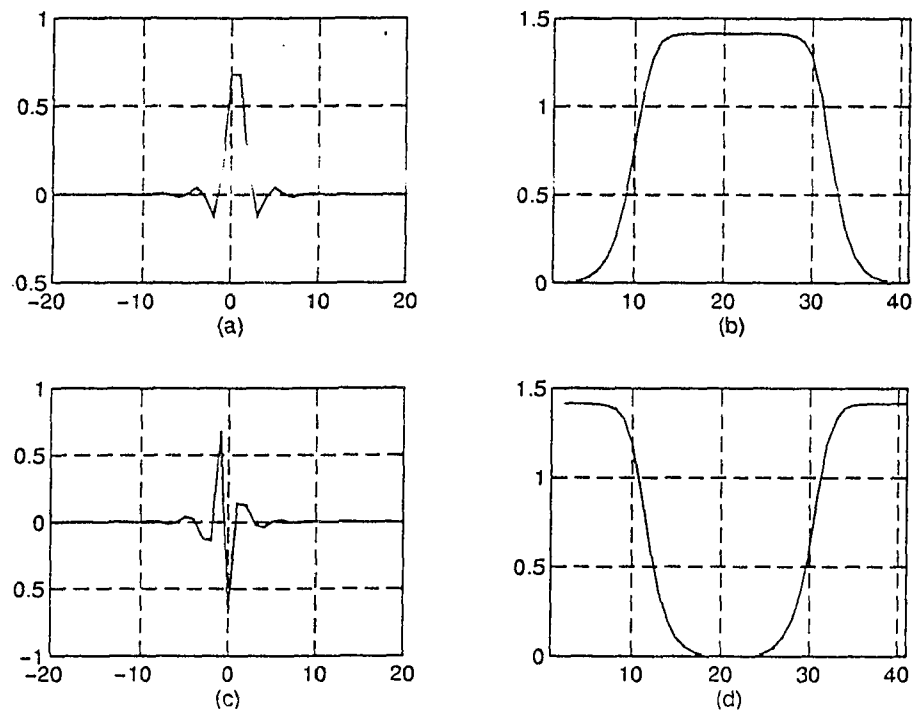


Figure 14 B-L wavelet in time domain and in frequency doman

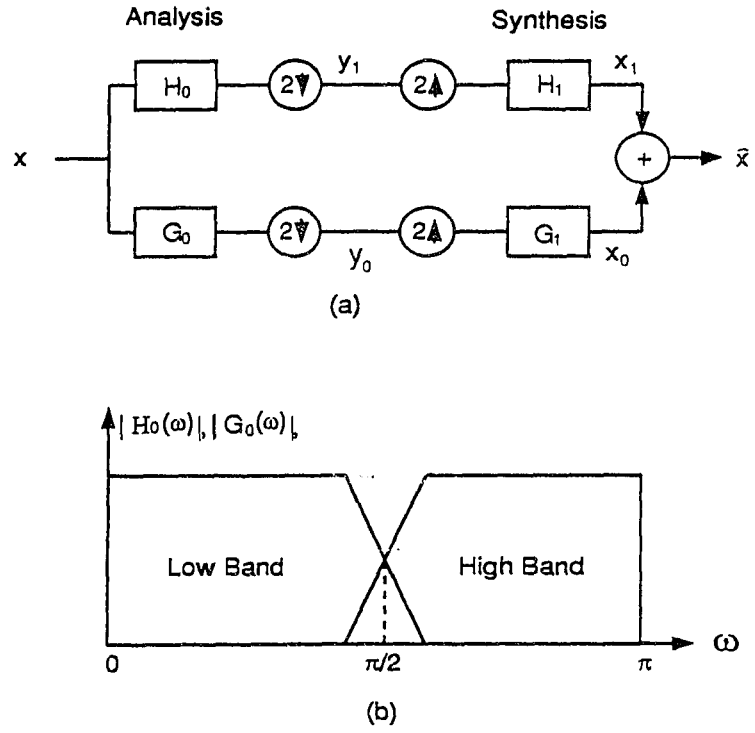


Figure 15. (a) Block diagram for two channel filter bank with analysis filters and synthesis filters.
(b) Spectrum splitting performed by the filter bank.

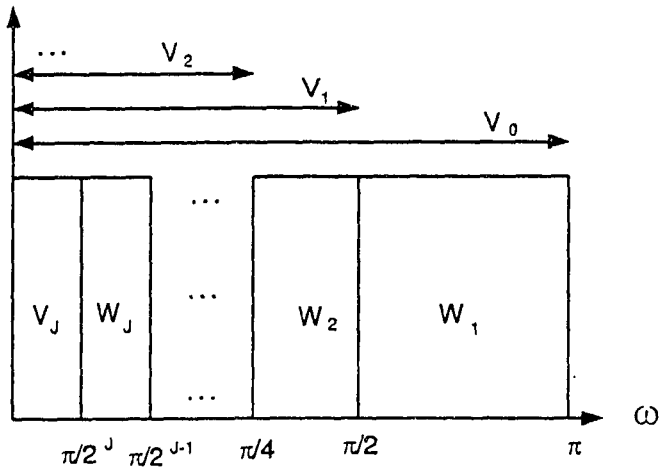
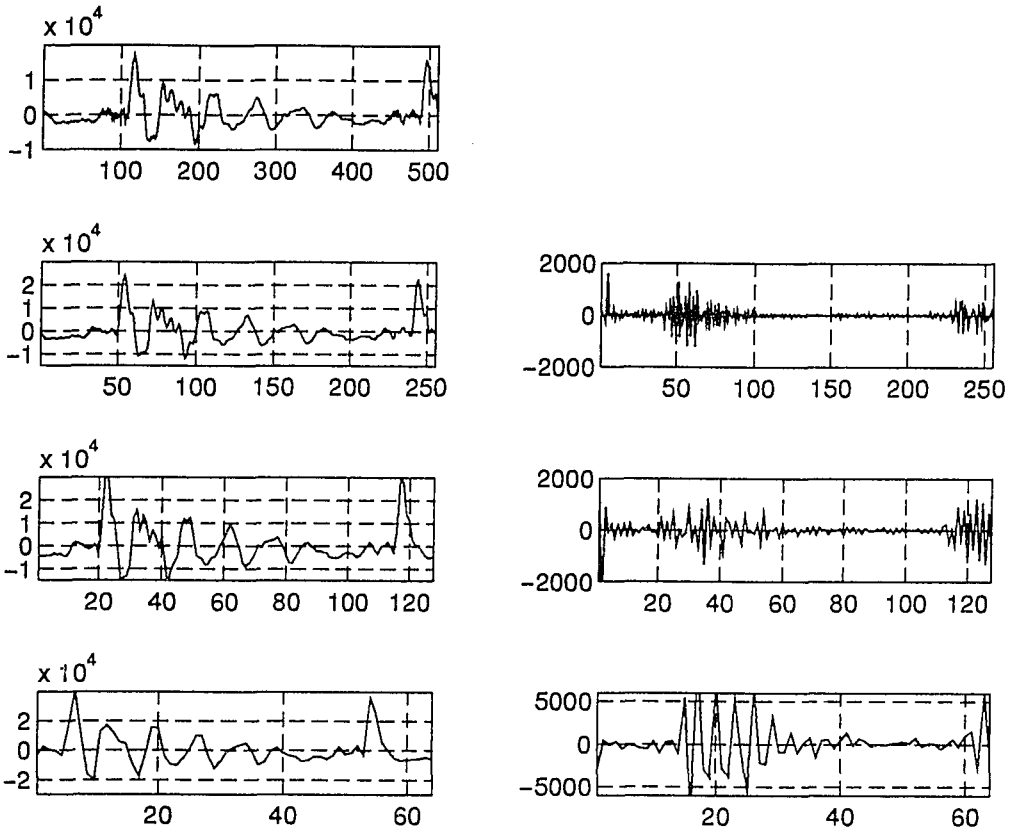


Figure 16. Half-band ideal filters. The spectrum space is divided into Vi spaces and Wi spaces. Only (0 - π) is shown. .



A segment of signal with 512 samples is decomposed using Daubechies Orthogonal Wavelet.

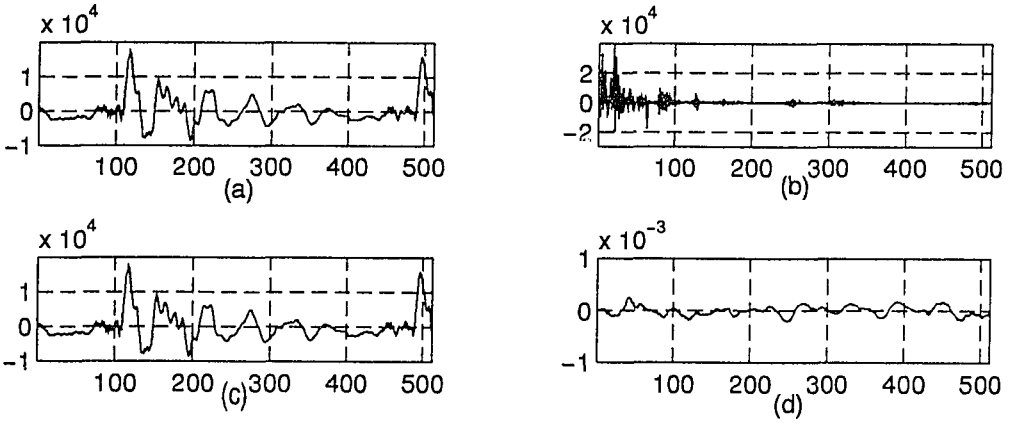


Figure 17. (a) A segment of signal with 512 samples. (b) Wavelet Transform (c) Reconstructed signal from Wavelet Transform, (d) Errors compare the original signals with reconstructed signals.

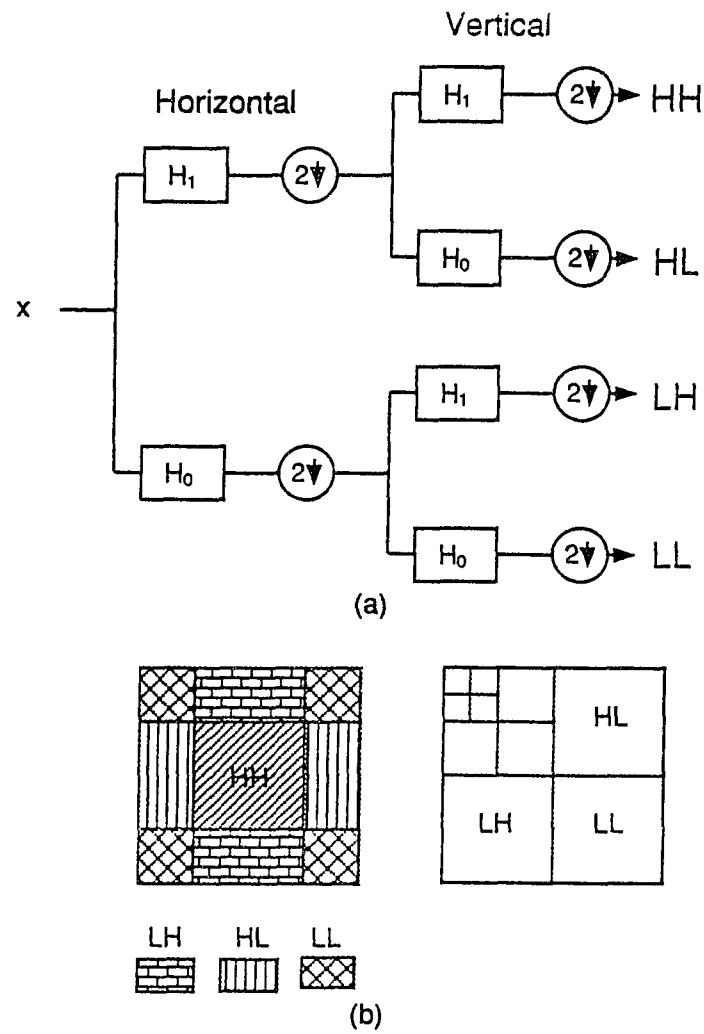
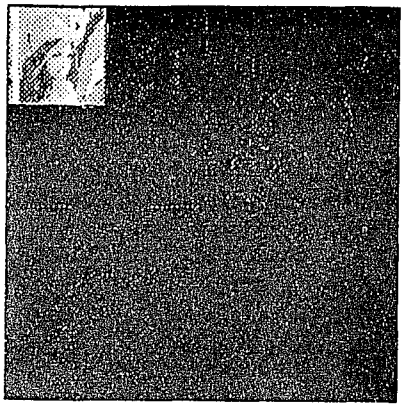


Figure 18. (a) Block diagram for two channel filter bank in two dimensions. Cascade of horizontal and vertical direction.
(b) Division of the frequency spectrum

(a) Lena, 512x512



(b) two stage decomposition

(c) Reconstructed

Figure 19: Image decomposition and reconstruction

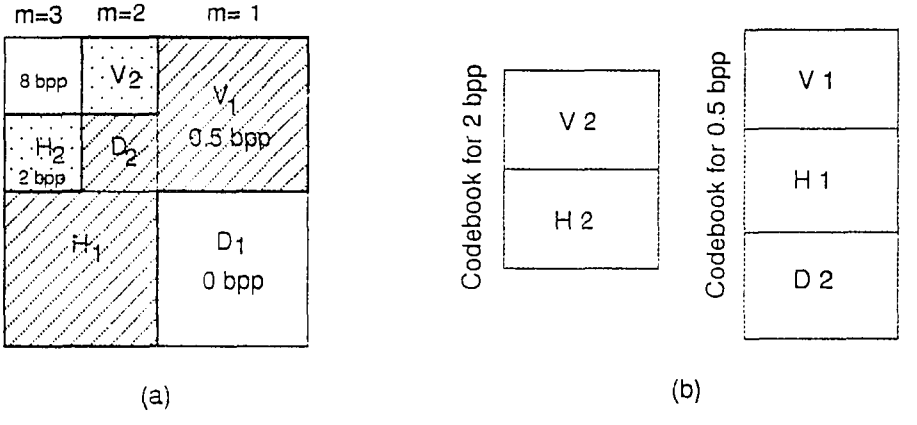


Figure 20. (a) Subimages bit rate allocation
(b) Multiresolution codebook.

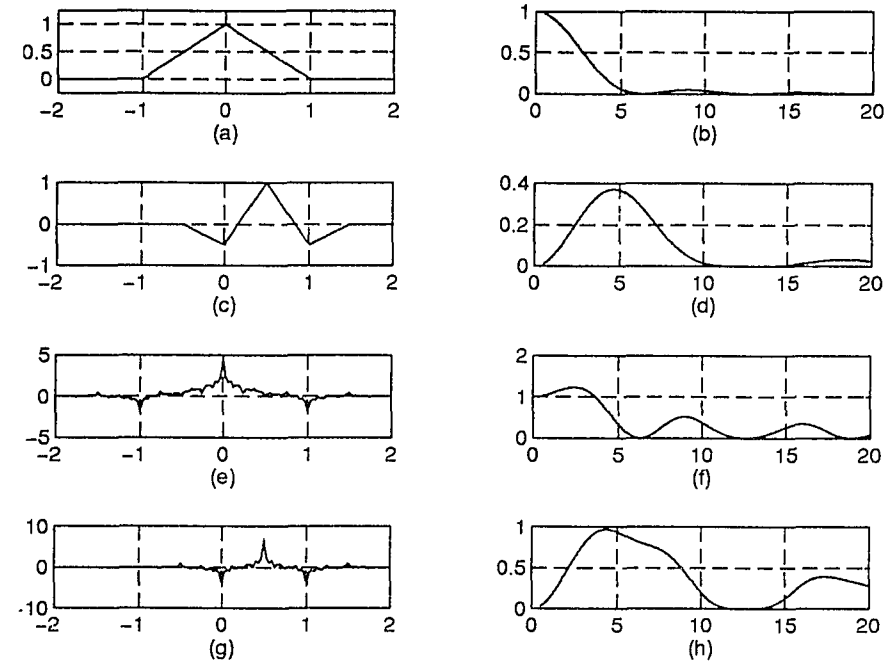


Figure 21. Biorthogonal basis for N=2, $\tilde{N}$=2.
(a),(e) two scaling functions, (b),(f) Fourier transform of (a),(e).
(c),(g) two wavelet W0 and W1. (d),(h) Fourier transform of (c),(g).

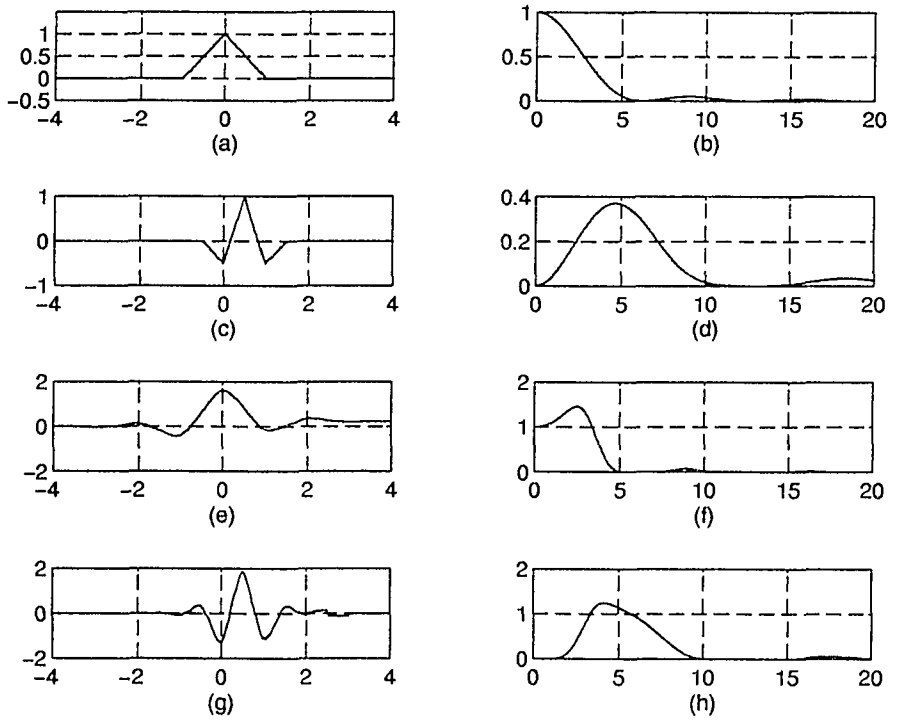


Figure 22. Biorthogonal basis for N=2, $\tilde{N}$=8.
(a),(e) two scaling functions, (b),(f) Fourier transform of (a),(e).
(c),(g) two wavelet W0 and W1. (d),(h) Fourier transform of (c),(g).

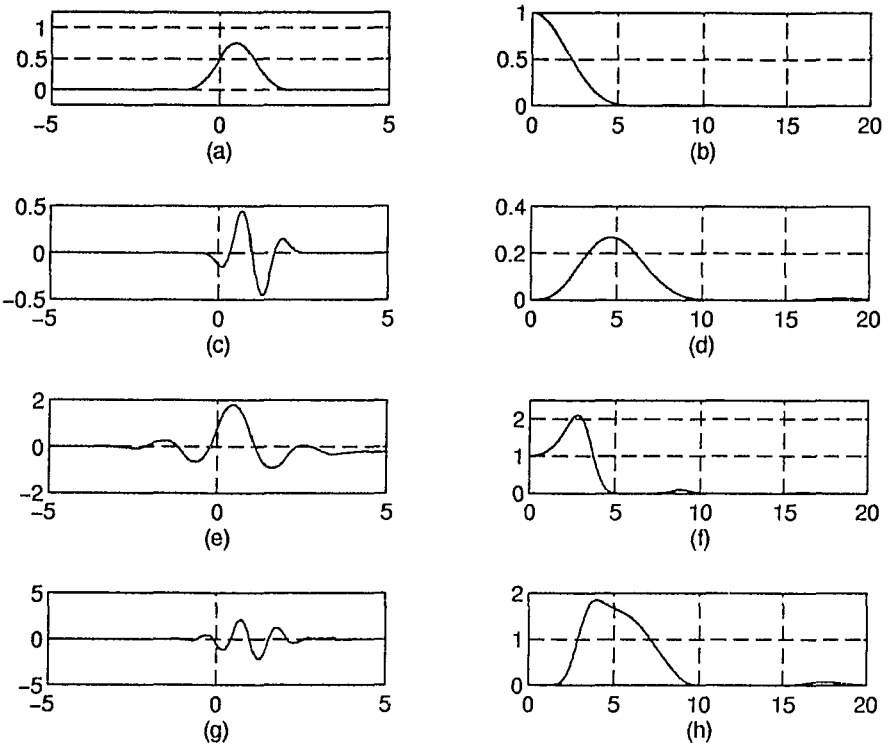


Figure 23. Biorthogonal basis for N=3, $\tilde{N}$=11.
(a),(e) two scaling functions, (b),(f) Fourier transform of (a),(e).
(c),(g) two wavelet W0 and W1. (d),(h) Fourier transform of (c),(g).

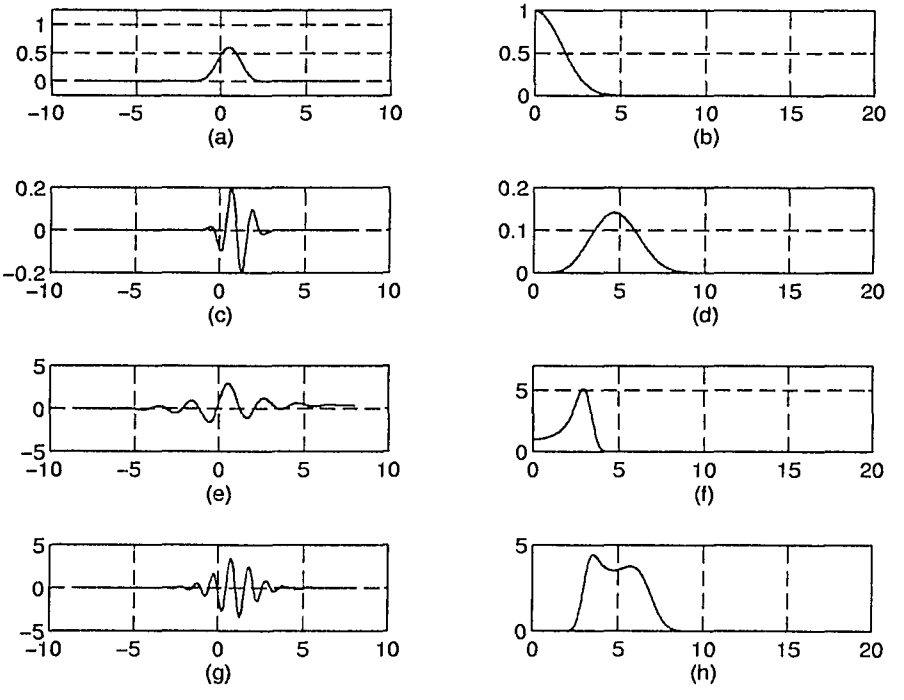


Figure 24. Biorthogonal basis for N=5, $\tilde{N}$=33.
(a),(e) two scaling functions, (b),(f) Fourier transform of (a),(e).
(c),(g) two wavelet W0 and W1. (d),(h) Fourier transform of (c),(g).

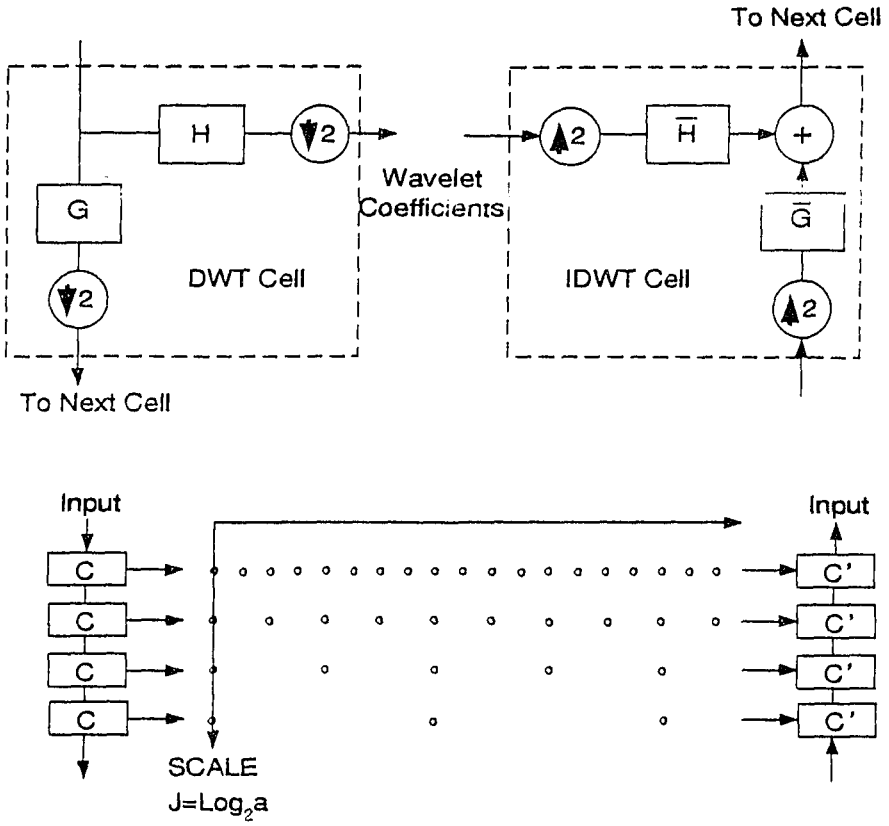


Figure 25. Basic computation cell of DWT and IDWT. Overall organization gives Wavelet Coefficient that correspond to a dyadic grid in the time-scale plane.

Table 1. Daubechies Orthogonal wavelet bases(N=4 and N=20)

Daubechies N=4 wavelet		Daubechies N=20 wavelet							
n	h(n)	n	h(n)	n	h(n)	n	h(n)	n	h(n)
1	0.230377	1	0.000779	11	0.228291	21	-0.013810	31	-0.000037
2	0.714846	2	0.010549	12	0.039850	22	0.006722	32	-0.000004
3	0.630881	3	0.063423	13	-0.155459	23	0.004420	33	0.000007
4	-0.027984	4	0.219942	14	-0.024717	24	-0.003581	34	-0.000001
5	-0.187034	5	0.472697	15	0.102292	25	-0.000832	35	-0.000001
6	0.030841	6	0.610494	16	0.005632	26	0.001393	36	0.000000
7	0.032883	7	0.361503	17	-0.061723	27	-0.000053	37	0.000000
8	-0.010597	8	-0.139211	18	0.005875	28	-0.000385	38	0.000000
		9	-0.326787	19	0.032294	29	-0.000101	39	0.000000
		10	-0.016727	20	-0.008789	30	-0.000068	40	0.000000

Table 2. B-L Orthogonal wavelet basis

B-L orthogonal wavelet basis							
n	h(n)	n	h(n)	n	h(n)	n	h(n)
-20	0.000023	-10	-0.002211	1	0.680369	12	-0.000107
-19	-0.000001	-9	0.000335	2	0.137965	13	0.000873
-18	-0.000057	-8	0.005705	3	-0.124835	14	0.000036
-17	0.000004	-7	-0.001123	4	-0.020747	15	-0.000349
-16	0.000140	-6	-0.015113	5	0.041982	16	-0.000013
-15	-0.000013	-5	0.004238	6	0.004238	17	0.000141
-14	-0.000349	-4	0.041980	7	-0.015113	18	0.000004
-13	0.000036	-3	-0.020740	8	-0.001123	19	-0.000057
-12	0.000873	-2	-0.124835	9	0.005705	20	-0.000002
-12	-0.000107	-1	0.137965	10	0.000334		
-11	-0.002221	0	0.680369	11	-0.002221		

Table 3. Daubiches Biorthogonal wavelet coefficients ($z = e^{-iw}$)

N	${}_N m_0$	$\tilde{N}$	${}_{N,\tilde{N}}\tilde{m}_0$
2	$1/4(z^{-1} + 2 + z)$	2	$2^{-3}(-8z^{-2} + 2z^{-1} + 6 + 2z - 1z^2)$
		4	$2^{-6}(-3z^{-4} - 6z^{-3} - 48z^{-2} + 38z^{-1} + 90 + \ldots)$
		6	$2^{-10}(-5z^{-6} + 10z^{-5} + 34z^{-4} - 78z^{-3} - 123z^{-2} + 324z^{-1} + 700 + \ldots)$
3	$1/8(z^{-1} + 3 + 3z + 2^2)$	1	$2^{-2}(-1z^{-1} + 3 + 3z - z^2)$
		3	$2^{-6}(-3z^{-3} - 9z^{-2} - 7z^{-1} + 45 + 45z - 7z^2 - 9z^3 + 3z^4)$
		5	$2^{-9}(-5z^{-5} + 15z^{-4} + 19z^{-3} - 97z^{-2} - 13z^{-1} + 175 + \ldots)$

(* The coefficents of ${}_{N,\tilde{N}}\tilde{m}_0$ are always symmetric; for very long ${}_{N,\tilde{N}}\tilde{m}_0$ only half the coefficients are listed, the others can be deduced by symmetry.)

4 Voice Pitch Detection using Multiresolution Analysis

4.1 Introduction

The wavelet transform is a useful tool for multiresolution analysis in signal processing. Wavelet transform decomposes a signal into a multiscale representation, and the details at different scales can be used to describe the variation of the signal. In this chapter, voice pitch extraction using a wavelet transform on compact support bases is described. Wavelet transforms of two consecutive scales are used to detect a sharp change in voice signals. Local maximum is found based on the modulus of wavelet transform. Examples are provided that show the success of wavelet application in voice pitch detection.

4.2 Multiscale Edge Representation

Many studies of image processing and interpretation use the important concept of edge features, which are considered to occur at those positions where the image intensity changes rapidly. In a 1-D signal, such as a voice signal, these points represent the glottal opening and closing; in a 2-D image, they provide the location of the structure edge, which is generally used to identify the objects in the image. Many edge detection methods have been studied and developed; the multiscale edge detection concept can be found in [26], [27], [28] and [29]. Several important edge detection algorithms look for local maxima of the gradient of various smoothed versions of the image. Edge features can be classified and indexed by the various resolutions or scales. Mallat and Zhong studied the properties of multiscale edges through the wavelet theory and proposed a new wavelet edge representation which uses wavelet transform modulus and gradient on different scales to represent the image, and which can reconstruct a close approximation of original signals from the multiscale edges [20]. In their approach, a particular class of wavelet – a non-orthogonal wavelet has been

used to realize the multiscale edge representation. Basically, there are two approaches used to detect the edge point: zero-crossings and the local maxima of the wavelet transform. The evolution of the wavelet transform local maxima across scales characterizes location of the signal sharp variation. The Canny edge detector is the equivalent of finding the local maxima of a wavelet transform modulus.

4.3 Wavelet Transform local maxima and Edge Detection

Instead of having constant time (space) and frequency domain windows, the wavelet transform uses short windows at high frequencies and long windows at low frequencies. The resolution of the wavelet transform varies with the scale parameter, making it a desirable tool for processing signals with features at different scales. The wavelet scheme decomposes the original signal into a coarse signal and a set of detail signals at the different scales. The set of detail signals can be used to describe the variation of the original signal in the time domain. Different wavelets can be selected depending on the application.

In the Dyadic Wavelet Transform(DWT), the transform coefficients of the signal $f(x)$ can be computed by:

$$C_{m,n} =< \psi_{m,n}, f >= \int_{-\infty}^{+\infty} f(x)\psi_{m,n}(x)dt \tag{127}$$

and the synthesis formula is:

$$f(x) = \sum_{m,n} C_{m,n}\psi_{m,n}(x) \tag{128}$$

The wavelet coefficients for different scales can be computed by the recursive algorithm:

$$A_{m,n}f(x) = \sum_k h_{2n-k} A_{m-1,k} f(x) \tag{129}$$

$$C_{m,n}f(x) = \sum_k g_{2n-k} A_{m-1,k} f(x) \tag{130}$$

where $A_{m,n}f(x)$ are approximation coefficients characterizing the projection of $f(x)$ onto V_m, the space spanned by the orthonormal functions $\phi_{m,n}$. The reconstruction is given by:

$$A_{m-1,l}f(x) = \sum_n h_{2n-l} A_{m,n} f(x) + g_{2n-l} C_{m,n} f(x) \tag{131}$$

In multiscale edge representation, a separable spline scaling function $\theta(x, y)$ is used as the smoothing filter, and its partial derivatives are used to construct the wavelets [20]. A smoothing function $\theta(x, y)$ is defined as:

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \theta(x) dx = 1 \tag{132}$$

$$\lim_{x,y \to \pm\infty} \theta(x) = 0 \tag{133}$$

Suppose that $\theta(x, y)$ is twice differential. Let ψ^1 and ψ^2 represent the first and the second order derivative of $\theta(x)$:

$$\psi^1(x) = \frac{d\theta(x)}{dx} \tag{134}$$

$$\psi^2(x) = \frac{d^2\theta(x)}{dx^2} \tag{135}$$

Because the integral of ψ^1 and ψ^2 equals zero,

$$\int_{-\infty}^{+\infty} \psi^1(x)dx = 0 \tag{136}$$

$$\int_{-\infty}^{+\infty} \psi^2(x)dx = 0 \tag{137}$$

these two functions can be considered to be wavelet functions. The wavelet transform of signal $f(x)$ using wavelet function ψ^1 and ψ^2 in discrete scale $[2^j]_{j\in Z}$ can be written as:

$$W_{2^j}^1 f(x) = f * \psi_{2^j}^1(x) = 2^j \frac{d}{dx}(f * \theta_{2^j})(x) \tag{138}$$

$$W_{2^j}^2 f(x) = f * \psi_{2^j}^2(x) = (2^j)^2 \frac{d^2}{dx^2}(f * \theta_{2^j})(x) \tag{139}$$

which are the first and the second derivatives of $f(x)$ smoothed at scale $[2^j]_{j\in Z}$. The local extrema of $W_{2^j}^1 f(x)$ correspond to the zero-crossings of $W_{2^j}^2 f(x)$. The extrema detection of the wavelet transform built from ψ^1 is equivalent to Canny edge detection [30]. The wavelet transform is related to multiscale edge detection and is equivalent to finding the local maxima of the wavelet transform modules. Mallet [20] has showen that the local maxima of $|W_{2^j}^1 f(x)|$ correspond to a sharp variation point of $f(x) * \theta_{2^j}(x)$, whereas the minima correspond to slow variations, and the zero crossing of the second derivative gives the position information. When detecting a local maximum of $|W_{2^j}^1 f(x)|$, both of the position x_0 and the amplitude of $W_{2^j}^1 f(x_0)$ have been recorded. This amplitude stabilizes the representation and is particularly important to describe the signal singularities. The wavelet transform of $f(x)$ decomposes $f(x)$ into $\{S_J^d f(x), W_{2^0}^d f(x), ...W_{2^J}^d f(x)\}, J\in Z$. The very sharp change in voice signal at time $t = t_0$ will be exhibited as a local maxima in the dyadic wavelet transform across several consecutive scales.

For most image and voice applications, selecting an antisymmetric wavelet is desirable because similarly it can operate in the right and the left neighborhood of a point. To build a wavelet $\psi(x)$ equal to a first order derivative of a smoothing function $\theta(x)$ implies that $H(\omega)$ is real and symmetrical around zero, and $G(\omega)$ must be antisymmetric around zero and imaginary. In our edge detection experiment we use the non-orthogonal wavelet introduced by Mallat and Zhong in [20]:

$$H(\omega) = e^{i\omega/2}(\cos(\omega/2))^3 \tag{140}$$

$$G(\omega) = 4ie^{i\omega/2}\sin(\omega/2) \tag{141}$$

Figure 26 is a quadratic spline wavelet of compact support, which we use in our experiment; the derivative of the cubic spline function $\theta(x)$ is shown in (b).

4.4 Voice Pitch Detection

Voice fundamental frequency f_0 is a particularly powerful parameter used in assessing the functional and anatomical status of the larynx. For instance, the prosodic information of an utterance is predominantly determined by this parameter. In voice pathology, local pitch-by-pitch analysis would help establish the importance of "jitters", defined as variation of pitch change and "shimmer", defined as variation of amplitude change within pitches. Pitch period information can also be used in voice disease diagnosis to measure improvement after a treatment or operation. It can also be used in speaker identification and verification, and in speech analysis and synthesis [31], [32]. The quality of vocoded speech is essentially influenced by the quality and faultlessness of the pitch measurement. The ear is more sensitive to variation in the fundamental frequency than to changes in other speech signal

parameters. The pitch period varies for different speakers based on factors such as age, reqional or accent, and gender. The emotional state of the speaker also strongly inflects the speech signal. The fundamental frequency of an arbitrary speaker can vary from 50 to 800 Hz, or four octaves. The momentary portion of the vocal test may change abruptly at any time. Additionally, the glottal excitation signal itself is not always regular. The voice may temporarily fall into vocal dry and the glottal waveform exhibits occasional irregularities. These factors produce a non-stationary speech process and make the task of pitch determination one of the most difficult problems in speech analysis.

The problem, in signal processing terms, is to extract relevant information from the possibly noise-corrupted measurement of a signal (e.g., extracting the pitch period of speech from a noisy data record). Pitch extractors are of three basic types [32]: event detectors, peak detectors, and time counters.

The event detector sets a marker whenever it detects that a significant event such as the instant of glottal closing or opening has occurred. When the glottal closes, a discontinuity in the derivatives of glottal airflow occurs. The pitch period is estimated by locating the instant of opening or closing either at the peak point or at the zero crossing point. One event detection method is to locate the maximum of the autocovariance matrix of a given signal to determine the instant at which the glottal is closed. Another event detector uses the occurrence of discontinuities in the derivatives of glottal airflow to detect the Glottal Closure Instant(GCI). The advantage of these methods is that the pitch period estimate is very accurate in the case of certain vowels. The disadvantage is this method can only be used for certain vowels, not suitable for all vowels; it is not suitable for nonstationary pitch periods.

The peak detector is a ordinary short-term pitch detector. The actual signal frame where all the periods are depicted is transformed into the interval $(0, t_0)$ in the spectral domain. The basic extractor picks up the peak at $T0$ (maximum or minimum according to the respective pitch detector) and locates its position with respect to the origin of the spectral function, then measures the distance from the starting point of that signal.

Time counters operate without a specified reference point. When the speech signal is in a particular state, the counter counts the elapsed time until the signal again reaches the same or very approximately the same state. When the measurement is activated, the extractor computes for each instant the Euclidean distance between the momentary state and the reference state. When the momentary state becomes sufficiently similar to the reference state, the Euclidean distance is at a minimum. It assumes that one period has elapsed. This method works in an incremental way. After it is positioned, it counts periods from that point on. If a setup measurement fails, it must be restarted. Usually having a fixed length window for calculating the average pitch period, and using the average pitch period to represent the given segment of voice signal, these pitch detectors assume that the pitch period is stationary within each segment. Hence, they are insensitive to nonstationary variations in the pitch period over the segment signal, and are unsuitable for speakers with both low and high pitch periods. The advantage is that the computation is simple.

4.4.1 Pitch Detection of Speech Signals Using DWT

The wavelet transform for pitch period detection is an event method. This method is based on the fact that shape changes in the original signal may exist in several different scales. The local maximum of the wavelet transform across scales may be used to indicated Sharpe changes, i.e. glottal opening and closing. The pitch period can be calculated from

one local maxima to the next local maxima. The wavelet transform property of changing window size in different frequency ranges, makes it possible for the DWT pitch detector to handle both stationary and nonstationary signals, and also signal segments with different pitch periods. A pitch detection technique using DWT was introduced by Kadambe and Boudreaus-Bartels in [31]. In their algorithm, they compute wavelet transform up to scale $a = 2^5$. Here we develop a modified DWT pitch detection algorithm using a non-orthogonal, antisymmetrical wavelet. This algorithm requires only the first and the second scale wavelet transform. The wavelet we selected can operate similarly on both sides of a signal, and we see that local maxima of the wavelet transform modulus are located at the sharp variations of the signal. A window is used to locate the local maximum on different scales. This algorithm can be implemented using relatively few computations and can be applied to real time applications. Pitch detection can be performed without prior information regarding the subject, and there is no need to pre-set the cut-off frequency range for evaluation.

The algorithm works as follows:

- Step 1. Voice signal $f(x)$ is decomposed into $\{S_{2^2} f(x), W_{2^1} f(x), W_{2^2} f(x)\}$.
- Step 2. We use $M_{2^J} f(x)$ to represent the modulus of difference signal in scale 2^J, where $J = 1, 2$. A lowpass filter is used to remove the high phenomena from $M_{2^J} f(x)$.
- Step 3. Select d, $d > 0$, for a window with window size $[-d, d]$, and select thresholding values C_{2^1} for scale 2^1, C_{2^2} for scale 2^2, respectively.
- Step 4. Locate the sharp change points which satisfy the following conditions:

  1. $M_{2^1} f(x) \geq C_{2^1}$ and $M_{2^2} f(x) \geq C_{2^2}$
  2. $M_{2^1} f(x_i) > M_{2^1} f(x_{i+1})$ and $M_{2^1} f(x_i) > M_{2^1} f(x_{i-1})$ for $0 \leq i \pm 1 \leq d$

3. $M_{2^2} f(x_i) > M_{2^2} f(x_{i+1})$ and $M_{2^2} f(x_i) > M_{2^2} f(x_{i-1})$ for $0 \leq i \pm 1 \leq \frac{d}{2}$

The flow chart of this algorithm is shown in Figure 27.

If there are $2n$ samples in a given segment voice signal, there will be n samples in scale 2^1 and $n/2$ samples in scale 2^2 after the wavelet transform with down sampling. The position of sharp variation in scale 2^1 may have shifted a little because the down sampling operation. This algorithm uses a small window in the first scale, locates the position of local maxima, then reduces the window size in half at scale 2^2 to locate the local maxima in the scale 2^2. If there exists a sharp variation in the original signal, the local maxima will be exhibited over the consecutive scales.

4.4.2 Experiment Results and Discussion

The goal of understanding vowel quality is the hope that measures can be obtained which directly reflect glottal function. In our experiment, we select vowel "a", and two sets of voice signal pronounced by patients to test our algorithm. Figure 28(a) shows the original voice signal vowel "a" in 2,048 samples in a time interval of 46.44 ms at sample frequency 44.1kHz. figure 28(b) shows the voice signal after applying the wavelet lowpass filter twice, resuting in a more smoothed version of the original. Figure 28(c) and (d) are the modulus of the DWT at the first and the second scales, respectively. Each peak in Figure 28(c) and (d) represent sharp changes in the original signal. A low-pass filter is used to remove the high frequency components in the DWT modulus and to exhibit the local maxima more clearly. A window of size $d = 14$ is used to locate the local maxima in scale 2^1. The "*" in Figure 28(c) represent the locations where a local maximum in both scales exists. Pitch period information are very useful in voice diseases diagnostic, and also can

be used to indicate the improvement of medical treatment. In our second experiment, we select two sets of vice signal pronounced by patients, before and after medical care. The original voice signals (4096 samples, 92.88 ms), before medical care are shown in Figure 29 and 31 for Patient One and Two, respectively. The results of applying wavelet lowpass filter twice are shown in 29, 31 (b). The wavelet transform modulus on the first and the second scales are shown in (c) and (d), respectively. Before medical treatment, the pitches of Patient One (in Figure 29) are varied and Patient Two's pitch (in Figure 31) is not even detected. The voice signal, wavelet lowpass filtering results, the first and the second scale wavelet transform modulus, after medical treatment, are shown in Figure 30 and 32, for Patient One and Two respectively. Both patients experience radical improvement after medical treatment. The "*" in figures 30(c) and 32(c) demonstrate the sharp changes in the original signals. The petch periods detected for patient one and two are listed in Table 4.

To process long continuous speech signals, the signals can be segregated into segments. Let L be the length of a voice singal segment segment, and S be the total samples of a continuous speech. Then we perform $\lfloor S/L \rfloor$ DWT for each segment. The boundary of each segment does not change the pitch period if we use the wavelet transform block from $L_w/2, L - L_w/2$, where L_w is the wavelet filter length. The local maximum on different scales may be allocated by applying a window with a threshold value. The selection of the threshold values C_1 and C_2 is based on the segment signal.In our experiment, the threshold C_1 is set to 0.75 of the maximaum in this segment, and $C_2 = 0.75C_1$. The sharp change in the original signal will generate a high peak in the difference scale and a low peak corresponding to a lower variation in the original signal.

Fast wavelet transform algorithms provide an efficient tool for performing the computation quickly. If the filter length is L_w, the voice signal segment length is L, and $L >> L_w$, straightforward DWT needs $2L_w(1-2^{-J})$ multiplication/point and $2(L_w-1)(1-2^J)$ additions/point, J is octaves [25]. DWT can be performed quickly, and the algorithm can be applied to real time applications.

The advantages of the DWT method are as follows. (1) The voice signals to be processed do not have to be stationary or quasi-stationary. The wavelet transform property of using narrow window at high frequency and wider window at low frequency allows the DWT pitch detection to handle non-stationary voice signals, and also to handle a wide range of pitch periods when pitch changes occur in the same segment. (2) The pitch periods detected by DWT are accurate. (3) The computation is simple and can be used in real time pitch detection. The wavelet transform for speech period detection is expected to benefit in the diagnosis of voice disease, indicate the change of glottal change, voice identification, pathologic voices such as spastic dysphonia, moderate-to-severe breathy voices associated with recurrent laryngeal nerve paralysis, polyps and nodules; and moderate-to-severe rough voices associated with different types of vocal lesions and with laryngeal cancer, in which current standard analysis techniques fail.

Table 4: The patch period detected for patient one and patient two

1 Patient One, after medical treatment											
n	1	2	3	4	5	6	7	8	9	10	11
Samples	343	337	342	343	341	340	337	338	338	332	336
Time(ms)	7.778	7.642	7.755	7.778	7.732	7.709	7.642	7.664	7.664	7.528	7.619
2 Patient Two, after medical treatment											
n	1	2	3	4	5	6	7	8	9	10	11
Samples	193	193	193	194	191	196	194	194	194	193	195
Time(ms)	4.376	4.376	4.376	4.399	4.331	4.444	4.399	4.399	4.399	4.376	4.421
n	12	13	14	15	16	17	18	19			
Samples	193	194	194	194	193	192	198	193			
Time(ms)	4.376	4.376	4.399	4.399	4.376	4.353	4.489	4.376			

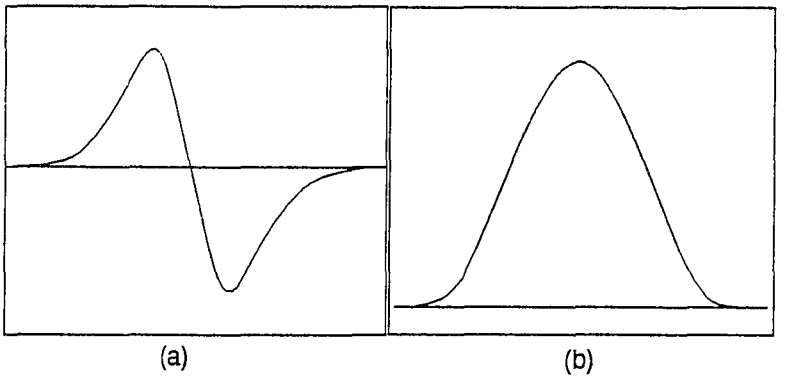


Figure 26. G(ω) and H(ω). Cubic Spline.
G(ω) is antisymmetric around zero,
H(ω) is symmetrical around zero.

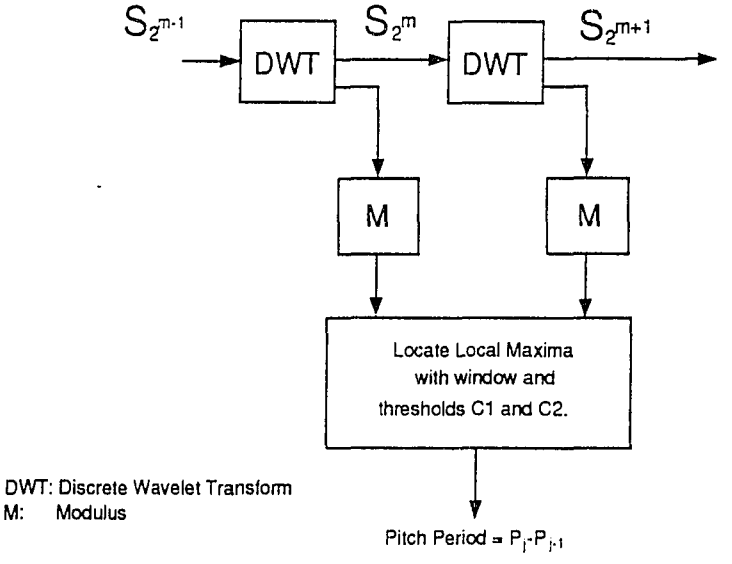


Figure 27. Block diagram of voice pitch period detection algorithm

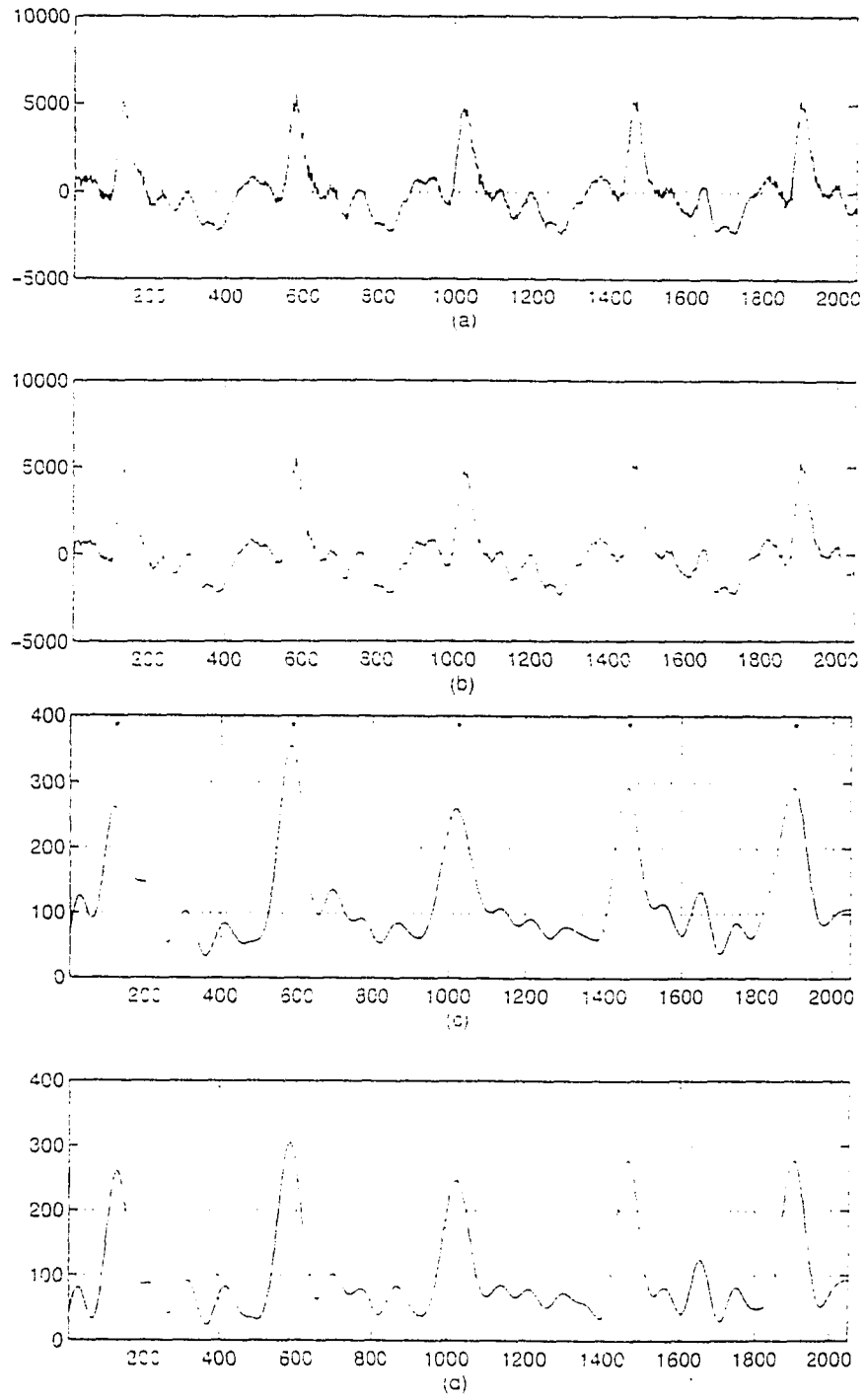


Figure 28. Voice vowel "a" and detected pitches. (a) "a" in 2048 samples (46.44 ms) (b) smoothed "a" by applying twice wavelet transform, (c) the first scale of wavelet transform, (d) the second scale of wavelet transform. "*" in (c) represent the local maximum found in (c) and (d).

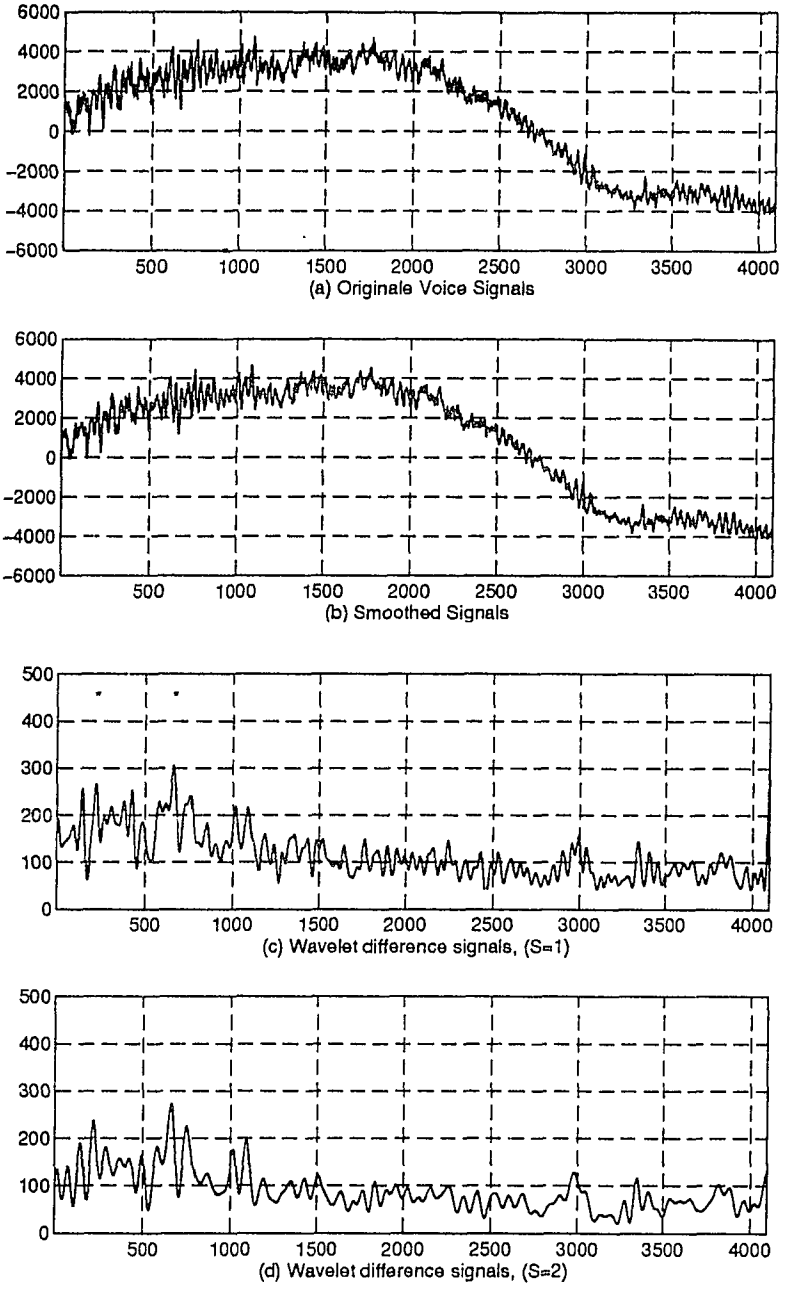


Figure 29. Voice pitch detected for patient one before medical treatment. (a) Original voice signals, pronounced by patient one, 4096 samples(92.88ms). (b) smoothed voice signals by applying twice wavelet transform. (c) wavelet difference signals on scale 1. (d) wavelet difference signals on scale 2. "*" in (c) represent the local maximum detected on both scales s=1 and s=2.

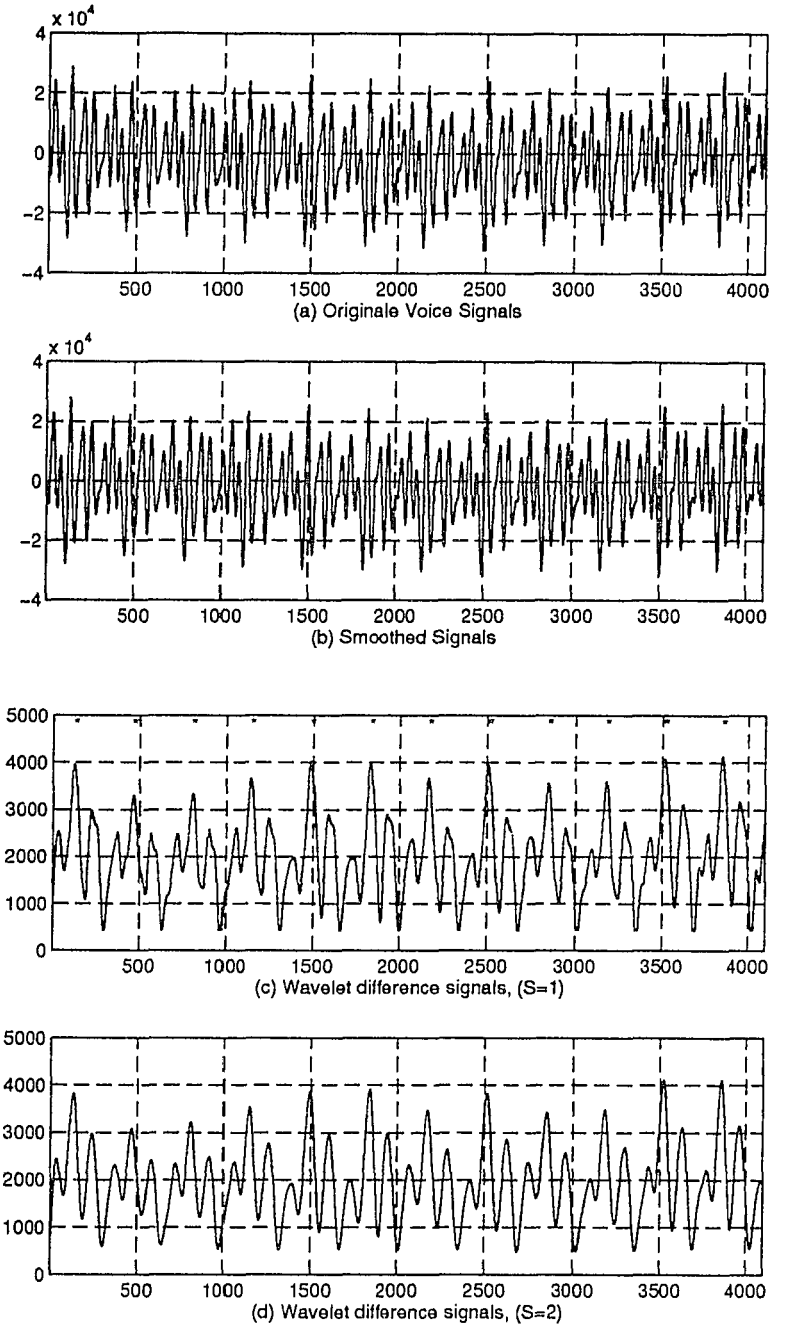


Figure 30. Voice pitch detected for patient one after medical treatment. (a) Original voice signals, pronounced by patient one, 4096 samples(92.88ms). (b) smoothed voice signals by applying twice wavelet transform. (c) wavelet difference signals on scale 1. (d) wavelet difference signals on scale 2. "*" in (c) represent the local maximum detected on both scales s=1 and s=2.

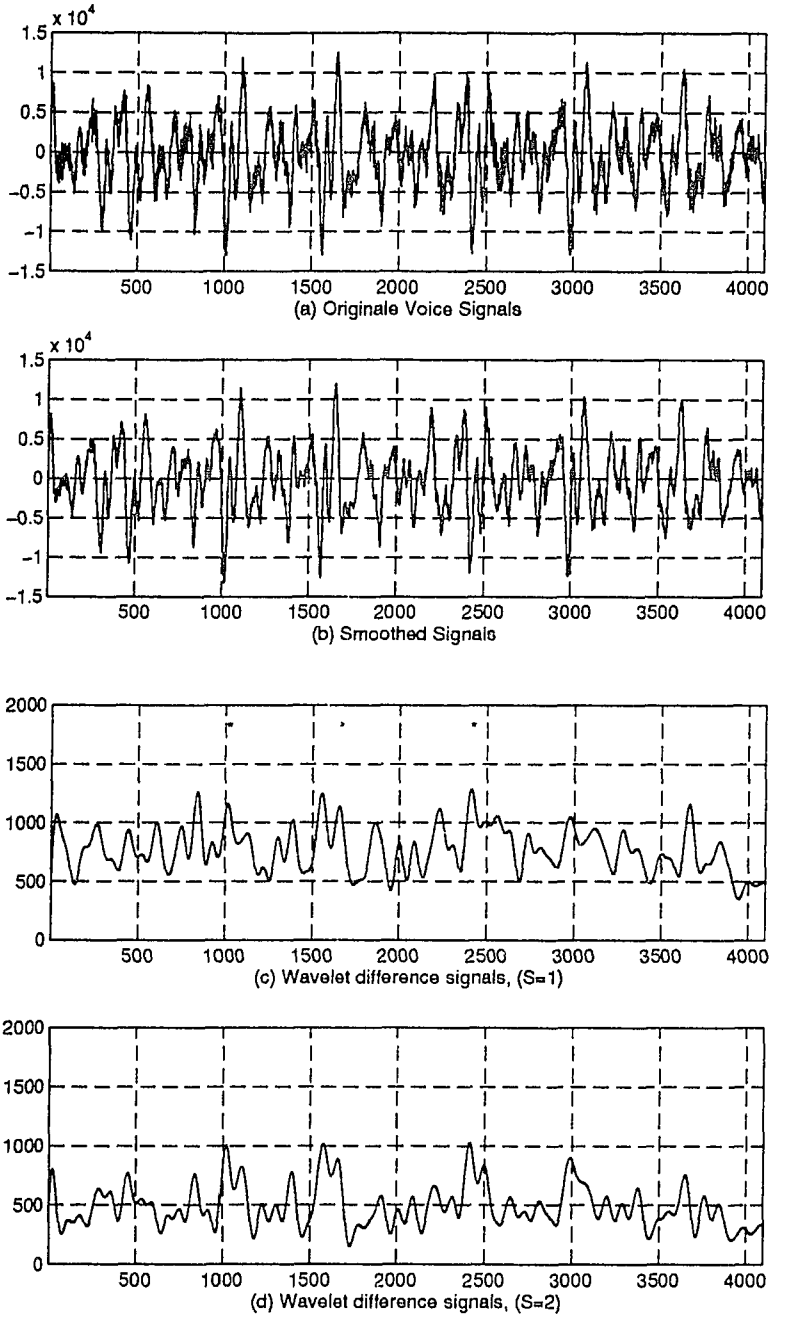


Figure 31. Voice pitch detected for patient two before medical treatment. (a) Original voice signals, pronounced by patient one, 4096 samples(92.88ms). (b) smoothed voice signals by applying twice wavelet transform. (c) wavelet difference signals on scale 1. (d) wavelet difference signals on scale 2. "*" in (c) represent the local maximum detected on both scales s=1 and s=2.

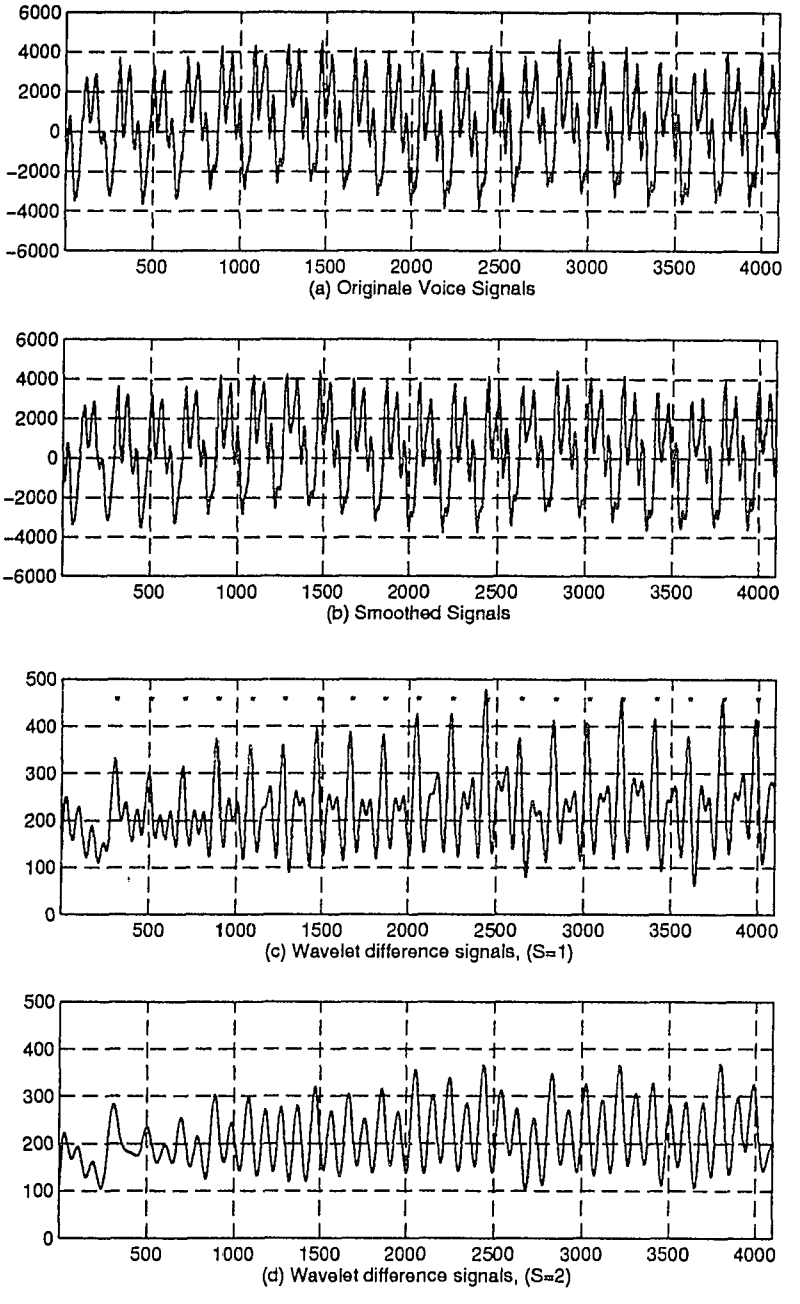


Figure 32. Voice pitch detected for patient two after medical treatment. (a) Original voice signals, pronounced by patient one, 4096 samples(92.88ms). (b) smoothed voice signals by applying twice wavelet transform. (c) wavelet difference signals on scale 1. (d) wavelet difference signals on scale 2. "*" in (c) represent the local maximum detected on both scales s=1 and s=2.

5 Cell Contour Extraction using Compactly Supported Wavelet

5.1 Introduction

An important application of image analysis is as a tool for pathologists to distinguish between normal, reactive and cancerous cells. This is routinely used by the trained pathologist observing cell morphology under a microscope. The pathologist is a specialist capable of integrating picture information and rendering an interpretation which may not be obvious. Implicitly, the pathologist learns to regard certain features as important while disregarding others as mere artifacts. Features which trained pathologists indicate are important in their diagnoses are cell and nuclear size and shape, and cell texture and coloration. The single most important task in the application of computer based image analysis as a diagnostic tool is the accurate segmentation of the relevant image structures. Errors in segmentation may lead to greater variance of subsequent extracted features, which may lead to wrong classifications and produce erroneous diagnoses. Additionally, tools developed for routine applications must have a fast implementation suitable for interactive applications.

Cell images were obtained by specifying a region of interest via a touch screen mounted directly over a video monitor [33]. The region of interest may contain one or multiple cells. Generally, the cell consists of a dark possibly textured nucleus surrounded by a lighter cytoplasm region. The cytoplasm region and the inside nuclear region usually exist as two separate aggregates. The background is generally lighter and may contain various darker artifacts which should not be recognized as meaningful regions due to their much smaller sizes. Factors preventing the straightforward segmentation and contour extraction of the desired structures include: 1) low contrast between background and structures to be seg-

mented, 2) boundaries having varying gray levels which are sometimes poorly defined or fragmented, 3) the presence of artifacts due to the specimen and/or preparation, 4) the presence of textures which may have greater contrast than the contours to be extracted, and 5) variability of size, shape and texture even within a specific specimen type. The image characteristics often provide inadequate histogram modes (global or local) or edge information for segmentation [34],[35]. Multi-thresholding technique can be applied to distinguish between false edges, but results are sensitive to the selected threshold [36] [37] [38]. Filtering may be used to reduce the effects of texture and artifacts but they also tend to distort the contours and substantially lower its contrast because of the amount of filtering required [35]. In some cases better results may be obtained if the image is re-quantized into fewer gray levels using a vector quantizer [39]. The extracted edges are generally fragmented and complex edge-linking algorithms have to be used to extract contours [40] [41]. Region-based methods have the advantage of not being as sensitive to noise and textural variation. However, their major drawback is that usually complicated split-and-merge operations are required, with a well defined rule to determine how far to proceed with the algorithm [42], [43].

5.2 Edge Detection using Wavelet Transform

Edge points are located where the image intensity has sharp transitions. In computer vision, a large class of edge detectors look for points where the gradient of the image intensity has a modulus which is locally maximum. Canny's [29] edge detector is a multiscale version of this approach. It has been shown by Mallat [20] that multiscale edges can be detected and characterized with a wavelet transform. For a particular class of wavelets, the local maximum of the wavelet transform modulus provides the edge locations in the image.

A smoothing function $\theta(x,y)$ is defined as:

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\theta(x,y)dxdy = 1 \tag{142}$$

$$\lim_{x,y\to\pm\infty}\theta(x,y) = 0 \tag{143}$$

Suppose that $\theta(x,y)$ is once partial differential. Let ψ^a and ψ^b represent the first partial derivative of $\theta(x,y)$.

$$\psi^a(x) = \frac{\partial\theta(x,y)}{\partial x} \tag{144}$$

$$\psi^b(y) = \frac{\partial\theta(x,y)}{\partial y} \tag{145}$$

Because the integrals of ψ^a and ψ^b are equal to zero,

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\psi^a(x)\psi^b(y)dxdy = 0 \tag{146}$$

these two functions can be considered to be wavelet functions. The associated 2-D dyadic wavelet transform of a function $f(x,y)\in L^2(R^2)$ at position (x,y) using wavelet functions ψ^a and ψ^b in discrete scale $[2^j]_{j\in Z}$ can be written as:

$$W_{2^j}^a f(x,y) = f * \psi_{2^j}^a(x) = 2^j\frac{\partial}{\partial x}(f*\theta_{2^j})(x,y) \tag{147}$$

$$W_{2^j}^b f(x,y) = f * \psi_{2^j}^b(y) = 2^j\frac{\partial}{\partial y}(f*\theta_{2^j})(x,y) \tag{148}$$

which is the first partial derivative of $f(x,y)$ smoothed at scale $[2^j]_{j\in Z}$ in the x and y directions. The two components of the wavelet transform of $f(x,y)$ form a gradient vector which can be used to locate the edge points in 2-D images.

$$\begin{pmatrix} W^a_{2^j} f(x,y) \\ W^b_{2^j} f(x,y) \end{pmatrix} = 2^j \begin{pmatrix} \frac{\partial}{\partial x}(f * \theta_{2^j})(x,y) \\ \frac{\partial}{\partial y}(f * \theta_{2^j})(x,y) \end{pmatrix} = 2^j \vec{\nabla} (f * \theta_{2^j})(x,y) \tag{149}$$

The modulus of the gradient vector is

$$M_{2^j} f(x,y) = \sqrt{\left|W^a_{2^j} f(x,y)\right|^2 + \left|W^b_{2^j} f(x,y)\right|^2} \tag{150}$$

and the angle in the horizontal direction is

$$\Theta_{2^j} f(x,y) = tan^{-1} \left(\frac{W^b_{2^j} f(x,y)}{W^a_{2^j} f(x,y)} \right) \tag{151}$$

A scaling function $\phi(x)$ is introduced:

$$\phi(x) = \sqrt{2} \sum_n h(n) \phi(2x - n) \tag{152}$$

The basic wavelet $\psi(x)$ is related to the scaling function via

$$\psi(x) = \sqrt{2} \sum_n g(n) \phi(2x - n) \tag{153}$$

where $g(n) = (-1)^n h(1-n)$.

The lowpass filter $h(n)$ and the bandpass filter $g(n)$ are required to implement the wavelet transform. The wavelet used in cell edge contour extraction is the same as that introduced in Chapter 4. One stage of the two-dimensional decomposition and reconstruction is illustrated in Figure 33. If the coefficient $h(n)$ has finite length, $h(n) = 0$ for $n < K_-$,

or $n > K_+$, then the corresponding basic wavelet given in (153) has compact support. One is interested in a scaling function with compact support in both the spatial and frequency domain. This necessitates the use of a scaling function which is as smooth as possible in the construction of the wavelet basis on $L^2(\mathbf{R})$. The regularity index of the scaling function represents the smoothness of the scaling function. The order of regularity increases linearly with the support width; when K is small, ϕ and ψ are not very regular. There is a compromise between the degree of regularity and the degree of compactness; ϕ and ψ have more regularity with a larger region of support, but this requires more computation.

The wavelet transform decomposes $f(x,y)$ into $\{S_J^d f(x,y), \{W_{2^J}^a f(x,y)\}, \{W_{2^J}^b f(x,y)\}\}_{J\in Z}$, where $\{S_J^d f(x,y)\}$ is the lowpassed signal to scale J, and $W_{2^J}^a$ and $W_{2^J}^b$ are wavelet difference signals in the x direction and the y direction, respectively. A very sharp intensity change in the cell image signal will be exhibited as a local maximum in the dyadic wavelet transform across several consecutive scales. For image and voice applications, selecting an antisymmetrical wavelet is desirable because it can operate similarly in the right and the left neighborhood of a point similarly. To construct a wavelet $\psi(x)$ equal to the first-order derivative of the smoothing function $\theta(x)$ implies that $H(\omega)$ is real and symmetrical around zero, and $G(\omega)$ must be antisymmetrical around zero and imaginary.

5.3 Cell Contour Extraction

In our experiment we use the wavelet introduced by Mallat in [20]. This wavelet offers much better edge detection because the signal is correlated across scale and the wavelet is an "edge detector".

$$H(\omega) = e^{i\omega/2}(\cos(\omega/2))^{2n+1} \tag{154}$$

$$G(\omega) = 4ie^{i\omega/2}\sin(\omega/2) \tag{155}$$

The corresponding Fourier transform of the scaling function ϕ and wavelet ψ are:

$$\hat{\phi}(\omega) = (\frac{\sin(\omega/2)}{\omega/4})^{2n+1} \tag{156}$$

$$\hat{\psi}(\omega) = i\omega(\frac{\sin(\omega/2)}{\omega/4})^{2n+1} \tag{157}$$

The cubic spline wavelet ψ described in [20], and the corresponding filters are given by $h(n) = \{\ldots, 0, 0.0625, 0.25, 0.375, 0.25, 0.0625, 0, \ldots\}$ and $g(n) = \{\ldots, 0, -0.0008, -0.0164, -0.10872, -0.59261, 0, 0.59261, 0.10872, 0.0164, 0.0008, \ldots\}$. A block diagram of the contour extraction algorithm is shown in Figure 34. The original image, represented at a fine scale, was decomposed into smoothed images at coarse scales and a set of detail difference images. The detail difference signal at each scale is used to detect the local maximum, corresponding to sharp changes in the image, as it propagates across scales. If there exist a local maximum at positions $P_{2^J}(x_0, y_0)$ and $P_{2^{J-1}}(x_1, y_1)$ with large amplitudes on scales 2^J and 2^{J-1} respectively, then these points are possible contour points, if they are spatially close. Since the two points, (x_0, y_0) and (x_1, y_1) are on different scales, the point at (x_1, y_1) on scale 2^{J11} is mapped to scale 2^J, denoted $P_{2^J}(x_1', y_1')$ to facilitate comparison. Two pixels are spatially close if their Euclidean distance is smaller than a given value d, i.e..

$$E[P_{2^J}(x_0, y_0), P_{2^J}(x_1', y_1')] = |(x_0 - x_1')^2 + (y_0 - y_1')^2|^{\frac{1}{2}} < d \tag{158}$$

The original image is decomposed into smoothed and difference images in the horizontal, vertical, and diagonal directions. The modulus of the difference signal at scale 2^J is:

$$M_{2^J} f(x,y) = \sqrt{\left|d^h_{2^J} f(x,y)\right|^2 + \left|d^v_{2^J} f(x,y)\right|^2 + \left|d^d_{2^J} f(x,y)\right|^2} \tag{159}$$

To locate the two-dimensional local maximum, corresponding to the significant image edges, the modulus signal is thresholded using threshold values C_{2^J} and $C_{2^{J-1}}$ for scales 2^J and 2^{J-1}, respectively. Using a window of size $2d$ x $2d$ to match pixels on the two scales, we process the image in two passes. First consider the y direction as constant. The local maximum at $(x_i, \cdot)$ propagating to the next scale within the window must satisfy:

1. $M_{2^J} f(x,\cdot) \geq C_{2^J}$ and $M_{2^{J-1}} f(x,\cdot) \geq C_{2^{J-1}}$

2. $M_{2^J} f(x_i,\cdot) > M_{2^J} f(x_{i+1},\cdot)$ and $M_{2^J} f(x_i,\cdot) > M_{2^J} f(x_{i-1},\cdot)$ for $0 \leq i \pm 1 \leq d$

3. $M_{2^{J-1}} f(x_i,\cdot) > M_{2^{J-1}} f(x_{i+1},\cdot)$ and $M_{2^{J-1}} f(x_i,\cdot) > M_{2^{J-1}} f(x_{i-1},\cdot)$ for $0 \leq i \pm 1 \leq \frac{d}{2}$

The same method is used in the y direction. The local maxima found along the x and y directions are mapped into a composite binary image. The binary image contains the desired contours but also fragmented contours segments due to the presence of artifacts; these appear as protruding branches when the artifacts are touching or close to the desired contours. Additionally, the thresholding and matching of the local maxima on the two scales may also fragment the desired cell contours. These problems are addressed in the post-processing step where the composite binary image is processed by a morphological closing operation which tends to connect contours separated by small gaps; the size of the gap closed is proportional to the size of the structuring element used in the closing operation. Removal of the contours fragmented by artifacts is accomplished by using a morphological pruning operation [44].

5.4 Results

Figure 35 demonstrates the use of wavelet to determine points of rapid signal change in a one-dimensional signal corresponding to the gray level along one row of a cell image. Figures 35(b) and (c) show the wavelet transform of the original signal at two subsequent scales. A low-pass filter was employed to remove the high frequency components in the DWT results and make the local maxima more clear. A sharp change in the signal in the original signal of Figure 35(a) is detected if there exist peaks at the same locations of Figures 35(b) and (c). Original images from cytology were obtained from the Department of Pathology at the Mount Sinai Medical Center [37] at a magnification of 100X. Our application on these types of images, which are primarily of isolated cells, indicate that only the first and second scales of the wavelet transform are sufficient to extract the desired contours. The algorithm requires selection of the threshold values C_{2^J} and $C_{2^{J-1}}$, which were empirically determined. Threshold value C_1 is set to $65 - 70\%$ of the local maximum in the first scale and C_2 is set to $50 - 55\%$ of the local maximum in the second scale. A 3x3 structuring element was used for the morphological filtering in the post-processing. With this structuring element the morphological closing algorithm connects contour segments separated by one pixel.

Figures 36 and 37 demonstrate application of the wavelet transform to cell contour extraction. The original cell in figure 37(a) has dark artifacts, some of which are touching the desired contour. Also, the nuclear and cytoplasm regions have textures which may present a problem during segmentation. To reduce the effects of the texture and background artifacts we preprocess the image by requantizing to three levels via the K-means algorithm. Three levels were chosen to coincide with the background, nuclear, and cytoplasm regions. Figure 36(b) is the preprocessing result. Figure 36(c) is the modulus of the wavelet transform. The

points corresponding to the local maximum of the wavelet transform is shown in Figure 36(d). The isolated points and branches are trimmed by application of a morphological pruning operation. Longer fragmented contours are removed by application of a length filter. The extracted contour is shown in Figure 36(e); for comparison, the extracted contour is superimposed on the original image in Figure 36(f).

Figure 37(a) shows a type of cell image which has relatively "constant" background with little if any background artifacts which would interfere with the segmentation. However, there are multiple cells, some of which are very close or are touching one another. In this case the wavelet filters are applied without preprocessing. Figure 37(c) shows the modulus of the wavelet transform. Local maximum binary image is shown in Figure 37(d). Figure 37(e) shows the extracted cell contours after pruning and 37(b) shows the original image with the extracted contours superimposed. We notice that cells on the right side of the image which are touching or have minute separation cannot be individually segmented using the wavelet filters. Additionally, other cell contours are not detected.

Figure 37 demonstrates the shortfalls of this approach. Contours that are very close or touching can not be extracted by this method. The wavelet filters are in essence low pass filters which tend to merge structures, thereby obliterating their individual contours. Also, if the separation of a fragmented contour is greater than one pixel, then the contour remains fragmented and the subsequent pruning operation removes the contour. Contour fragmentation is caused by using a window with a small size and/or setting the threshold value too high. In our experiment, we only connect fragmentations that are one pixel wide. It is possible to link fragmentations that are n pixels wide, but the trade-off is that the probability of incorrectly connecting two fragments increases with n.

5.5 Conclusion

The wavelet transform can be used to extract the sharp gray level change which represents the edges in an image. The difference between the wavelet transform at two consecutive scales can be used to detect the sharp change in the original signal. This chapter has demonstrated the application of compact support wavelet to cell contour extraction. If the image is corrupt with background artifacts, or if the nucleus and cytoplasm contain texture, preprocessing the image via the K-means algorithm with three levels facilitates segmentation. All the cell and nuclear edges can be extracted simultaneously. However, cells that are touching or very close together cannot be extracted by this technique.

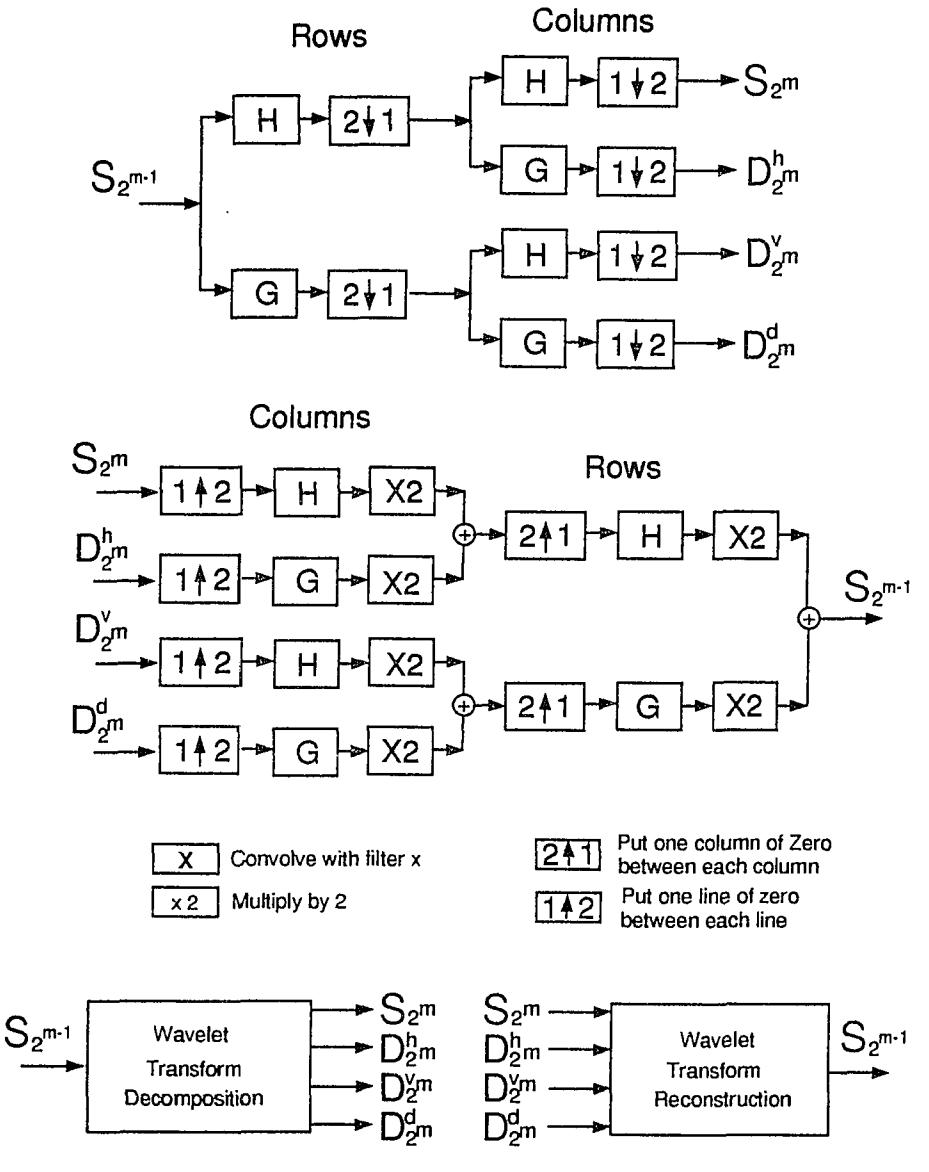


Figure 33. One stage of two-dimensional decomposition and reconstruction

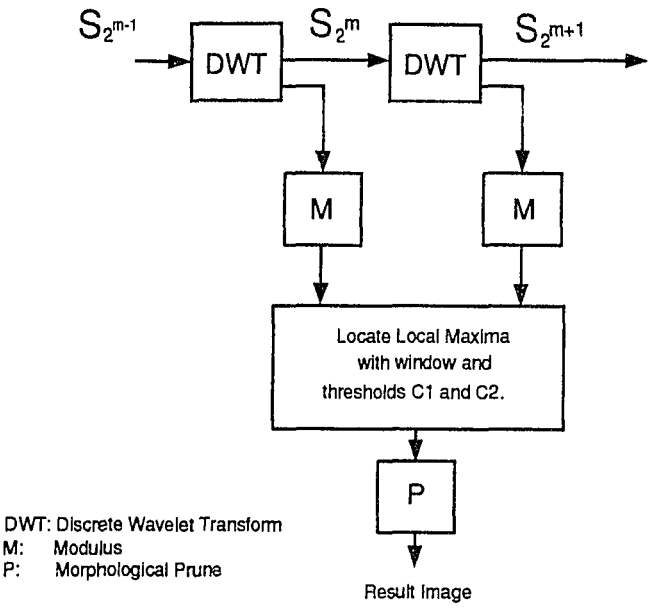


Figure 34. Block diagram of cell contour extraction algorithm

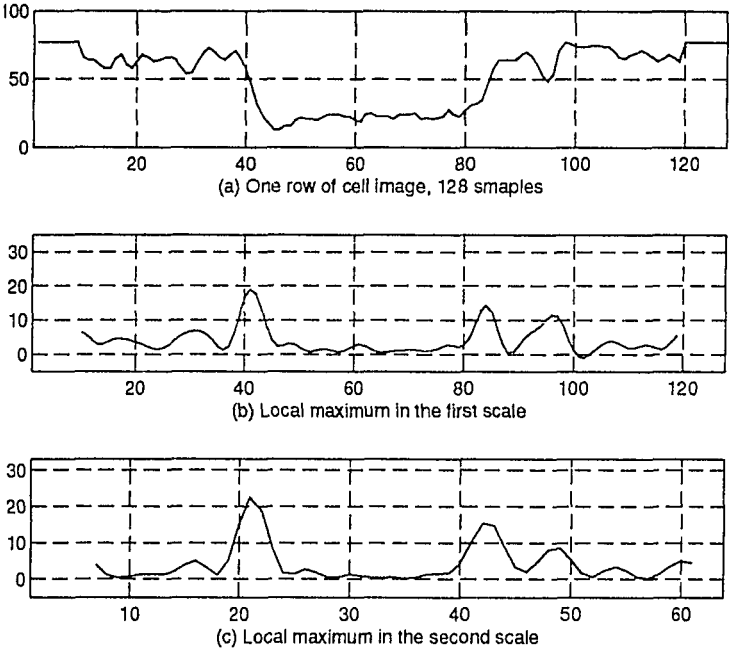


Figure 35. One row of cell image in 128 samples is in (a). Local maximum in Scale 1 and 2 are shown in (b) and (c).

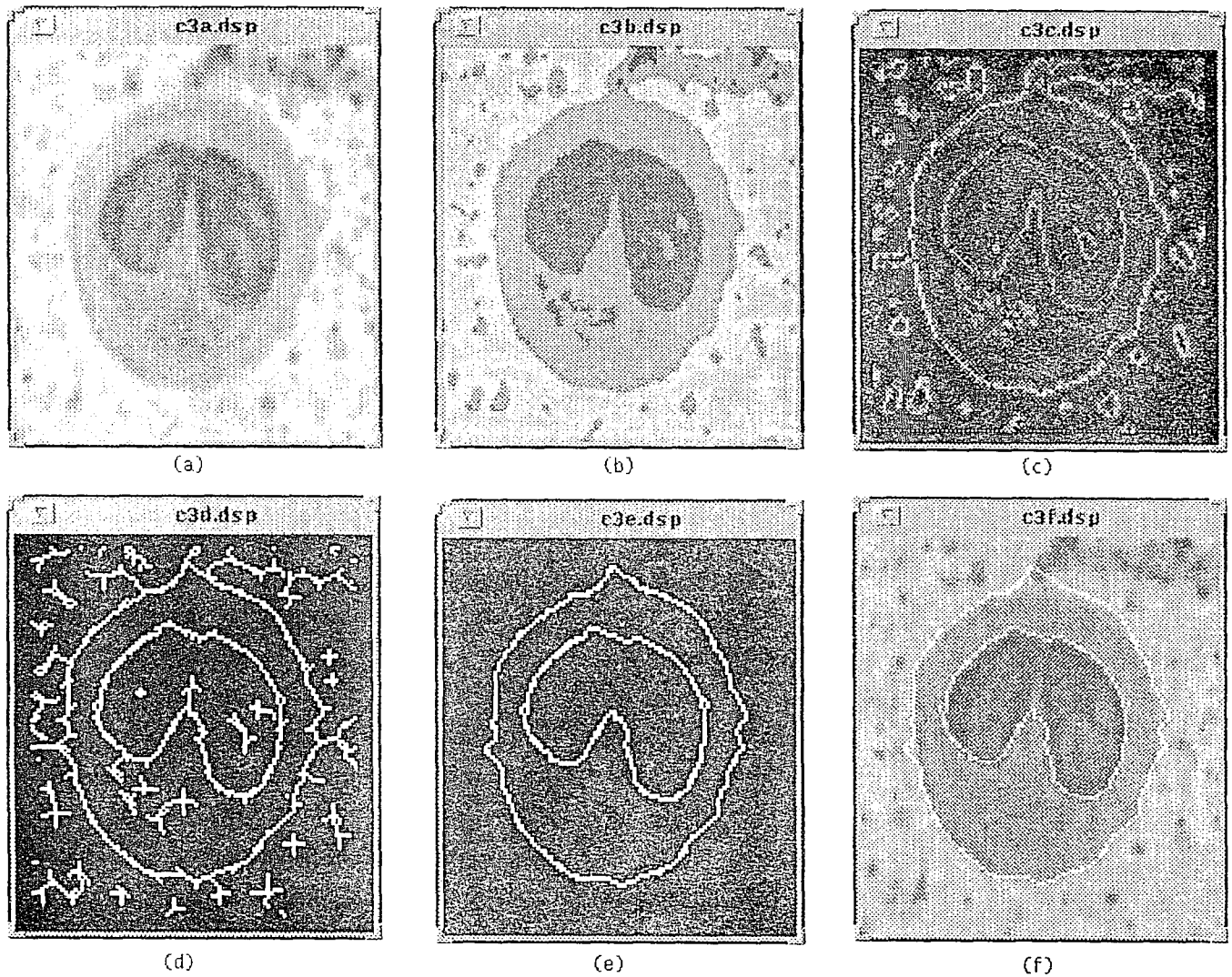


Figure 36: One cell contour extraction

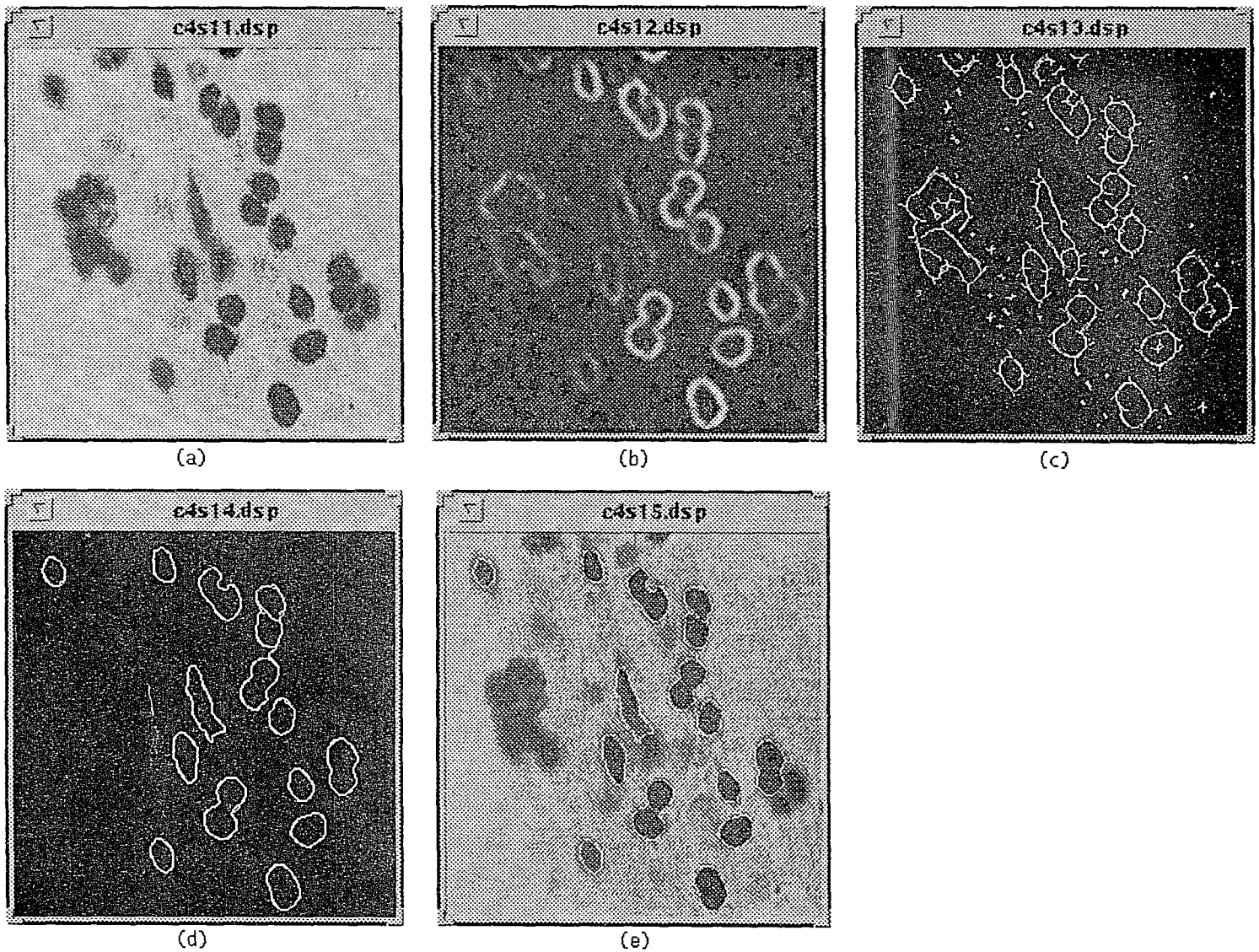


Figure 37. Cell contours extraction (2)

6 Contrast Enhancement Cell Contour Extraction

6.1 Introduction

A segmentation algorithm using wavelet transform and enhancement to extract cell contours is here introduced. The enhancement is applied to the difference signal of the wavelet transform to increase the contrast of cell structure with background. The regions containing cell edges are detected on each scale of the wavelet transform. A median filter is applied to remove isolated variations. A morphological dilation operation is applied to correct the cell regions. The experimental results in fuzzy cell segmentation are provided and the results demonstrate that the proposed algorithm is simple and successful.

Identifying the edges of low contrast structures is one of the most common tasks performed in medical image processing. It is commonly practiced in digital medical imaging areas such as magnetic resonance (MR), ultrasound, computed tomography(CT), nuclear medicine, and others, etc. In certain digital medical imaging procedures, such as CT, X-ray, high contrast images can be obtained by rasing the number of photons at the time of examination. In other areas, it is still very expensive to obtain high contrast image, especially may have to be processed again. Image enhancement clarifies details within an image. Many different enhancement methods have been proposed; among these post-enhancement processes, linear intensity scaling, clipping, and histogram equalization are most commonly used [44]. Linear intensity scaling and clipping maps a partial range (window) of the grayscale to the dynamic range of the display device. Prior knowledge or intensive human interaction is needed to choose the right window. Histogram-modeling modifies an image so that its histogram has a desired shape, but it is not very effective for images containing large homo-

geneous regions. Filter-based approaches use "unsharp masking" to amplify the magnitude of edge gradients so that the intensity on both sides of the edge is increased. Leu [45] increased the intensities difference between neighboring regions of nearly constant intensity. A histogram equalization function from pixels on edge contours is driven and applied the same transformation to the entire image. The contrast increase must be perpendicular to the edge contours. In Beghadai and Le Negrate's method [46], a spatial localization of edges was constructed and used to index the design of intensity transform functions. Pixel intensities on the different sides of the edge contour were changed by using the transform function.

Edge-based segmentation methods are based on the idea of image intensity discontinuities along the boundaries between different objects [47], [29]. For a fuzzy structure boundary, image intensity discontinuity information is weak. Enhancing the edge information may help to locate precisely the image intensities at structure boundaries. Region-based algorithms used in cell contour extraction [48], [49] employ region growth and merging to separate different objects with homogeneous intensities.

Lu et al. [50] proposed a multiscale edge enhancement method using non-orthogonal wavelet bases with stretching factors on each scale to stretch the gradient maxima of the image. A reconstruction algorithm for non-orthogonal wavelet bases with stretching is needed to reconstruct the approximation of the original image. In our application of non-orthogonal wavelet, we have observed a correspondence between the details of non-orthogonal wavelet transform and variations in the original image. This observation has led us to develop a contrast enhancement technique for locating blurred cell edges. The following presents a simple and effective method for region-adapted local contrast enhancement.

In some types of cell images, the background is relatively homogeneous without many artifacts and variations in the background are relatively smaller, the boundaries of the cells are blurred. Typically, cells get darker around the nuclear region of the cell. The method we propose here is based on enhancing the wavelet difference signals to locate the regions which contain more variations. By contrast-enhancing these regions, cell contours can be more effectively extracted. A threshold is selected for each wavelet transform scale in order to reduce noise.

6.2 The Method

In Mallat and Zhong's [20] introduction of a wavelet edge representation, local maxima and gradients were defined based on the wavelet transform difference signals. Structure edge can be located based on the local maxima on consecutive wavelet scales. The magnitude and the location of non-orthorognal wavelet transform difference signals characterize variations in the original signal. The locations indicate where the change occurred, and the magnitude of the difference signal represent how sharp the change is. Lu [50] introduced a method which enlarged the magnitude of edge gradients to enhance the difference between objects. The observing this wavelet difference signal led us to develop a new method for enhancing the cell edge and then extracting the cell contours. With multiscale edge representation, the edge features of location, magnitude, gradient and sign enable us to emphasize the changes in the image. Certain spatial regions within spatial/scale domain may be located. These regions contain the edge information.

As before, a multiscale wavelet transform decomposition of a given image $f(x, y)$ can be written as:

$$\{S_2^J f(x,y), \{W_{2^j} f(x,y)\}, 1 \leq j \leq J\} \tag{160}$$

where S_{2^j} is the lowpassed subimage of a given image $f(x,y)$. $W_{2^j} f(x,y), 1 \leq j \leq J$ are the wavelet transform difference signals on a scale from 1 to J. A stretching factor $k_j \geq 1$, for $1 \leq j \leq J$, can be selected for each scale of the wavelet transform. The stretched difference signal set simply multiplies the stretching factor to each difference value.

$$k_j \cdot W_{2^j} f(x,y) \tag{161}$$

The stretching factor can be a constant on each scale, $k = k_1 = k_2 = \cdots = k_n$, and can also be varied on each scale. Our goal is to find the regions on different scales where sharp changes occur. Figure 38 shows the change when applying the different stretching factor on lowpassed signal and/or difference signal. The original signal with 256 samples is shown in (a). In (b), (c), and (d), we show the reconstructed signals which the stretching factor is applied on decomposed signals. In (b), the stretching factor $k = 2$ is applied on both average and detail signals. It is almost the same as multiplying (a) by 2. In (c), we only apply stretching factor $k = 2$ to the detail signal on scale 2 and 3. In (d), we only apply $k = 2$ to the average signal. One can find out stretching factor applied on the difference signal will affect the detail variation. To illustrate our algorithm, one row of a cell image is shown in figure 39(a). It contains 512 samples and crosses two cells. The mean of the sample signals has been set to zero. The wavelet difference signal, with the position, magnitudes, and sign, is shown in Figure 39(b). Then we only stretch those difference signal greater than a threshold C_t by stretching factor k, see (c). Directly multiplying by the stretching factor, the wavelet difference signals not only enhance the structural contrast

with neighboring regions, but also increase the variation in the whole image. To efficiently control the enhancement, a threshold is used to block out the lower variation. A threshold c_j, for $1 \leq j \leq J$, is defined for each wavelet transform scale. On each scale, the difference signals lower than the threshold are removed. The stretching factor multiplies the different signals which is greater than c_j. We use $R^S_{2^j}$ to represent the stretched difference signal which greater than threshold c_j on each scale $1 \leq j \leq J$. $R^S_{2^j}$ is given by:

$$R^S_{2^j} = \begin{cases} k_j \cdot W_{2^j} f(x,y) & \text{if } |W_{2^j} f(x,y)| > c_j \\ 0 & \text{otherwise} \end{cases}$$

Most non-zero values of $R^S_{2^j}$ are located at the region where relatively more sharper changes occur. Regions having denser non-zero values in $R^S_{2^j}$ correspond to cells in the original image. There may still exist some isolated or relatively smaller regions with non-zero values in $R^S_{2^j}$, but the corresponding density is much less. These isolated non-zero values may represent noise, small structure, or unsharp changes in the image. We are more interested in locating those regions with denser non-zero values in $R^S_{2^j}$. A median filter is an ideal tool to remove those isolated non-zero values. We use MF to represent the median filter operator and use $R^m_{2^j}$ to represent the result of median filtering of $R^S_{2^j}$. The $R^m_{2^j}$ is defined as:

$$R^m_{2^j} = \begin{cases} 1 & \text{if } MF(R^S_{2^j}) > 0 \\ 0 & \text{otherwise} \end{cases}$$

Figure 39(d),(e) show the result after applying the median filter on the scale 2 and scale 3, respectively. As the wavelet transform scales increased the average signals are more smoothed and $R^m_{2^j}$ for their scales vary less. We use R^A to represent the union of the $R^m_{2^j}$, where $1 \leq j \leq J$.

$$R^A = \bigcup_{j=1}^{J} R_{2^j}^m \tag{162}$$

Figure 39(f) shows the union of (d) and (e). R^A represent the area in which the relative sharper change has occurred in the image. These areas form a thick belt around the cells but exclude the center of the cells. This is due to threshold c_j on each scale not only removing the small variation in difference signals but also eliminating the changes around the zero-crossing. In order to fill in the center of the cells, we use the minimum of the smoothed signal to locate the seeding position and apply a morphological dilation operation. The dilation is only applied to the minimum of the smoothed signals. The dilation operation cannot be applied to R^A directly because it would not only fill in the center but also enlarge the area on the outside. If cells are very close, this operation may connect the cells and make it more difficult to extract the separate cell contours.

The R^A after the dilation operation is:

$$R^A = R^A \cup DIL(min(S_{2^J}(x, y)) \tag{163}$$

The final contours of cells can be obtained by subtracting the morphological erosion of R^A from R^A.

$$\text{Cell Contour} = R^A \text{ - } ER(R^A) \tag{164}$$

where DIL is the morphological dilation operation and ER is the morphological erosion operation.

Because of the down-sampling of wavelet transform, the subimage is one-quarter the size of the prior scale. We simply double the size of the lowpassed subimage to make it

comparable to the prior image.

6.3 The Algorithm

The outline of the computational steps for enhanced cell contour extraction is given below. The wavelet we used in the experiment is introduced in the work of Mallat and Zhong [20], same as in Chapter 5.

1. A non-orthogonal wavelet transform is applied to the cell image to obtain the image wavelet decomposition representation:

$$\{S_{2^J}, \{Wj(x,y)\}, 1 \leq j \leq J\}$$

2. The magnitude of the wavelet difference signal is stretched by a stretching factor k_j at each scale only if the signal is greater than the threshold value c_j. Others magnitudes will not be stretched.

3. A median filter is applied to the result of step (2) to remove isolated variations and to detect the denser area.

4. The area corresponding to the cell contour is formed by the union of all results of step (3) for each scale.

5. Passible holes in the centers of cells are filled in by applying morphological dilation operation on the minimum of smoothed signal.

6. Cell Contours are extracted from the result of step (5) by using Erosion and Subtraction.

6.4 Experiment Results

Two images with blurred cells of size 512 x 512, in Figures 40(a) and 41(a), are tested in our experiment. The median filter window size we use in our experiment is 5 x 5. And we use a constant stretching factor $k = 2$. Comparing these two images, the cells in figure 41(a) are more fuzzy than those in figure 40(a), and are not textured. Two scales of the wavelet transform are performed to extract the cell contours in Figure 40(a). Figure 40(b) is the stretched difference signal in scale 1; the result of median filtering is shown in Figure 40(c). Figures 40(d) and (e) show the stretched difference signal and result of median filtering in scale 2, respectively. The union of (c) and (e) is shown in Figure 40(f). Figure 40(g) shows the result of filling these holes using morphological dilation. The extracted cell contours are shown in Figure 40(h). The original cell image superimposed by the extracted cell contours is displayed in Figure 40(i).

In the second image the cell edges are more blurred and require more wavelet transform. Six scales of the wavelet transform are performed. The original cell image is shown in Figure 41(a); R^A for $J = 3$ is shown in Figure 41(c); in Figure 41(d) is shown R^A for $J = 4$; in Figure 41(e) is shown R^A for $J = 5$; and in Figure 41(f) is shown R^A for $J = 6$. The result after filling in the center holes is shown in Figure 41(g). The final extracted cell contours are shown in Figure 41(h). To show the performance of the segmentation visually, the cell image superimposed by the extracted cell contours is displayed in Figure 41(b).

6.5 Discussion and Conclusion

The method we propose here combines the robust characteristics of both edge-based and region-based segmentation methods. The variations of the wavelet transform can be used as pattern features to discriminate cell regions from their background. It locates

sharp changes based on wavelet transform difference signals, which are then mapped to a special domain with homogeneous intensities. This method has the flexibility to selectively enhance features of different size by providing scale-variable stretching factors, and also has the ability to control noise magnification by thresholding.

To extract contour of blurred cells with texture is also a difficult topic because the intensity difference used as a dissimilarity measure in many existing segmentation techniques in not suitable for textured cell image. A simple thresholding method may indicated a boundary for fuzzy cells, but also may include the background intensity range. Using local variation provides a more efficient method to extract fuzzy textured cell contours because wavelet multiresolution analysis can "zoom in" as long the resolution changes and detect more details on the local variation. In Sec. 5, a regular cell contour extraction method using wavelet multiresolution method is introduced. One of the shortfall in this method is that the cell contour may appear disconnected if the original object has a blurred edge. In Figure 37 (a), the edge of the object in the middle left portion of the picture is not very clear. Such fragment contours may be extracted by using the method introduced in Sec. 5. Whereas the application of several erosion operation may remove the branches along the extracted cell contours, the fragmented contours will be removed also. The contrast enhancement method can overcome this shortfall. Figure 42(a) is a subimage cropped from Figure 37. The subimage is of a cell object with a contour which cannot be extracted properly using regular wavelet multiresolution method but which can be extracted by using the contrast enhancement method. This subimage is resized to 128 x 128 in order to process the contrast enhancement. Figure 42 (b) shows the result of R^A while $J = 3$. The extracted cell contour is shown in Figure 42 (c). The cell image superimposed by the extracted contour is displayed

in Figure 42(d).

The morphological or gray-level morphological method is sensitive to image noise. The morphological dilation operation may enlarge the homogeneous region, but on the other hand, it is non-linear operation. The dilation operation may cause cell connection if cells are very close and it is impossible to separate the cells using an erosion operation.

The experimental results of the proposed algorithm for extracting fuzzy cell contours are provided. The extracted cell contours are superimposed on the original images. The results show that the extracted cell contours fill the cells in the original image very well. The algorithm can successfully extract blurred cell contours. In the case of those cells which are more textured, we observe that the homogenous area detected after applying the median filter is thicker than that of non-textured cells. This indicates that this method is also suitable for extracting textured cell image contours.

The method presented here is based on stretching the wavelet transform coefficients on different scales and mapping them to a binary domain to locate the boundaries of blurred cell contours. It is also possible to stretch all the wavelet transform coefficients and smoothed signal, then reconstruct to achieve an image with enhanced contrast and clarity. Figure 43(a) is a 256 x 256 image cropped from a digitized pelvis x-ray image. Both pelvis and vessels are low-contrast in this image. Two levels of wavelet decomposition are applied to the image. Not only are the different signals on different scales stretched, but also the smoothed low-passed signal. The reconstructed image shows the objects in this image clearly enhanced, with both pelvis and vessels more clearly defined. Compares as follows with the γ enhancement method. Different γ value are selected. Figure 43(b), (c) and (d) show $\gamma = 0.25$, 0.50 and 0.75, respectively. It can be seen that neither pelvis nor vessels

are satisfactorily enhanced. when the γ value has been changed. Figure 43(e) and (f) are the reconstructed images after stretching the wavelet coefficient and the smoothed with different values. In Figure 43(e), the stretching value $k_1 = 2$, $k_2 = 3$, and the smoothed signal is stretched by $k_0 = 2$. In Figure 43(f), the stretching value $k_1 = 3$, $k_2 = 3$, and $k_0 = 1.5$, respectively.

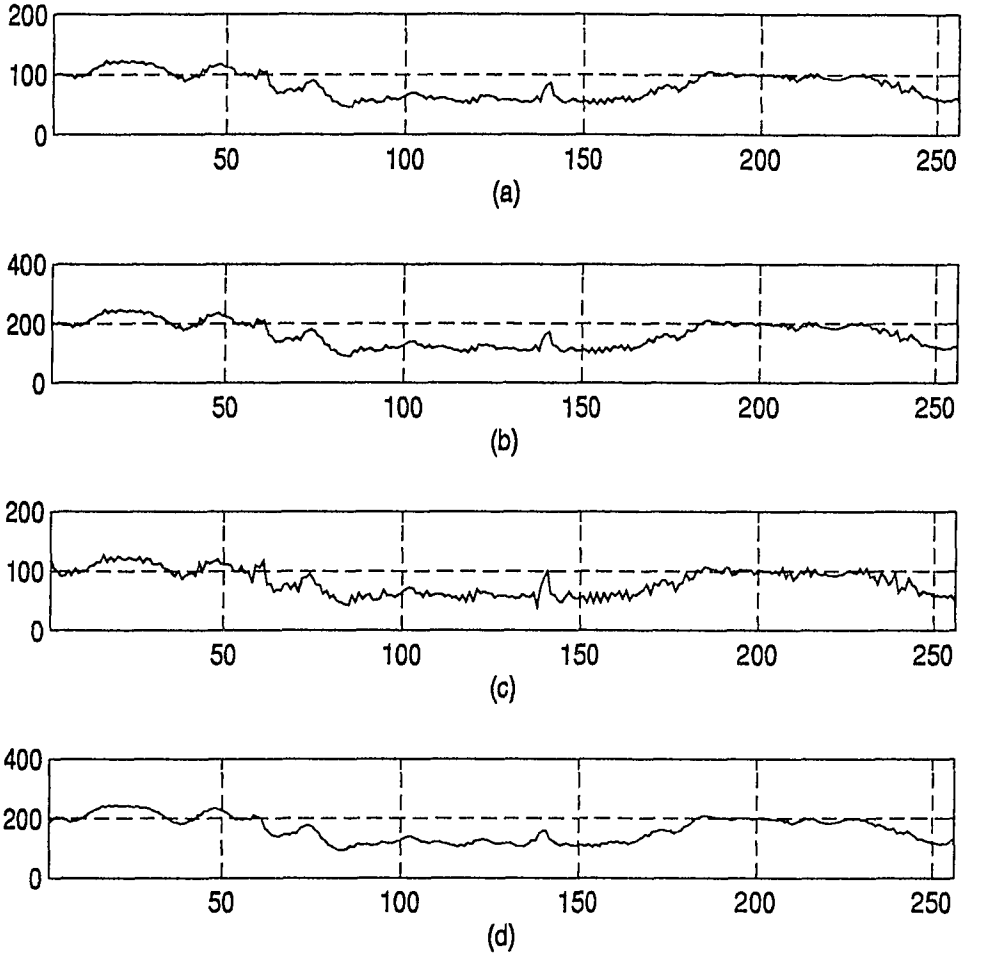


Figure 38. (a) Original signal with 256 samples (b) original signal stretched by factor 2. (c) Reconstructed signal from wavelet decomposition, the different signal on scale 2 and 3 are stretched by factor k = 2. (d) Reconstructed signal from wavelet transform, the lowpassed signal is stretched by 2.

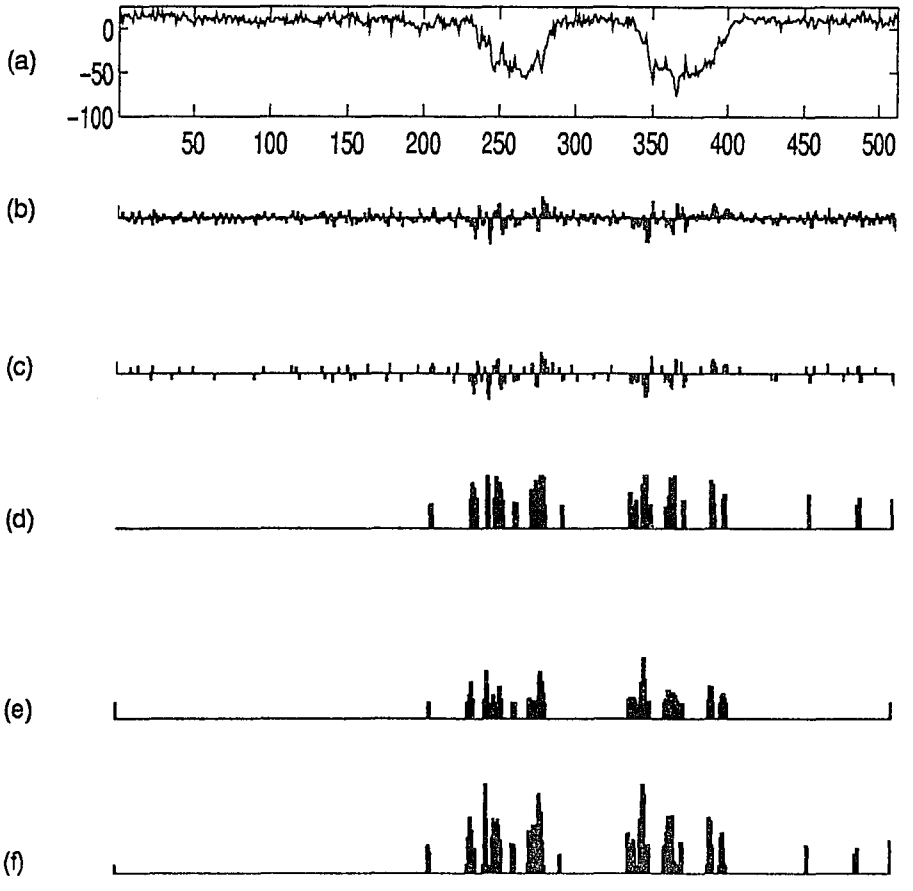


Figure 39 (a) Original signal, 512 samples. (b) wavelet different signal for the first scale. (C) Stretched (b) by factor k1 = 4, (d) the result after applying median filter. (e) the result after applying stretch with factor k2 = 4 and median filtering. (f) the union of (d) and (e).

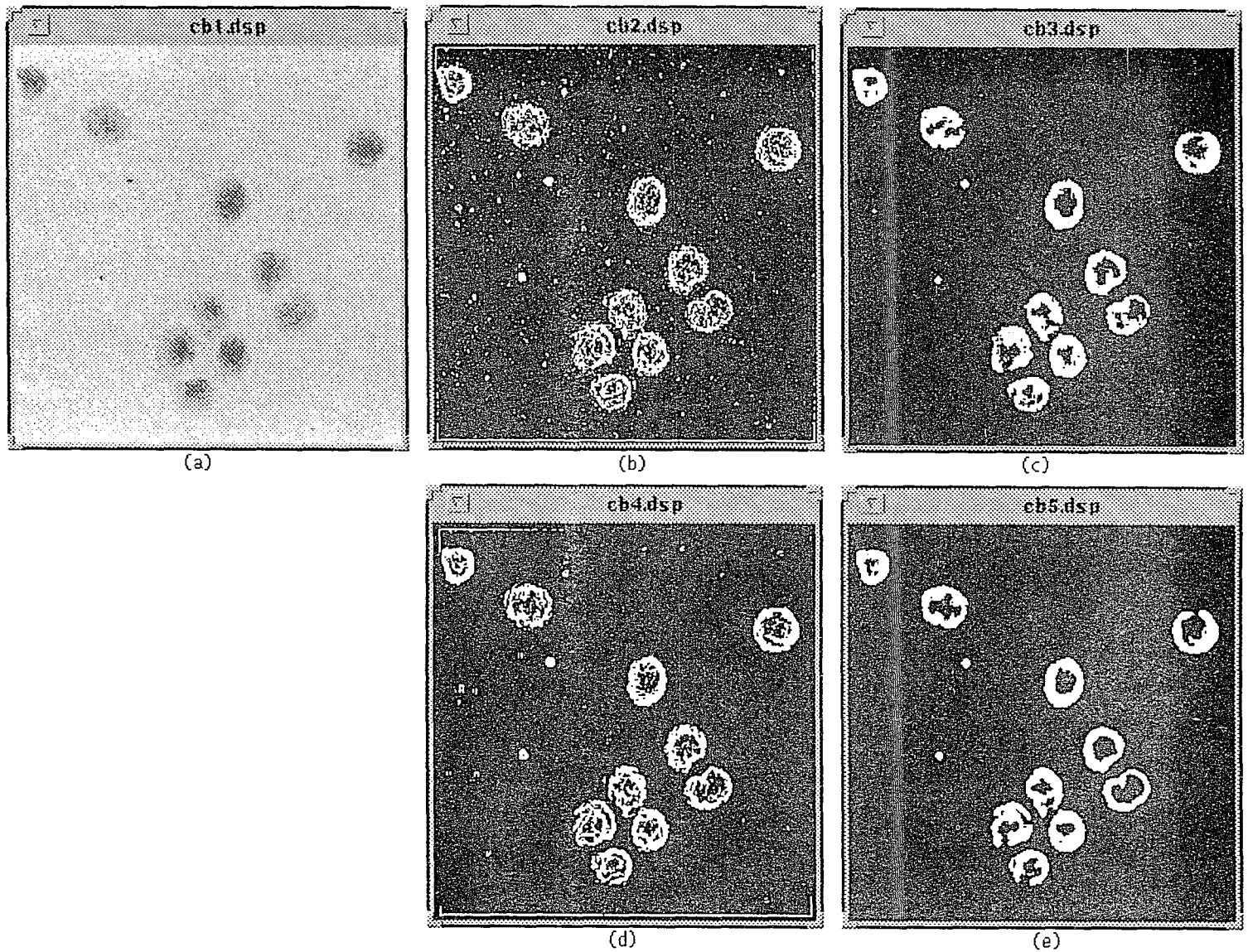


Figure 40. (a) -- (e) Cell contours extraction (3)

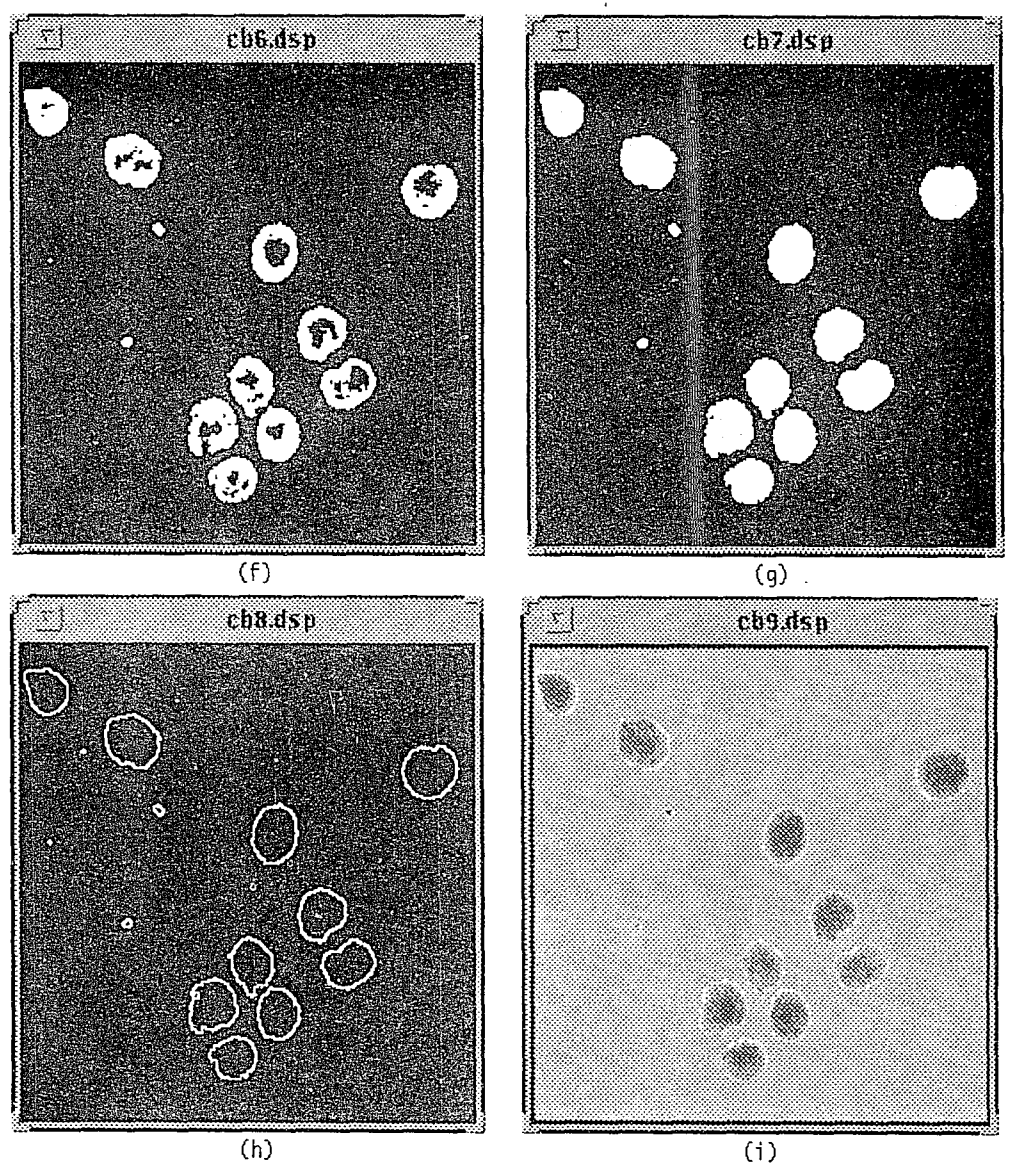


Figure 40, (f) -- (i) Cell contours extraction (3)

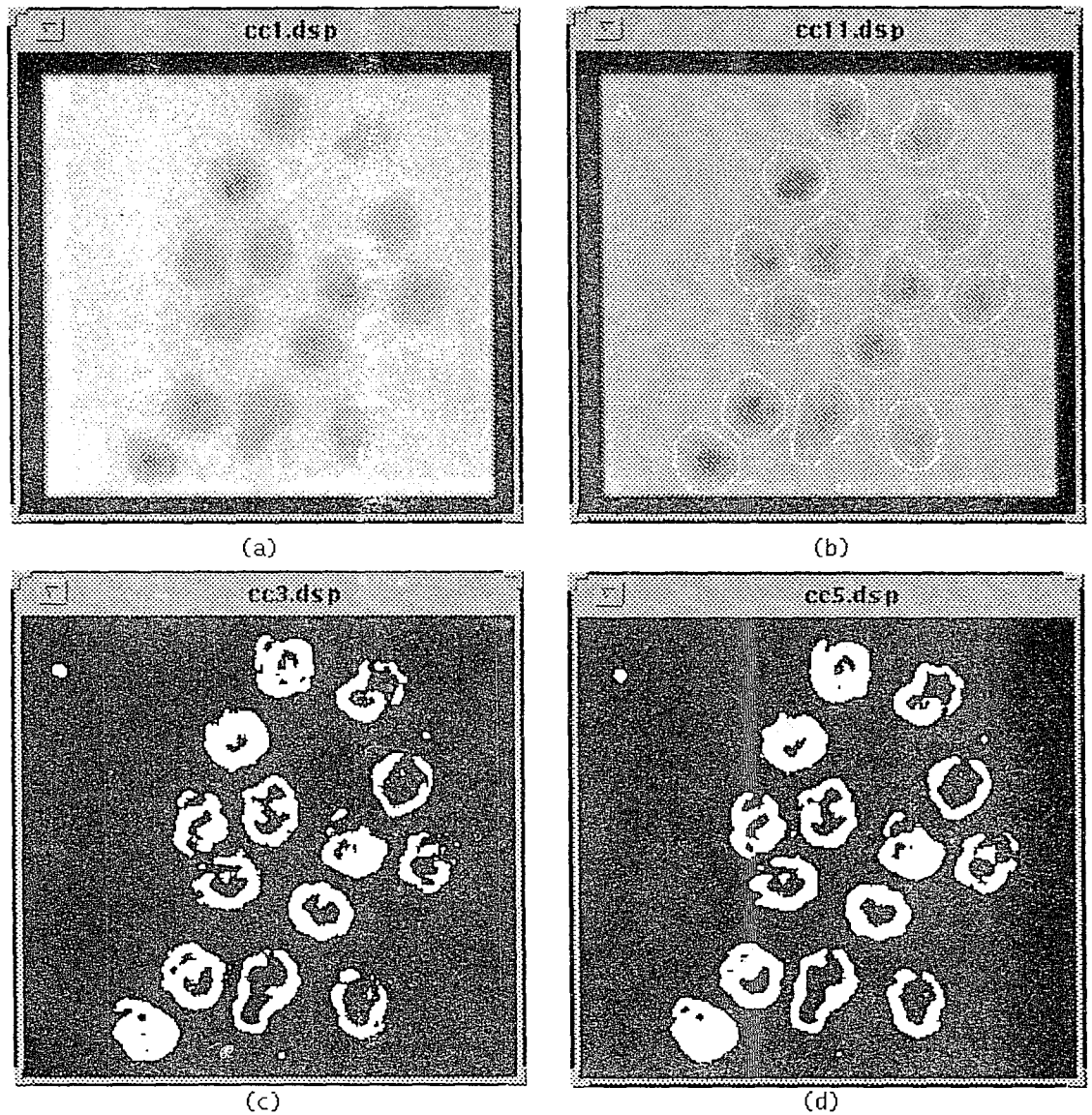


Figure 41, (a) -- (d) Cell contours extraction (4)

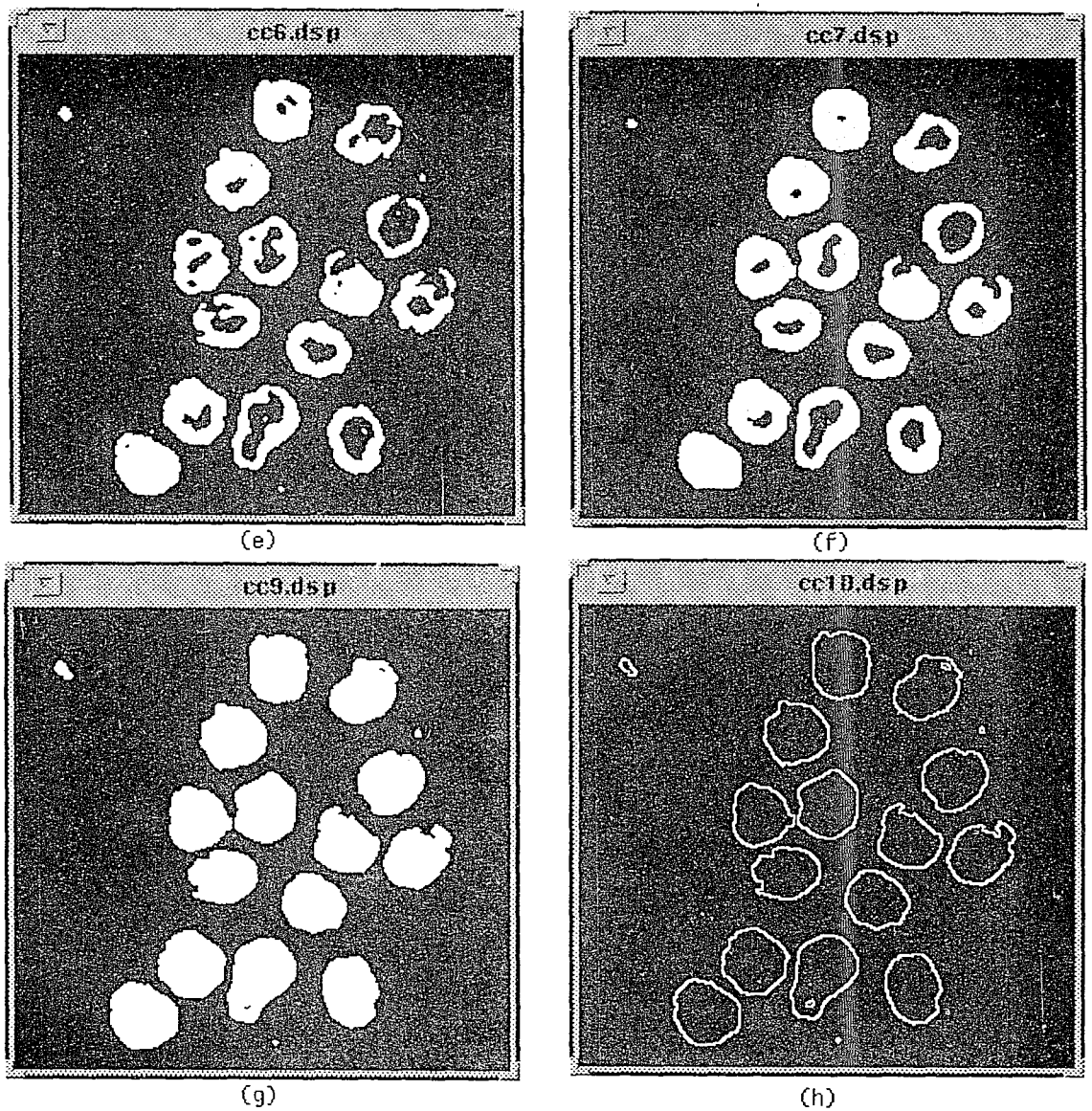


Figure 41, (e) -- (h) Cell contours extraction (4)

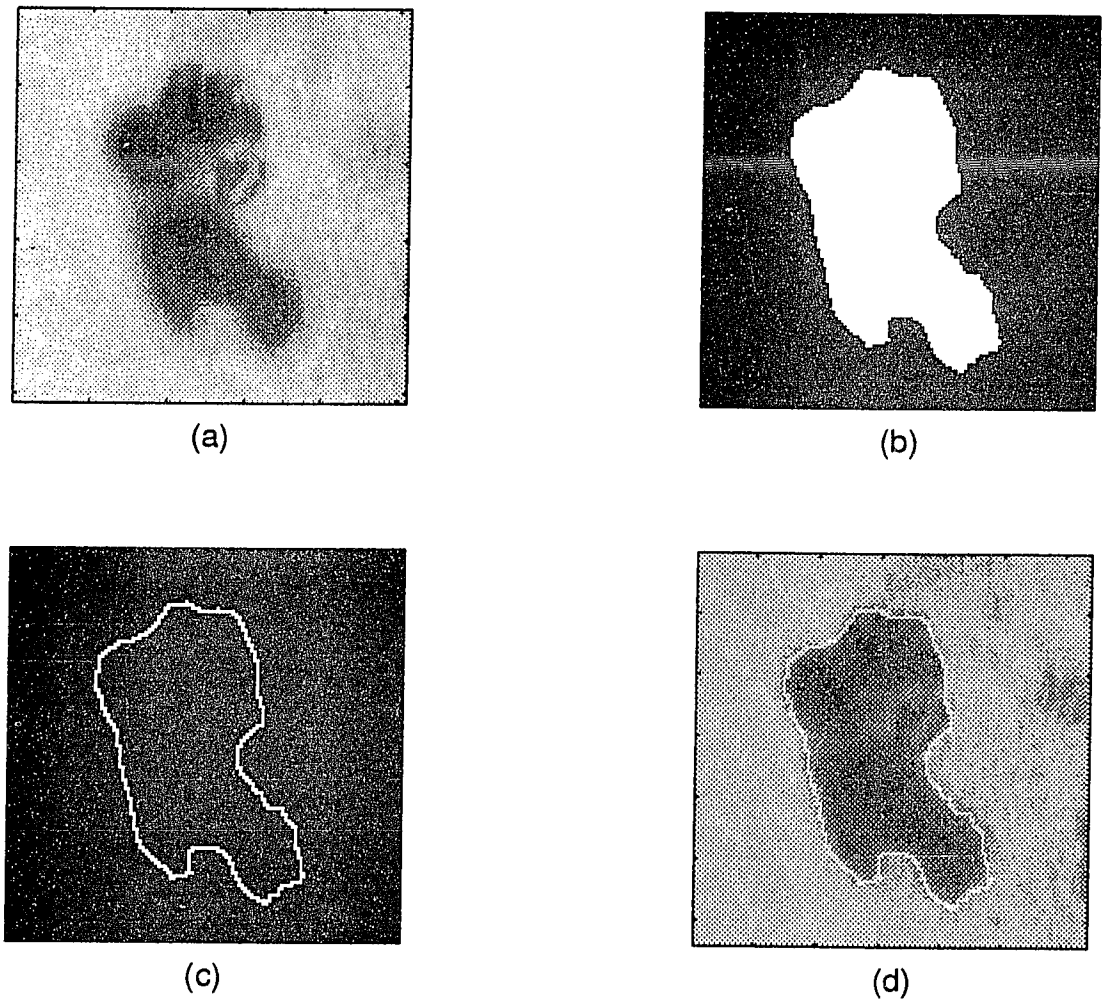
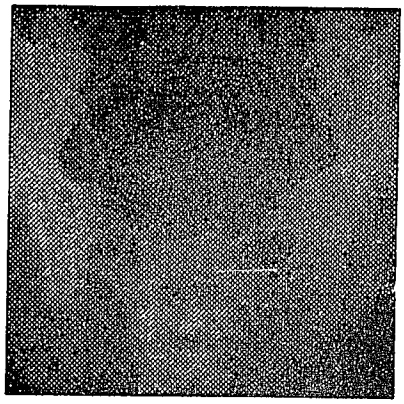


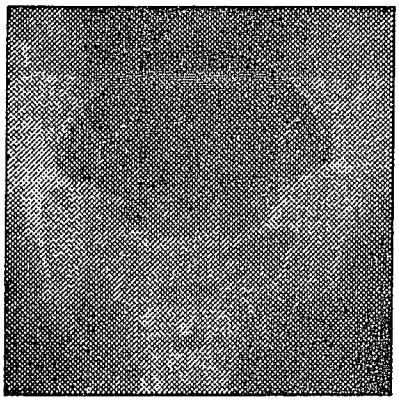
Figure 42: Subimage from Fig. 37 and extracted contour



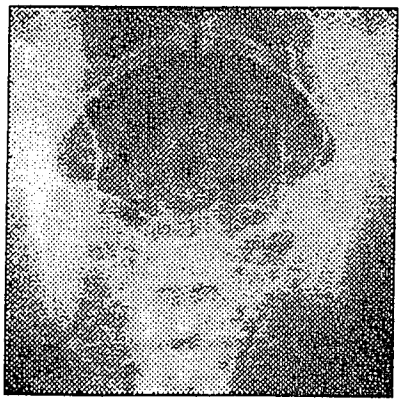
(a) Original Pelvis (256x256)

(b) Gamma = 0.25

(c) Gamma = 0.50



(d) Gamma = 0.75



(e) Reconstructed pelvis

(f) Reconstructed pelvis

Figure 43. Low-contrast pelvis and enhanced result

7 Texture Segmentation

Texture energy representation using wavelet transform is a new approach which can be used in texture classification and discrimination. In this chapter, we develop a texture segmentation method based on the wavelet transform energy pattern. The representation features at each scale are obtained by decomposing a signal (image) on to an orthonormal basis, thus avoiding correlation between scales. Direction selectivity is built into the two-dimensional orthonormal wavelet. Test images, which contain structured or/and unstructured images, can be classified and segmented into homogeneous textural regions. Also, features of rotated textures are extracted and compared with those of the original to demonstrate the change of features. Examples are provided that demonstrate the success of the algorithm.

7.1 Introduction

Image segmentation is the process of separating an image into regions that are homogeneous with respect to particular characteristics, such as gray tone, texture, or color. It assumes that the statistics of each region are stationary and each region extends over a significant area. In each region the local statistics are constant, or slowly varying, or approximately periodic [51]. When processing small neighborhoods, the differences in the mean gray level, color, or apparent texture are not always sufficient to distinguish the differences of textures. On the other hand, using large block subimages may also lead to incorrect segmentation when the window function overlays texture boundaries. The actual texture boundaries are not easy to detect in either case, and therefore texture segmentation needs both local and global information.

In the mammalian visual system, the visual signal is received by the retina and then transmitted through the optic nerve and into the lateral geniculate nucleus and striate cortex. The major known classes of neurons, leading from the retina to the early stages of visual cortical processing in the striate cortex, have filtering properties that accomplish on increased conjoint spatial and spatial-frequency resolution. The visual stimulus is processed in parallel by a number of independent spatial-frequency channels in the striate cortex. Experiments on human and mammalian vision support the notion of spatial-frequency(multiscale) analysis that maximizes the simultaneous localization of energy in both the frequency and spatial domains [52] [53]. The mammalian visual system can preattentively segment textures robustly, performing some form of local spatial-frequency analysis on the retinal image. Experimental analysis is done by a bank of bandpass filters [53], [54]. Statistics based on second-order distribution of gray levels can be used as discriminators in identifying texture.

A number of texture segmentation methods have been proposed [55],[56], [57], [58], [21], [59], [60], [61]. These techniques have demonstrated substantial success for a wide variety of images. Generally, the texture representation approaches can be divided into two categories [62], [63]: statistical approaches and spatial/spatial-frequency approaches. In statistical approaches random fields, generally a Markov Random Field(MRF), with smooth spatial characteristics, the discrete valued field containing the classification of each pixel in the image. A region in the image is distinguished from the neighborhood regions if there exists sufficient evidence to separate these regions. Each point in an MRF is dependent on its neighbors, and the complexity of the model is restricted. The random field models provide a powerful means of characterizing different texture classes [64], [65].

Local region-based textured image processing methods have achieved remarkable effects. Joint spatial and spatial-frequency representations have been applied to image processing problems since they find their analog in the encoding of visual information in the human visual system. Bovik [66] used complex Gabor filters to compute texture feature from the envelope and phase information extracted from two quadratic components of each channel output. A. K. Jian [59] chose real Gabor filters for analysis. The advantage of these methods is that Gabor filters have orientation selectivity and filter bandwidth properties and the Gabor filters are easy to design. The drawbacks are that they are computational inefficient, requiring a large number of channels to cover the complete frequency plane. Although the Gabor function is an admissible wavelet, it does not result in an orthogonal decomposition [59], the energy distributed in multichannels may correlated. The developments of wavelet theory provide an alternative tool in multichannel filter banks. Wavelet theory has several potential advantages in that wavelet filters cover the complete frequency domain and fast algorithms are available. Several recent studies have successfully applied wavelets to different image processing fields, such as compression, edge detection, texture classification and segmentation, [19].

The different textures in an image usually appear as such to a human observer. There is no mathematical definition that can capture the diversity of textural families. Basically, textures can be classified into two categories: structured and unstructured. The structured texture is produced by the placement of texture primitives according to certain placement rules [67]. Methods that can handle the more structured textures require defining a good set of texture primitives and placement rules. Only regular patterns can be handled in such a way. Unstructured texture is composed of random texture elements. There is no

placement rule for the large class of such textures. Relative to the entire image, textures in the same region have similar characteristics and may exhibit quasi-periodicity.

A large class of natural textures can be modeled as a quasi-periodic pattern and detected by highly concentrated spatial frequencies and orientations. Texture segmentation requires simultaneous measurements in both the spatial and the spatial/frequency domains. Filters with smaller bandwidths in the spatial/frequency domain can make finer distinctions among different textures. Spatial/frequency representation, which preserves both global and local information, is adequate for quasi-periodic signals. The criteria of a good segmentation are: (1) that regions should be homogeneous, (2) that regions should be simple and without many interior holes, (3) that adjacent regions should be significantly different, and (4) that boundaries should be smooth, accurate and well defined [68].

7.2 The Method

In general, texture segmentation consists of two phases: feature extraction and segmentation. Features for texture representation are of crucial importance in accomplishing segmentation [69]. A classification algorithm employing wavelet packets to obtain an energy feature vector of moderate length was proposed in [55] for texture images with large sizes. In this study, a subimage of size 128 x 128 was extracted from the original sample texture with a fixed 43 pixels overlapping subimage, from which 64 sub-samples were evenly extracted. Even though the classification method achieves a high correction rate, the drawback is that a fairly large window size is required for extraction of the feature vector components, and hence only texture images of fairly moderate size can be classified. In fact, the smallest window size deemed suitable is 16 x 16. The algorithm presented here employs a window size only large enough to capture the quasi-periodicity of the texture image. Since only

wavelet decomposition is performed, the proposed algorithm is less compositionally expensive compared with the algorithm in [55]. The energy of an image $f(x, y)$ with n x n pixels is defined as:

$$E_n = \sum_{i}^{n} \sum_{j}^{n} f(i, j)^2 \tag{165}$$

In orthogonal wavelet decomposition the image energy is preserved. We assume the energy distributed on different channels as a distinguishing feature for texture classification and segmentation. In our study the tree structures are defined in two ways: standard tree structure and quadtree tree structure.

The traditional pyramid-type wavelet decomposes an image into four subimages (low frequency subimage, horizontal orientation subimage, vertical orientation subimage and diagonal orientation subimages) by applying wavelet transform filters h and g. Figure 44(a) shows the decomposition in two scales, $s = 2$. The energy of a given image is the sum of the energy of the four subimages in scale 1.

$$E^0 = E_L^1 + E_H^1 + E_V^1 + E_D^1 \tag{166}$$

where the superscription is for the wavelet decomposition scale and the subscripts represent the subimages.

In the standard wavelet tree structure, the low-frequency subimage is decongested recursively in m times. The standard decomposition tree structure up to scale m is shown in Figure 44(b). The energy of the image distributed on the different channels can be written as:

$$
\begin{aligned}
E^0 &= E_L^1 + E_H^1 + E_V^1 + E_D^1 \\
&= E_L^2 + E_H^2 + E_V^2 + E_D^2 + E_H^1 + E_V^1 + E_D^1 \\
&= E_L^m + E_H^m + E_H^{m-1} + \cdots + E_H^1 + E_V^m + E_V^{m-1} + \cdots + E_V^1 + E_D^m + E_D^{m-1} + \cdots + E_D^1 \\
&= E_L^m + \sum_p^m E_H^p + \sum_p^m E_V^p + \sum_p^m E_D^p
\end{aligned}
$$

In the quadtree structure, each subimage is decomposed into four subimages. Figure 44(c) shows the quadtree structure decomposed for three scales ($m = 3$). The sum of the energy on a leaf note is equal to the energy of image $f(x, y)$. We use e_i to represent the energy of leaf i, $1 \leq i \leq n$. For $m = 2$, there are 16 leaves, ($n = 16$). When $m = 3$, there are 64 leaves, ($n = 64$). The general form for quadtree decomposition energy representation is:

$$E^0 = \sum_{i=1}^{4^m} e_i \tag{167}$$

where m is the wavelet transform decomposition scale and i represents the leaf on scale m.

We use the wavelet transform energy vector as the feature of the image. The wavelet transform energy vector is defined as follows. An image block of size of $2k$ x $2k$, which represents the pixel $p(i, j)$ in the image, is decomposed, and the energy distributed on each channel is computed. The energy on each channel constructs a wavelet transform energy vector. The block size is selected as a power of 2 since the wavelet transform decomposed the image into four subimages. The pixel $p(i, j)$ cannot be in the center of the block since the block size is an even number. We consider the $p(i, j)$ as $(k - 1, k - 1)$ of the block. To

overcome the boundary problem, we mirror the boundary of the image of size k in order to make sure that the center of the block will cover the entire image when the block moves from the upper left corner to the lower right corner. Each pixel in the image is represented by a wavelet energy vector. We use $E(i,j)$ to represent the energy vector of pixel $p(i,j)$.

The difference in texture may not be sufficient to distinguish if the block size is too small and the wavelet transform has adequate texture data in the decomposition. If the block is too large, it may lead to incorrect segmentation when the block overlays the texture boundaries. In our experiment, we used a block of size 16 x 16 to construct the energy vectors in a standard tree structure with 13 components in each energy vector. We also used blocks of size 32 x 32 to construct energy vectors with scale $m = 2$ resulting in 16 components in the energy vector. See Figure 45(a),(b) and (c).

$$E^a(i,j) = \{e_1, e_2, \cdots, e_{13}\};$$

$$E^b(i,j) = \{e_1, e_2, \cdots, e_{16}\};$$

$$E^c(i,j) = \{e_1, e_2, \cdots, e_{64}\};$$

The formations of energy vectors will not affect the segmentation results whether the energy vector is formed from left to right of the last wavelet decomposition scale or from right to left.

7.2.1 The Wavelet Selected

The computational structures of the discrete wavelet transform (DWT) are shown in Figure 44 (showing decomposition stages only in our proposed algorithm). The general

structure and computational framework of the DWT are similar to those of subband systems. The main difference lies in the filter design where the wavelet filters are required to be regular [19]. In our experiment, we have selected Daubechies (N=4) orthogonal wavelet (coefficients are listed in Table 1) for texture segmentation, and $N = 10$ nearly symmetric orthogonal wavelet bases introduced in [24] for rotated texture segmentation (coefficients of the lowpass filter are listed in Table 5). The highpass filter, $G(\omega)$, is obtained by frequency-shifting $H(\omega)$ by:

$$g(n) = (-1)^n h(n), \qquad \text{or,}$$

$$G(\omega) = H(\omega + \pi)$$

These filters have flat in-band frequency response and sharp off-band attenuation (see figure 46). In image processing, symmetric/antisymmetric wavelets are preferred, because they can operate on both sides of a pixel evenly. However, a compactly supported orthogonal wavelet cannot by symmetric or antisymmetric [13]. The wavelet we selected have rapid decay speed and are almost symmetric. The orthogonality of the wavelet basis is very important in our experiment, since we use the energy distributed in each channel as the feature of the texture.

7.2.2 The Clustering Method

This classifier is based on the assumption that each pattern class p_k is representable by a prototype pattern C_k (class center). The minimum distance classifier is assigned a pattern x of unknown classification to the class P_k, if the distance D_k between x and C_k is minimum among all possible class prototypes. The Euclidean distance D_k is defined by:

$$D_k = \|x - C_k\| = \sqrt{\sum_{j=0}^{N-1} (x - j - C_{K,j})^2} \tag{168}$$

The true class centers C_K are also unknown in real practice. In our segmentation method, the center of each class C_K is estimated by using the mean of the training samples for each class.

7.2.3 Proposed Algorithm

The segmentation algorithm works as follows:

1. A subimage of appropriate size is extracted from the texture image. This subimage is then decomposed onto orthogonal channels.

2. A feature vector is constructed from the energy values on these channels.

3. In the training phase, a group of samples from each texture image is presented and the average feature vector computed. This average vector is used as the centroid for that particular texture in the feature space.

4. In the classification phase, let $D(\cdot, \cdot)$ be a distance measure and $D(E_j, E_{c_i})$ the distance between energy vectors E_j and centroid E_{c_i}. The pixel at the center of the block is assigned to the class c_i for which $D(E_j, E_{c_i})$ is minimum.

7.3 Experimental Results and Discussion

Structured and unstructured textures were selected from the Brodatz Album [70]. Each test image was 128 x 128 and contained four different textures (see Figure 47[a]). Three test images were used in our segmentation algorithm. The first 50 energy vectors E^a in training texture for D_6, D_{21}, D_{29}, and D_{76} are shown in Figure 48(a), (b), (c), and(d). We used the average l_1 norm of the training samples as the center of the texture.

A majority filter was used as a post-processor to remove the small isolated points and to smooth out the contours of the segment image. In the majority filtering, the window size selected was $(2p+1)$ x $(2p+1)$, $p \in Z$. The label of the central pixel in the window was replaced by the label of the majority of the pixels in the window.

- Training Phase

  1. 100 16 x 16 block size subimages were selected from original texture, then decomposed using wavelet transform. A completed decomposition created energy vector E^a or E^b.
  2. Centers of each training texture were computed using average l_1-norm.

- Segmenting Phase

  1. Test images were decomposed the same as above.
  2. Every energy vector was compared with each cluster center E_{c_i}, and get the minimum distance d_l.
  3. The cluster l for the smallest d_l was assigned to the unknown texture.

- Post-processing using a Majority filter with window size (5 x 5) was applied to achieve better segmentation.

The segmentation results shown in Figure 49(a),(b), and(c) are our experiments results. The result obtained by using the majority filter post-processor are shown in figure 49(d),(e), and (f).

7.4 Image Rotation Segmentation

7.4.1 The Rotated Image

In this section, the important problem of rotation texture identification is addressed. The importance of rotation invariant texture identification was first recognized by Kashyap el al. [71]. Since then, few researchers have paid attention to the rotation invariant texture identification problem. The rotation invariance can be achieved through rotation invariant features, or through an appropriate transformation of the power spectral density of the Gauss Markov Random Field (GMRF) model of the texture [72]. The rotation of texture changes the subband distribution values. If the classifier or the model that is used is able to capture the trend of changes in the channels, the texture still can be classified properly. In texture segmentation, if a correct feature of texture can indicate the trend of the changes when the texture is rotated, this texture can be segmented correctly.

To demonstrate the segmentation algorithm performance for rotated texture, we create two test set images:

Set 1: Texture D57 with texture D77 in 0, 30, and 60 degrees, as shown in Figure 50(a), (e), and (j). Set 2: texture D6 with texture D21 in 0, 30, and 60 degrees as shown in figure 52(a), (e), and (j), respectively. Each one is in the image size of 128 x 128.

7.4.2 Experimental Results and Discussion

In Set 1, D57 is an unstructured texture, and D77 is an structured texture. In this experiment, we selected the block size of 32 x 32 and used a quadtree structure for $m = 2$ to construct the texture feature. The segmentation results are displayed in black and white regions to distinguish the different categories. Figure 50(b),(f) and (j) shows the segmentation results for the $D57/D77$ pair with different angles shown in Figure 50(a),

(e) and (i), respectively. The results of applying majority filtering for (b), (f), and (j) are shown in (c), (g), and (k), respectively. The segmentation errors for each rotation angle are shown in Figure 50(d),(h), and (l), respectively. In Set 2, a block subimage of size 32 x 32 and a standard tree structure are used to construct the energy feature. 100 samples are used to compute the centroid of the test image. After applying wavelet lowpass filter four times, most the energy of the texture is concentrated at the first channel and is much greater than the energy distributed in channel 2 to channel 13. We omit the first channel to show more details in channel 2 to channel 13. The energy vector for D6 is shown in Figure 51(a). Figure 51(b), (c) and (d) show the energy vectors for D21 at 0, 30, and 60 degrees, respectively. It can be observed that the energy distributed in channels 2 - 13 in (b), (c), and (d) are changed but compared with the energy vector in (a), there are still significant difference. The variation of extracted features in the rotated texture is not wide.

Figure 52(b),(f) and (j) shows the segmentation results for the $D6/D21$ pair with different angles shown in Figure 52(a), (e) and (i), respectively. The results of applying majority filtering for (b), (f), and (j) are shown in (c), (g), and (k), respectively. The segmentation errors for each rotation angle are shown in Figure 52(d),(h), and (l), respectively.

Wavelet energy representations computed from the standard tree structure and quadtree structure are sufficient for texture segmentation, but finer segmenting may be achieved by additional space information. Higher scale decomposition breaks down the distribution channel into more detail channels. A more efficient classifier, such as the Mahalanobis Distance, or Bayes Distance, improves segmenting. In this study, analyzing functions were constructed by computing the tensor product of the one-dimensional function, which exhibited sensitivity along the principal axes. Finer segmentation may be achieved by using

a two-dimensional nonseparable analyzing function, which has more orientation selectivity.

7.5 Summary

In this chapter, we presented a multiresolution energy distribution-based texture segmentation technique that used orthogonal wavelet transform to characterize the features of textures. The energy of the image was preserved. Different tree structures were used to construct the energy vector. The center of the cluster was assigned by computing the average of selected training samples. The effect of size on border performance was discussed. A clustering algorithm was then used to identify the texture categories. A simple procedure for inclusion of spatial information in the clustering process was used to reduce misclassification. A number of test images, with structured and unstructured images, and rotated textures, were used to demonstrate the algorithm is success of the algorithm in texture segmentation.

For rotated image segmentation, we only selected 0, 30, and 60 degree rotated images. It is assumed that a good texture rotation invariant identification can handle arbitrary orientation. Further study of rotated texture orientation in wavelet energy representation can be based on capturing the trend of feature changes in the channels caused by rotation, providing effective information on dependence among the subbands. Accompanied by a Hidden Markov Model (HMM), the segmentation can be more flexible at different rotations and the error rate decreased.

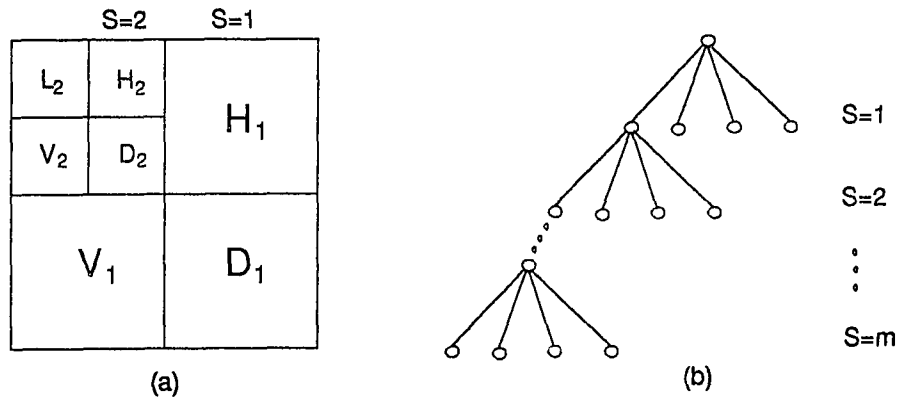


Figure 44. (a) Decomposition in 2-D and (b) Standard tree structure

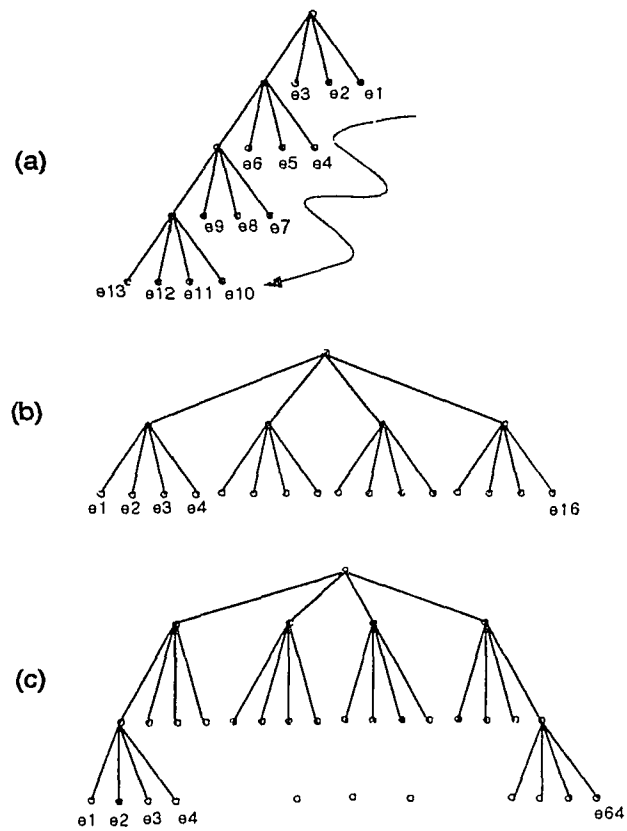


Figure 45. (a) 13 components energy vector.
(b) 16 components energy vector.
(c) 64 components energy vector

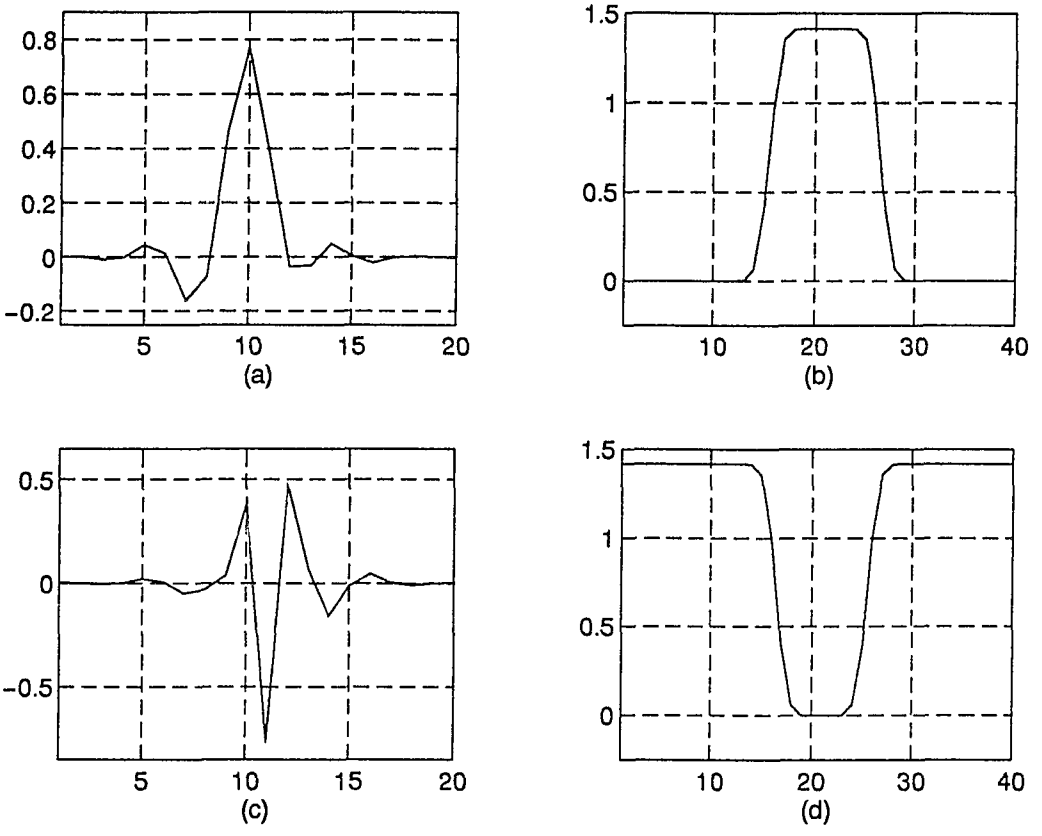


Figure 46. Daubechies Orthogonal Basis, nearly symmetric
(a) Lowpass filter in Time domain
(b) Lowpass filter in Frequency domain
(c) Bandpass filter in Time domain
(d) Bandpass filter in Frequency domain

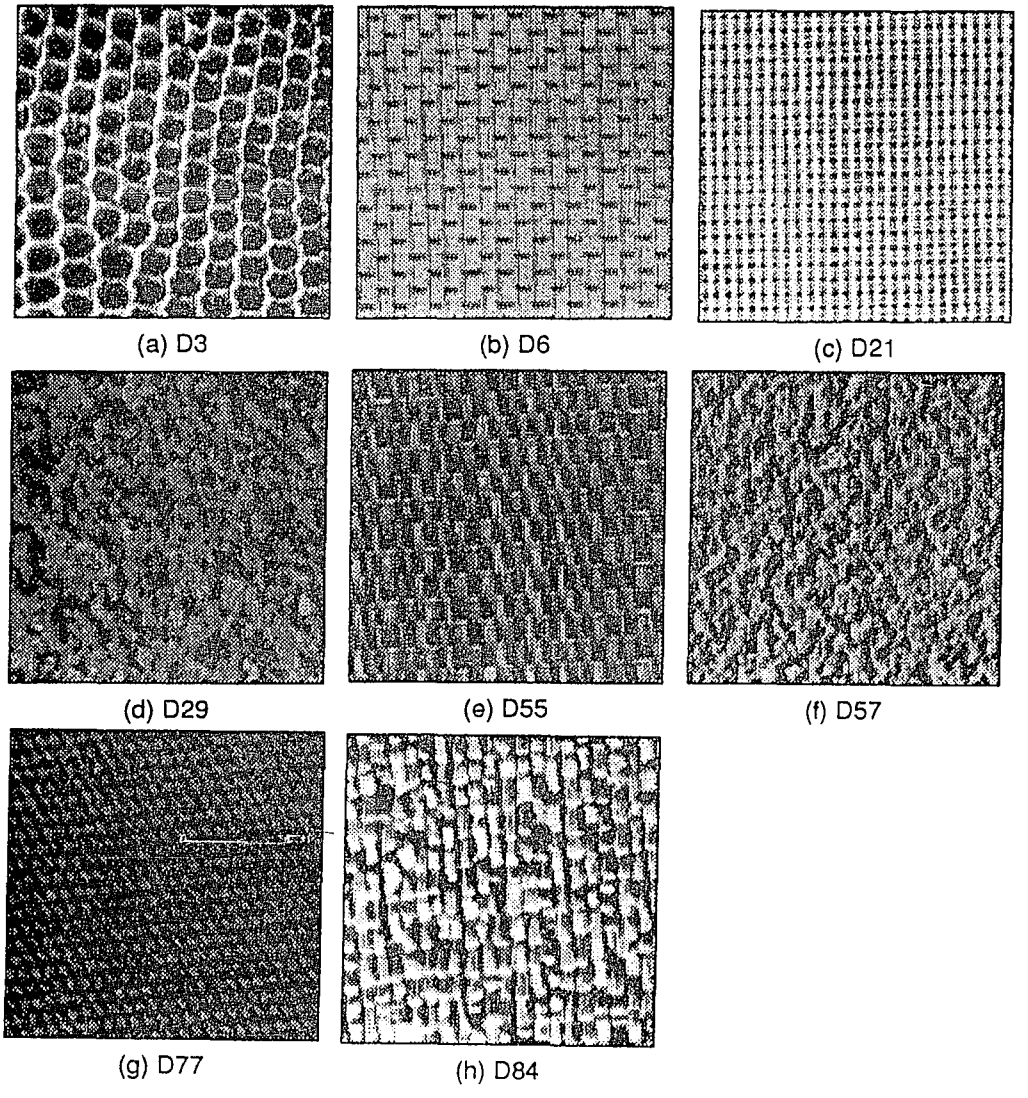


Figure 47. Textures selected from P. Brodatz's Album[6].

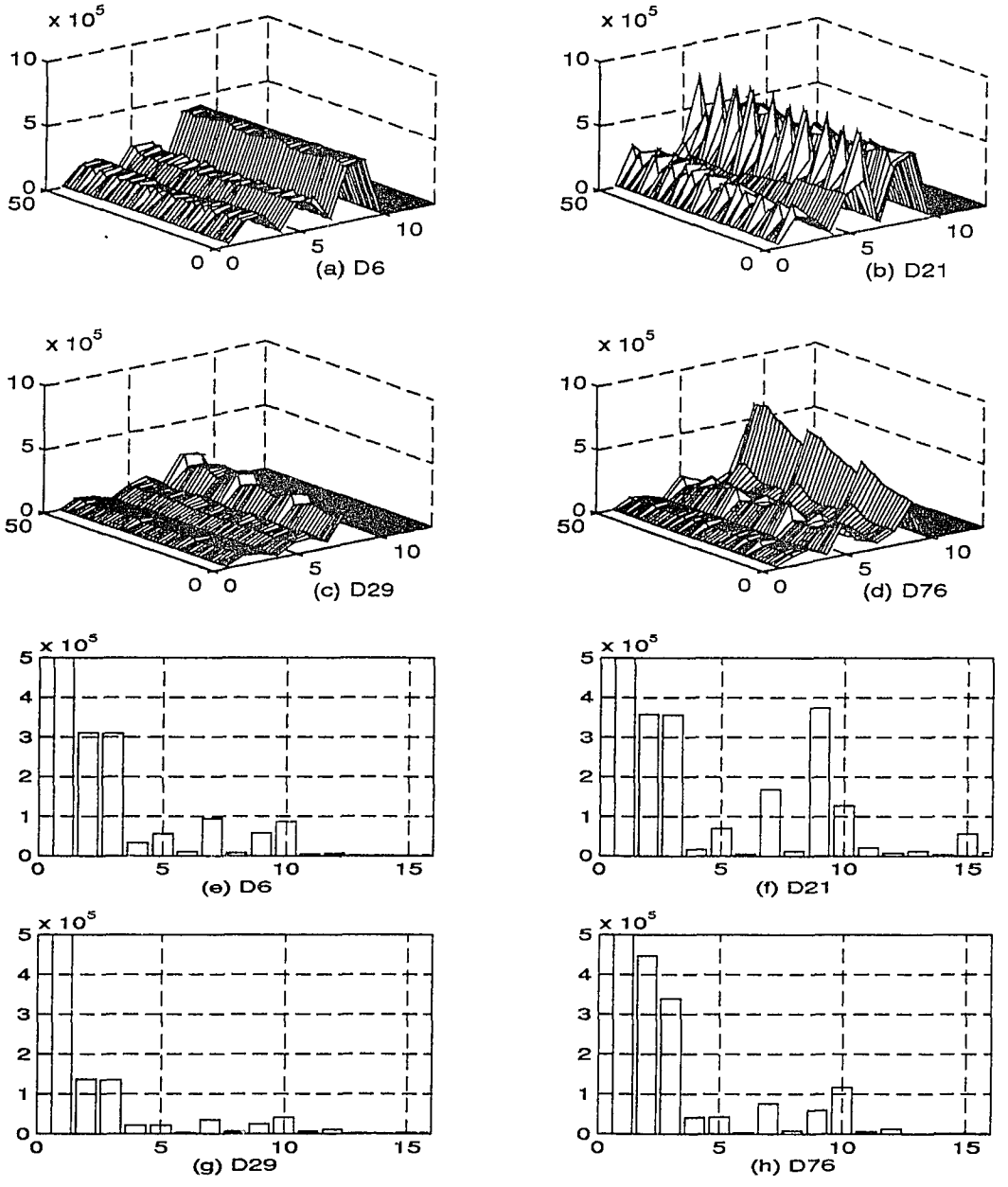


Figure 48.
(a) 50 energy vectors with 13 components for D6
(b) 50 energy vectors with 13 components for D21
(c) 50 energy vectors with 13 components for D29
(d) 50 energy vectors with 13 components for D76
(e) Energy vector with 16 components for D6
(f) Energy vector with 16 components for D21
(g) Energy vector with 16 components for D29
(h) Energy vector with 16 components for D76

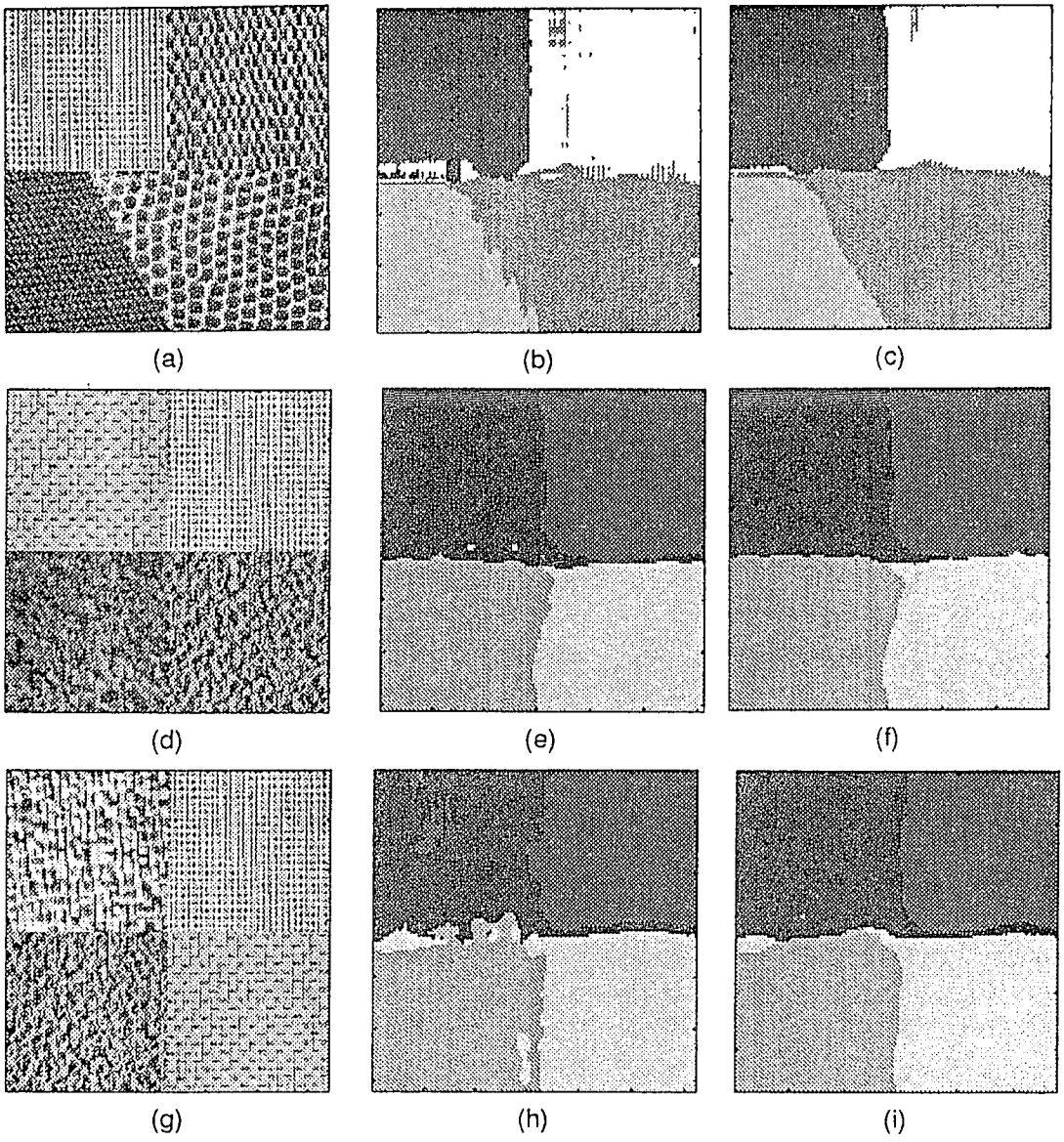


Figure 49. Three test images and the segmentation results (a),(d) and (g) are original test images. (b), (e) and (h) are segmentation results. (c), (f) and (i) are results after post processing.

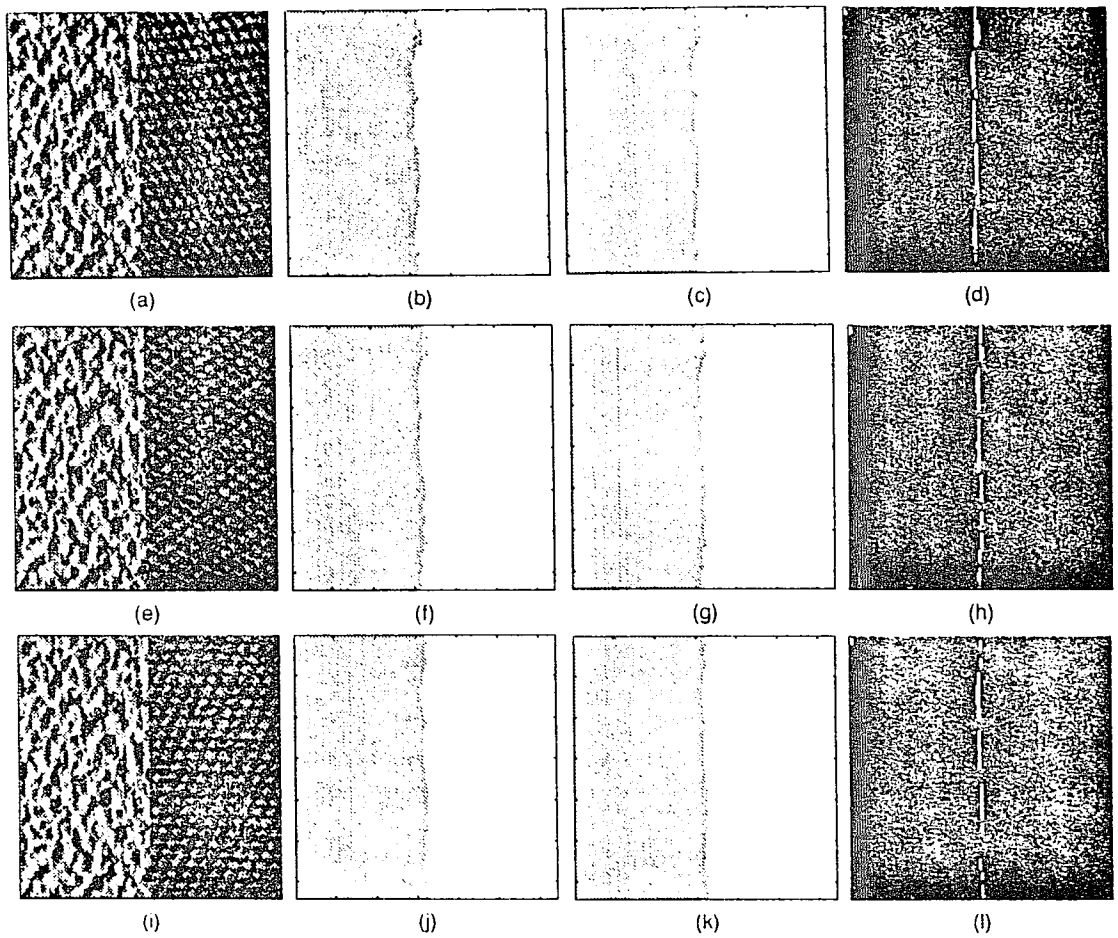


Figure 50. Rotated test image set 1 and segmentation results.
(a) is D57 and D77 with 0 Deg, (b) is segmentation result
(c) is the result after post processing, (d) is the segmentation errors. (e) is D57 and D77 with 30 Deg. (i) is D57 and D77 with 60 Deg. (f) and (j) are segmentation results, (g) and (k) are the results after post processing, (h) and (l) are segmentation errors, respectively.

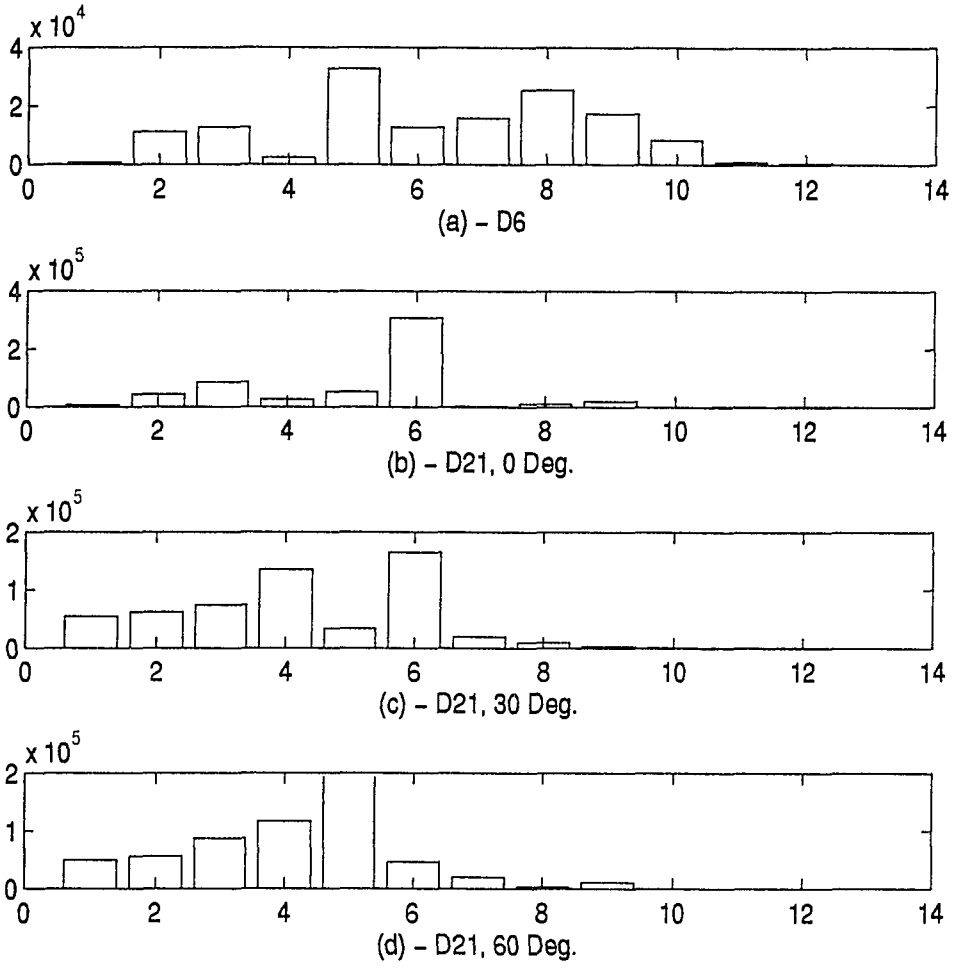


Figure 51. Energy vector for D6, and D21 with 0, 30 and 60 Deg. Only e2 to e13 are shown.

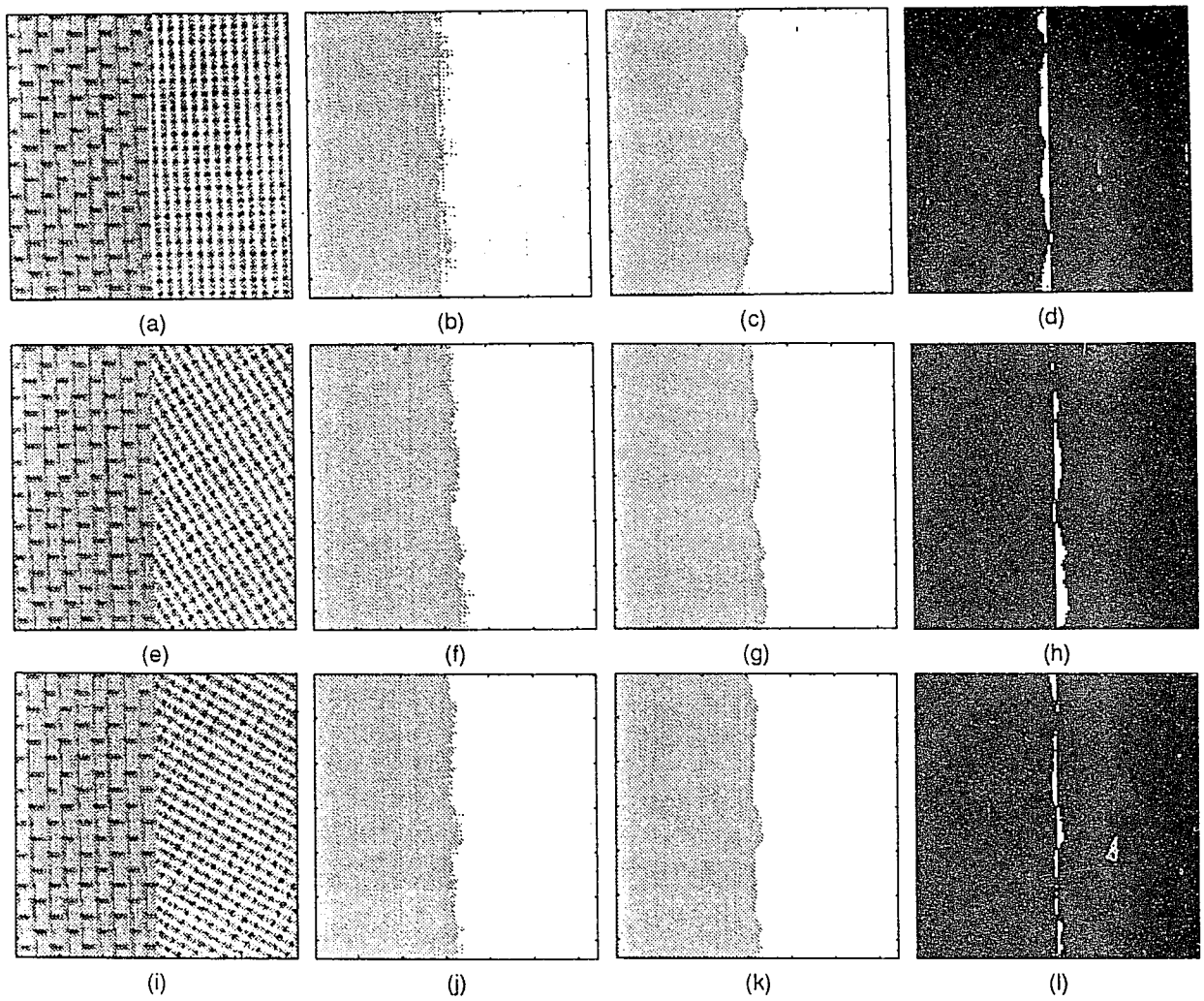


Figure 52. Rotated test image set 2 and segmentation results. (a) is D6 and D21 with 0 Deg, (b) is segmentation result (c) is the result after post processing, (d) is the segmentation errors. (e) is D6 and D21 with 30 Deg. (i) is D6 and D21 with 60 Deg. (f) and (j) are segmentation results, (g) and (k) are the results after post processing, (h) and (i) are segmentation errors, respectively.

Table 4. Orthogonal wavelet coefficients used in rotated texture segmentation

n	h(n)	n	h(n)
1	0.00077	11	0.38383
2	0.00096	12	-0.03554
3	-0.00864	13	-0.03199
4	-0.00147	14	0.04999
5	0.04593	15	0.00577
6	0.01161	16	-0.02036
7	-0.15949	17	-0.00081
8	-0.07088	18	0.00459
9	0.47169	19	0.00006
10	0.76951	20	-0.00046

Table 5. Error Rate for rotated texture segmentation

Texture	Deg.	Before Post Processing		After Post Processing	
		Num.	%	Num.	%
D57/D77	0	153	0.93%	143	0.87%
	30	133	0.81%	128	0.78%
	60	110	0.67%	100	0.61%
D6/D21	0	448	2.73%	326	1.99%
	30	302	1.84%	278	1.70%
	60	214	1.31%	164	1.00%

References

[1] S. Mallat, "A Theory for Multiresolution Signal Decomposition: the Wavelet Representation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, pp. 674–693, 1989.

[2] P. J. Burt and E. H. Adelson, "The Laplacian Pyramid as a Compact Image Code," *IEEE Trans. Commun.*, vol. Com-31, pp. 532–540, Apr. 1983.

[3] J. Crowley, "A Representation for Visual Informationf," Tech. Rep. CMU-RI-TR-82-7, Robotic IJ. Inst. Carnegie Mellon Univ., 1987.

[4] S. Mallat, "Multifrequency Channel Decomposition of Images and Wavelet Model," *IEEE Trans. Acoust. Speech, Signal Processing .*, vol. 37, pp. 2091–2110, Dec. 1989.

[5] S. Mallat, "Multiresolution Approximations and Wavelet Orthonormal Bases of $l^2(r)$," *Trans. of American Mathematical Society*, vol. 315, pp. 69–87, Sep. 1989.

[6] E. O. Brigham, *The Fast Fourier Transform.* Englewood Cliffs, New Jersey: Prentice-Hall,Inc., 1974.

[7] R. N. Bracewell, *The Fourier Transform and Its Applications.* New York: McGraw-Hall Book Company, 1986.

[8] D. F. Elliott and K. R. Rao, *Fast Transfroms Algorithms, Analyses, Applications.* New York: Academic Press, 1982.

[9] D. Gabor, "Theory of Communication," *J. Inst. Elec. Eng.*, vol. 93, no. 3, pp. 429–457, 1946.

[10] A. Grossman and J. Morlet, "Decomposition of Hardy Functions Into Square Integrable Wavelets of Constant Shape," *Siam J. Math. Anal.*, vol. 15, pp. 723–736, July 1984.

[11] Y. Meyer, "Ondelettes et fonctions splines," tech. rep., Śeminaire EDP,École Polytechnique,Paris, 1986.

[12] J. Stromberg, "A Modified Haar System and Higher Order Spline Systems," in *Conf. in Harmonic Analysis*, vol. II, (Honor of Antoni Zygmund), pp. 475–493, 1982.

[13] I. Daubechies, "Orthonornal Bases of Compactly Supported Wavelets," *Comm. Pure Appl.Math.*, vol. XLI, no. 41, pp. 909–996, 1988.

[14] G. Battle, "A Block Spin Construction of Wavelets, part i lemarie functions," *Comm. Math. Phys.*, vol. 110, pp. 601–615, 1987.

[15] P. Lemarié, "Une Nouvelle Base," *J.de Math. Pures et Appl.*, vol. 67, pp. 227–236, 1988.

[16] A. C. Ingrid Daubechies and J. Feauveau, "Biorthogonal Bases of Compactly Supported Wavelets," *Comm. Pure Appl.Math.*, vol. XLV, no. 41, pp. 485–560, 1992.

[17] Y. Meyer, *Wavelets and Operators.* Paris: Cambridge University Press, 1992.

[18] C. K. Chui, *Wavelets: A Tutorial in Theory and Application.* Academic Press Inc., 1992.

[19] A. Riour and M. Vertelli, "Wavelets and Signal Processing," *IEEE Trans. Signal Processing*, pp. 14–38, 1991.

[20] S. Mallat and S. Zhong, "Characterization of Signals from Multiscale Edges," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 14, pp. 710–732, July 1992.

[21] M. Unser and M. Eden, "Multiresolution Feature Extraction and Selection for Texture Segmentation," *IEEE. Trans. Pattern Anal. Machine Intell.*, vol. 11, July 1989.

[22] J. R. Willians and K. Amaratunga, "Introduction to Wavelets in Engineering," *preprint.*

[23] Y. Meyer, "Principe D'incertitude, Bases Hilertiennes et Algerbres D'operateurs," *Seminaire Bourbake.*, vol. 662, 1985.

[24] I. Daubechies, *Ten Lectures on Wavelets.* Philadelphia: Siam, 1992.

[25] O. Rioul and P. Duhamel, "Fast Algorithms for Discrete and Continuous Wavelet Transforms," *IEEE Trans. Info. Theory*, vol. 38, March 1992.

[26] D.Marr and E.Hildreth, "Theory of Edge Detection," *Proc. Royal Soc. London*, vol. 207, pp. 187 - 217, 1980.

[27] A.Rosenfeld and M. Thurston, "Edge and Curve Detection for Visual Scene Analysis," *IEEE Trans. COmput.*, vol. C-20, pp. 572 - 569, 1971.

[28] A. Witkin, "Scale Space Filtering," in *Int. Joint Conf. Artifical Intell*, 1983.

[29] J. Canny, "A Computational Approach to Edge Detection," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 8, pp. 679–698, 1986.

[30] S. Mallat, "Zero-crossing of a Wavelet Transform," *IEEE Trans. Info. Theory*, vol. 37, pp. 1019–1032, July 1991.

[31] S. kadambe and G. Boudreaux-Bartels, "Application of the Wavelet Transform for Pitch Detection of Speech Signals," *IEEE Trans. on Information Theory*, vol. 38, no. 2, pp. 917–924, 1992.

[32] W. Hess, *Pitch Determination of Speech Signals: Algorithms and Devices.* Berlin: Springer-Verlag, 1983.

[33] J. B. H. Jeanty and J. Gil, "On the Components and Specifications of Personal Computer Based Image Processing System," *ACTA Stereol*, vol. 6/2, pp. 233–246, 1987.

[34] C. Garbay, "Image Structure Representation and Processing: A Discussion of Some Segmentation Method in Cytology," *IEEE Trans. Pattern Anal. and Machine Intell.*, vol. PAMI-8, no. 2, pp. 140–146, 1986.

[35] P. F. J. Barba, J. Hemrick and J. Gil, "The Use of Local Entropy Measure in Edge Detection for Cytological Image Analysis," *Microscopy*, vol. 156, pp. 125–134, Oct. 1989.

[36] L.Herts and R.Schafer, "Multilevel Thresholding Using Edge Matching," *Comp. Vis. Graph. Ima. Process.*, vol. 44, pp. 279–295, 1988.

[37] Y. L. J. Barba and J. Gil, "Cell Image Contour Extraction by Multithresholding," in *6th International Symposium on Diagnostic Quantization Pathology*, (Basel/Swetyerland), 1990.

[38] Y. L. J. Barba and J. Gil, "Edge Detection in Cytology Using Morphological Filters," *Digital Image Processing Applications*, vol. 1075, pp. 311–318, Jan. 1989.

[39] Y. Li and J. Barba, "Segmentation and Texture Representation with Vector Quantizers," (San Diego), SPIE, 1990.

[40] A. W. J.M. Lester, A. Williams and J. Brenna, "Two Graph Searching Technique for Boundary Finding in White Blood Cell Image," *Comp. Biol. Med.*, vol. 8, pp. pp. 293–308, 1978.

[41] G. Ashkar and J. Modestero, "The Contour Extraction Problem with Biomedical Applications," *Comp. Graphics, Image Proc.*, vol. 7, pp. 331–355, 1978.

[42] G. B. C. Anrabay, J.M. Chassery, "An Iteration Region Growing Process for Cell Image Segmentation Based on Local Color Similarity and Global Shape Criteria," *Anal. and Quant. Cytology and Hist.*, vol. 8, no. 1, pp. 25–34, 1986.

[43] A. W. H.M. Roafat, "A Texture Information Directed Region Growing Algorithm for Image Segmentation and Region Classification," *Computer Vision and Image Processing*, vol. 43, pp. 1–21, 1988.

[44] A. Jain, *Fundamentals of Digital Image Processing.* Prentice Hall, 1989.

[45] J.-G. Leu, "Image Contrast Enhancement Based on the Intensities of Edge Pixels," *CVGIP: Graphical Models and Image Processing*, vol. 54, no. 6, pp. 497–506, 1992.

[46] A. Beghdadi and A. L. Negrate, "Contrast Enhancement Techbnique Based on Local Detection of Edges," *Computer Vision, Graphics, and Image Processing*, vol. 46, pp. 162–174, 1989.

[47] M. J.Parkinen, G.Cohen and N.Andreasen, "Segmentation of MR Brain Images," in *Int. Conf. Medicine Biology*, pp. 71–72, Proc. IEEE, 1991.

[48] X. Wu, "Adaptive Split-and-Merge Segmentation Based on Piecewise Least-Square Approximation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 15, pp. 808–815, 1993.

[49] H.-S. Wu and J. Barba, "An Algorithm for Noisy Cell Contour Extraction via Area Merging," *J. of Imaging Science and Technology*, vol. 38, pp. 604–607, Nov. 1994.

[50] D. M. H. J. Jian Lu and J. B. Weaver, "Contrast Enhancement of Medical Images Using Multiscale Edge Representation," *Optical Engineering.*

[51] J. Sklansky, "Image Segmentation and Feature Extraction," *IEEE Trans. on System and Man Cybern*, vol. 8, pp. 237–247, 1978.

[52] B. Julesz, "A Theory of Preattentive Texture Discrimination Based on First-order Statistics of Textons," *Biol. Cybern.*, vol. 44, pp. 131–138, 1981.

[53] J. Daugman, "Spatial Visual Channels in the Fourier Plane," *Vision Research*, vol. 24, pp. 891–910, 1984.

[54] A. S. J. Beck and R.Ivry, "Spatial Frequency Channels and Perceptual Grouping in Texture Segregation," *Comput. Vision, Graphics, Image Processing*, vol. 37, pp. 299–325, 1987.

[55] A. Laine and J. Fan, "Texture Classification by Wavelet Packet Signature," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 15, pp. 1186–1197, Nov. 1993.

[56] C. Bouman and B. Liu, "Multiple Resolution Segmentation of Textured Images," *IEEE trans. Pattern Anal. Machine Intell.*, vol. 3, pp. 99–113, Feb. 1991.

[57] T. Chang and C. Kuo, "Texture Analysis and Classification with Tree-Structured Wavelet Transform," *IEEE Trans. Image Processing*, vol. 2, pp. 429–441, Oct. 1993.

[58] J. Liu and Y. Yang, "Multiresolution Color Image Segmentation," *IEEE. Trans. Pattern Anal. Machine Intell.*, vol. 16, Jan. 1994.

[59] A. K.Jain and F. Farrakhnia, "Unsupervised Texture Segmentation Using Gabor Filters," *IEEE Trans. Pattern Recog.*, vol. 24, pp. 1167–1185, May 1991.

[60] C. Won and H. Derin, "Segmentation of Noisy Textured Images using Simulated Annealing," *ICASSP 87*, pp. 14.4.1–14.4.4, 1987.

[61] H. Derin and C. Won, "A Parallel Image Segmentation Algorithm using Relaxation with Varying Neighborhoods and its Mapping to Array Processors," *Comput. Vision Graphics Image Processing*, vol. 40, pp. 54–78, Oct. 1987.

[62] J. Hsiano and A. Sawchuk, "Supervised Textured Image Segmentation using Feature Smoothing and Probabilistic Relaxation Techniques," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, pp. 1279 – 1292, 1989.

[63] T. Reed and H. Wechesler, "Segmentation of Textured Images and Gestalt Organization using Spatial/Spatial-Frequency Representations," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, pp. 1–12, 1990.

[64] G. Cross and A. Jain, "Markov Random Field Texture Models," *IEEE. Trans. Pattern Anal. and Machine Intell.*, vol. 5, pp. 25–39, Jan. 1983.

[65] R. Chellappa and S. Challerjee, "Classification of Texture using Gaussian Markov Random Fields," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. Assp-33, pp. 959–964, Aug. 1985.

[66] A. Bovik, "Analysis of Multichannel Narrowband Filters for Image Texture Segemtnation," *IEEE Trans. Signal Processing*, vol. 38, pp. 2025–2043, Sept. 1991.

[67] R. F.M. Vilnrotter and K. Price, "Structural Analysis of Natural Texture," *IEEE. Trans. Pattern Anal. and Machine Intell.*, vol. 8, pp. 76–89, 1986.

[68] R. Haralick and L. Shapiro, "Survey - Image Segmentation Techniques," *Comput. Vision, Graphics, and Image Processing*, vol. 29, pp. 100–132, 1985.

[69] M. J.M.H.DU Buf and M.Spann, "Texture Feature Performance for Image Segmentation," *Pattern Recognition*, vol. 23, pp. 291–309, 1990.

[70] P. Brodatz, "Textures: A Photographic Album for Artists and Designers." Dover, New York, 1966.

[71] R. Kashyap and A.Khotanzad, "A Model-based Method for Rotation Invariant Texture Classification," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-8, pp. 472–481, July 1986.

[72] Z. F. F.S. Cohen and M. Patel, "Classification of Rotation and Sealed Textured Images using Gauss Markov Random Field Models," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, pp. 192–202, Feb. 1991.