A Wavelet-Based Technique for Reducing Noise in Audio Signals

by

K. Allen Comer

Thesis submitted to the faculty of the

Virginia Polytechnic Institute and State University

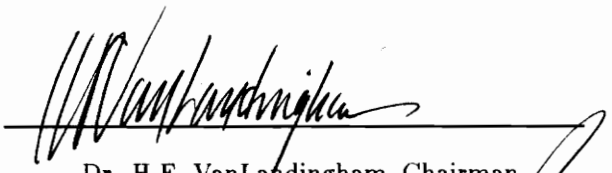
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

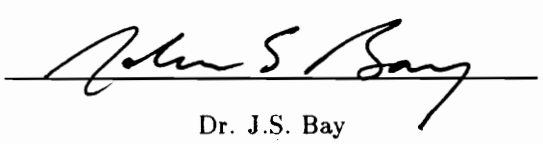
in

Electrical Engineering - Signals and Systems

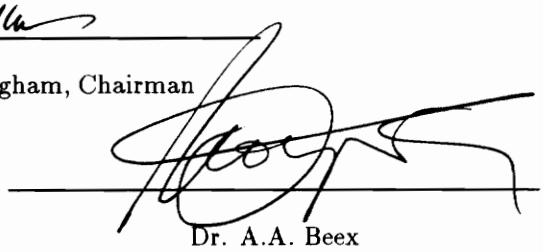
APPROVED:



Dr. H.F. VanLandingham, Chairman



Dr. J.S. Bay



Dr. A.A. Beex

April, 1996

Blacksburg, Virginia

A Wavelet-Based Technique for Reducing Noise in Audio Signals

by

K. Allen Comer

Committee Chairman: Dr. H.F. VanLandingham

Bradley Department of Electrical Engineering

(ABSTRACT)

Wavelets have received considerable attention in recent general signal processing, image processing, and pattern recognition literature, as a new method of signal analysis. This marks a transition in wavelet study from theoretical investigation to application-driven research. In this paper, wavelets and wavelet transformations are presented in a context intended to be appropriate as a first exposure to the engineer. The wavelet transform, more specifically the discrete wavelet transform, and its relationship to multiresolution analysis is then explored in a framework familiar to those versed in multirate digital signal processing concepts. Elements of the perspective offered by wavelet analysis, in contrast to the features of more conventional Fourier techniques, are examined. General procedures for wavelet-based signal processing applications are discussed and the specific application of reducing noise in audio signals examined. Within the context of this application, considerations unique to wavelet analysis are revealed and trade-offs analyzed. Finally, the results obtained from implementing the noise reduction system are presented and extensions to the technique proposed.

ACKNOWLEDGEMENTS

I would like to acknowledge my advisor, Dr. H. F. VanLandingham, and the remaining members of my committee, Dr. J. S. Bay and Dr. A. A. Beex. The knowledge gained in courses taken under these instructors was essential throughout all phases of the research. As much of the wavelet literature currently available originates from the mathematical community, a strong fundamental framework was necessary. I thank my committee members for providing me with much of the material needed to construct this framework. Additional thanks are extended to these individuals for renewing my appreciation of, and interest in, the engineering discipline. Acknowledgement is also given to Mr. Farooq-e-Azam, for both his convincing argument regarding the use of LaTeX as well as general assistance and good conversation. Finally, thanks are offered to members of the statistics department of Stanford University, responsible for the WaveLab software which was integral to completing this project.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

*

Chapter 1

INTRODUCTION

Wavelets and signal analysis based on transformation under the wavelet basis constitute a new paradigm for signal processing referred to as *analysis-by-scale*. In this technique, signals are viewed at varying degrees of resolution (or scales) allowing simultaneous observation of gross and local behavior. This type of decomposition is mirrored in the human auditory system where sensitivity to frequency varies roughly by octave intervals over a significant portion of the audio spectrum. The human vision system may also operate on a multiscale principle with visual receptors functioning as edge detectors of different sizes and varying spatial acuity. The analogy between these natural processes and wavelet analysis has created a good deal of interest from researchers in many areas.

The relationship between wavelets and wavelet transforms is analogous to that of sinusoids and the Fourier transform. Furthering the analogy, wavelet transforms accommodate both the continuous-time and discrete-time cases with efficient algorithms existing for the latter situation. Unlike Fourier transforms, wavelet transforms segment the frequency spectrum into octave intervals, leading to a multiresolution perspective. In Chapter 2, wavelets and wavelet transforms are introduced. Following this introduction and a short example, the construction of wavelets through recursive dilation is examined. As a result of the rigid mathematical constraints that govern this construction, an important relationship between wavelets and their associated 'mother' function exists. This relationship is defined in Chapter 2 and further investigated in Chapter 3.

The primary focus of Chapter 3 is the wavelet basis and the discrete wavelet transform (DWT). Specifically, the wavelet basis and the DWT are discussed within the context of Mallat's multiresolution analysis (MRA) ideas. This directly leads to a perspective on the DWT presented in terms of Quadrature Mirror Filter (QMF) pairs which are borrowed from multirate digital signal processing. The QMF model, in turn, leads to the efficient FWT algorithm for computing the DWT. Chapter 3 concludes by introducing the wavelet packet library (of which the wavelet basis is a subset) and the Wavelet Packet Transform (WPT).

Chapter 4 seeks to highlight the differences between wavelet and Fourier analysis. Critical to appreciating the capabilities and limitations of wavelet analysis is an understanding

of the elements that distinguish these two types of signal decomposition. A preliminary related discussion concludes this chapter.

Wavelets have only recently received attention from engineering and other application-oriented disciplines. In Chapter 5, several current uses of wavelets are examined. The variety of wavelets and the number of bases found in the WPT introduced in Chapter 3 provide a good deal of flexibility for signal analysis. However, with such a large degree of freedom offered by the analysis tool, techniques for selecting the best-suited wavelet and wavelet packet basis are necessary. Chapter 5 provides guidelines for choosing an analyzing wavelet and describes the 'best basis' algorithm within the greater context of a wavelet-based procedure for noise reduction and data compression - 'Wavelet Packet Shrinkage.' A signal analysis paradigm with both Fourier and wavelet components, local trigonometric decomposition, is introduced at the end of Chapter 5.

The procedures of Chapter 5 are young and many variations of the basic idea have been proposed. The result is a host of adjustable parameters that should be considered when implementing the shrinkage algorithms. The effects of varying these parameters, as well as general performance issues, are discussed in Chapter 6 which also presents the results of applying shrinkage as an audio signal noise reduction techique.

Wavelet theory is developed in this paper within the context of function decomposition by orthogonal transformation. This perspective is necessary in order to fully appreciate the signal analysis possibilities offered by wavelets. However, it may be beneficial to first examine wavelets strictly qualitatively.

Perhaps most critical in beginning the study of wavelets is disassociating conventional ideas regarding signal analysis from the wavelet-based paradigm. Most engineers, scientists, and mathematicians develop notions of the utility and mechanics of signal analysis based on exposure to the well-known Fourier transform. As a result, the concepts of transform and distinct frequencies become intimately linked. While information regarding a signal's frequency content can be obtained, the notion of scale is the distinguishing element in wavelet analysis. Wavelet transforms are calculated by level, each tier revealing the amount

of signal detail lost by looking through a window with half the resolution capability as that at the previous level.

The fundamentals of a compression algorithm are already apparent. Moving from one level of the wavelet transform to another offering less detail, a more compact representation is possible if the important information remains. This application highlights the nature of wavelet signal processing.

Traversing the levels of a wavelet transform, a sense of what is critical is obtained. This multiresolution representation is even more useful since there is a common factor relating the amount of detail across levels. Specifically, a given level in the wavelet transform affords a factor of two greater or less resolution than seen in the adjacent levels. This type of decomposition is realized by essentially applying a bank of filter pairs with each stage in the cascade consisting of symmetric high- and low-pass components. Repeatedly filtering the low-pass outputs with these symmetric-response filters strips increasing amounts of detail. As the input signal at each level (the previous low-pass output) has half the bandwidth as the input at the preceding stage, the mirror image filters are seen to satisfy a constant-Q (center frequency / bandwidth) condition.

The multiresolution respresentation is obtained by expressing the signal as a linear combination of a particular wavelet and shifted, as well as, dilated variations thereof. As mentioned earlier, wavelets are discussed in greater detail in Chapters 2 and 3. Let it suffice, at present, to describe wavelets as functions which exist over a finite interval. Dilating and shifting the wavelets yields a family of functions capable of completely representing a signal of interest in multiresolution fashion.

Briefly, and somewhat coarsely, these are the fundamentals of wavelet analysis. The important features revealed in the above discussion include: (1) the idea of *analysis-by-scale* and (2) the finite interval over which wavelets exist. It is these properties that provide the greatest distinction of wavelet-based analysis when compared to Fourier-based techniques. Given this qualitative background, it is now necessary to turn toward a more detailed and mathematically-oriented examination before considering potential applications.

Chapter 2

WAVELETS AND THE WAVELET TRANSFORM

2.1 Wavelets and the Wavelet Basis

Wavelets are so-named as they: (1) integrate to zero ('waver' about the x-axis) and (2) are well-localized in time (wavelet) [1]. A function that satisfies these two conditions can be used as the fundamental element in constructing an orthogonal basis for the space of square-integrable, or energy, functions, $\mathbf{L}^2(\Re)$. This basis consists of a mother function (also known as a scaling function or 'mother' wavelet) and translated/dilated versions of the mother function (the entities to which the term wavelets generally refers). The Haar wavelet and translated/dilated versions of the Daubechies (D4), Coiflet (C3), and Symmlet (S8) wavelets are shown in Figure 2.1.

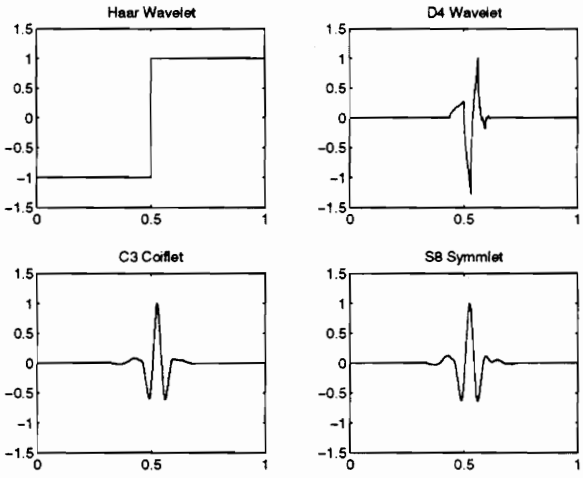


Figure 2.1: Wavelet Examples

The shorthand convention used in referencing wavelets is derived from the number of vanishing moments in a given wavelet. A function $f(x)$ with α vanishing moments satisfies:

$$\int_{-\infty}^{\infty} f(x)x^p dx = 0 \quad \text{for } 0 \leq p \leq \alpha. \tag{2.1}$$

It is useful to note, at this point, that the wavelet basis incorporates short, high-frequency, and long, low-frequency, analysis functions (Figure 2.1) as a result of the translations and dilations on the mother wavelet.

2.2 The Continuous Wavelet Transform (CWT)

The majority of practical applications, involving wavelets, are performed using the discrete wavelet transform (DWT). However, for completeness, the CWT of a signal, s(t), is given below:

$$S(b,a) =< s(t), \psi_{b,a}(t) >, \tag{2.2}$$

or,

$$S(b,a) = |a|^{-\frac{1}{2}} \int_{-\infty}^{\infty} s(t)\psi(\frac{t-b}{a})dt \tag{2.3}$$

where,

$$\psi_{b,a}(t) = |a|^{-\frac{1}{2}}\psi(\frac{t-b}{a}). \tag{2.4}$$

The corresponding inverse transform is defined by the following expression:

$$s(t) = C_{\psi}^{-1} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (< s(t), \psi_{b,a}(t) >)(\psi_{b,a}(t))\frac{da\,db}{a^2} \tag{2.5}$$

where,

$$C_{\psi}^{-1} = 2\pi \int_{-\infty}^{\infty} |\hat{\psi}(\xi)|^2 |\xi|^{-1} d\xi \tag{2.6}$$

and $\hat{\psi}(\xi)$ indicates the Fourier transform of $\psi(\xi)$:

$$\hat{\psi}(\xi) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \psi(\xi)e^{-j\omega\xi}d\xi. \tag{2.7}$$

The wavelet, by virtue of its localization in time, behaves somewhat like the window functions used in conventional Fourier analysis. Therefore, it is not surprising that the CWT and ICWT exhibit similarity to their short-term Fourier counterparts. The short-term Fourier transform (STFT) is given by:

$$S_{\tau}(\omega) =< s, g_{\tau,\omega} > \tag{2.8}$$

or equivalently,

$$S_{\tau}(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} s(t)g(t-\tau)e^{-j\omega t}dt \tag{2.9}$$

where g(t) is the window function, τ is the delay parameter, and,

$$g_{\tau,\omega} = g(t-\tau)e^{j\omega t}. \tag{2.10}$$

where g(t) is the window function, τ is the delay parameter, and,

$$g_{\tau,\omega} = g(t-\tau)e^{j\omega t}. \tag{2.10}$$

The inverse STFT is computed from:

$$s(t)g(t-\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_\tau(\omega)e^{j\omega t}dt. \tag{2.11}$$

2.3 The Discrete Wavelet Transform (DWT) and an Example

The decomposition of a discrete-time signal, $s(x)$, resulting from wavelet transformation, is given by:

$$s(x) = \sum_{k=-\infty}^{\infty} c_{\phi,k}\phi(x-k) + \sum_{j=0}^{\infty}\sum_{k=-\infty}^{\infty} d_{j,k}\psi(2^j x - k). \tag{2.12}$$

where $\phi(x)$ denotes the mother function and $\psi(2^j x - k)$ represents the wavelets.

Before examining the DWT and its development in more detail, it is illustrative to consider a particular wavelet (Haar), its associated basis, and the wavelet transform of a data vector, $\mathbf{y}$, of length 2^N. If the indices on the summations are such that $0 \leq j \leq N$ and $0 \leq k \leq 2^j$, the family of Haar functions constitutes a basis for the interval [0,1). The mother function, $\phi(x)$, for the Haar basis (Figure 2.2) is defined to be:

$$\phi(x) = 1 \ (0 \leq x < 1). \tag{2.13}$$

while the untranslated/undilated Haar wavelet, $\psi(x)$, is represented by:

$$\psi(x) = \begin{cases} -1 & 0 \leq x < \frac{1}{2} \\ 1 & \frac{1}{2} \leq x < 1 \end{cases} \tag{2.14}$$

Figure 2.2 depicts the Haar mother function and several translated and dilated Haar wavelets.

The family covers the interval with successively smaller-scale Haar functions at each level of the decomposition (each increment of the index j). However, the Haar basis is

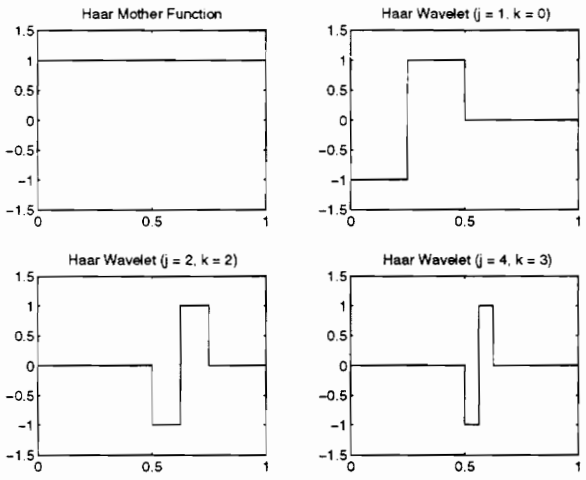


Figure 2.2: Mother Function, $\phi(x)$, and Haar Wavelets, $\psi(2^j x - k)$

useful only after mapping $\mathbf{y}$ onto $[0, 1)$ through:

$$f(x) = \sum_{k=0}^{2^N-1} y_k \quad [k2^{-N} \leq x < (k+1)2^{-N}]. \tag{2.15}$$

Given the size of the data vector (length 2^N), the mapping of (2.15), and the Haar basis, the limits on the summation indices of (2.12) are adjusted. The sum with respect to j now ranges from 0 to $(N-1)$ which scales the wavelets at each level of decomposition. The sum with respect to k now extends from 0 to $2^j - 1$ which corresponds to the translations needed to cover $[0, 1)$ at the j^{th} level of the decomposition. The wavelet decomposition of $f(x)$ is now:

$$f(x) = c_{\phi,0}\phi(x) + \sum_{j=0}^{N-1} \sum_{k=0}^{2^j-1} d_{j,k}\psi_{jk}(x). \tag{2.16}$$

where,

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

In the following example [1], consider the specific vector $\mathbf{y}$,

$$\mathbf{y} = [\ 1 \quad 0 \quad -3 \quad 2 \quad 1 \quad 0 \quad 1 \quad 2\ ]$$

Before continuing, it should be stated that the role of the mother function, $\phi(x)$, is discussed in 2.5 and Chapter 3. However, concerning the mother function, it is not unreasonable

that there is only one $\phi(x)$ term (untranslated and undilated) in the expansion of (2.16). Referring to Figure 2.2 it is seen that the mother function exists, with unity magnitude, over its domain ($[0,1)$) which is also the domain of $f(x)$.

The matrix representation of (2.16) is:

$$\begin{bmatrix} 1 \\ 0 \\ -3 \\ 2 \\ 1 \\ 0 \\ 1 \\ 2 \end{bmatrix} \begin{bmatrix} 1 & -1 & -1 & 0 & -1 & 0 & 0 & 0 \\ 1 & -1 & -1 & 0 & 1 & 0 & 0 & 0 \\ 1 & -1 & 1 & 0 & 0 & -1 & 0 & 0 \\ 1 & -1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & -1 & 0 & 0 & -1 & 0 \\ 1 & 1 & 0 & -1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 & -1 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_{\phi,0} \\ d_{0,0} \\ d_{1,0} \\ d_{1,1} \\ d_{2,0} \\ d_{2,1} \\ d_{2,2} \\ d_{2,3} \end{bmatrix} \tag{2.17}$$

where the dilated and translated wavelets appear in the matrix columns (orthogonality of the basis is easily verified). Solving this system of equations leads to the result:

$$\begin{bmatrix} c_{\phi,0} \\ d_{0,0} \\ d_{1,0} \\ d_{1,1} \\ d_{2,0} \\ d_{2,1} \\ d_{2,2} \\ d_{2,3} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ \frac{5}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \tag{2.18}$$

which, when expanded, yields,

$$f(x) = \frac{1}{2}\phi(x) + \frac{1}{2}\psi_{00}(x) - \frac{1}{2}\psi_{10}(x) + \frac{1}{2}\psi_{11}(x) - \frac{1}{2}\psi_{20}(x) + \frac{5}{2}\psi_{21}(x) - \frac{1}{2}\psi_{22}(x) + \frac{1}{2}\psi_{23}(x). \tag{2.19}$$

As a check, the data point $x = \frac{5}{16}$ is examined. Using the mapping of (2.15), this point corresponds to the third segment $[\frac{1}{4}, \frac{3}{8})$ of the interval $[0, 1)$. Therefore, the associated data point in $\mathbf{y}$ is $y(3) = f(\frac{5}{16}) = -3$.

$$\begin{aligned} f(\tfrac{5}{16}) &= \tfrac{1}{2}\phi(\tfrac{5}{16}) + \tfrac{1}{2}\psi_{00}(\tfrac{5}{16}) - \tfrac{1}{2}\psi_{10}(\tfrac{5}{16}) + \tfrac{1}{2}\psi_{11}(\tfrac{5}{16}) - \tfrac{1}{2}\psi_{20}(\tfrac{5}{16}) + \tfrac{5}{2}\psi_{21}(\tfrac{5}{16}) - \tfrac{1}{2}\psi_{22}(\tfrac{5}{16}) \\ &\quad + \tfrac{1}{2}\psi_{(\frac{5}{16})} \\ &= \tfrac{1}{2}(1) + \tfrac{1}{2}(-1) - \tfrac{1}{2}(1) + \tfrac{1}{2}(0) - \tfrac{1}{2}(0) + \tfrac{5}{2}(-1) - \tfrac{1}{2}(0) + \tfrac{1}{2}(0) \\ f(\tfrac{5}{16}) &= -3. \end{aligned} \tag{2.20}$$

The above example is atypical in that the data vector is small, lending a simple solution to the matrix equation. This is generally not the case and the Fast Wavelet Transform (FWT), analogous to the FFT, exists to handle these longer data sequences. The example is also unusual in that the Haar basis is used in the transform. Generally speaking, the Haar basis is not practical due to its piecewise-constant character which makes it a poor approximator for continuous functions. The problem is similar to that seen in Fourier analysis when high-frequency signals are considered - the number of terms needed for acceptable accuracy becomes excessive. The construction of better wavelets is the subject of the following section.

2.4 The Construction of Wavelets

There are several well-known and widely used wavelets for which closed-form expressions exist, to include the Morlet and Meyer functions. The properties of these wavelets are addressed in many references, such as [2]. Recently, however, attention has been given to the construction of wavelets for which there is no closed-form description. An examination of this process, which yields the wavelet used in the application described in Chapter 6, also reveals much about wavelets in general. The procedure assumes the form of an iterative

system based on the 'dilation equation,' which is given below:

$$\phi_j(x) = \sum_{k=0}^{W-1} a_k \phi_{j-1}(2x - k). \tag{2.21}$$

A qualitative foundation for this equation is presented in 3.1 where the wavelet basis is examined. From that perspective, it is clearly seen why the dilation equation assumes the form of (2.21). The expression is given at this point out of its necessity in discussing the mathematical properties that characterize wavelets.

The iteration of (2.21) always begins with the unit box and is continued until some limiting shape is reached. Furthermore, the number of coefficients, W, is always chosen to be even. The $W = 2$ case corresponds to the mother function found in the Haar basis. The desirable properties of compact support (existence of the wavelet over a finite interval only) and piecewise-continuous behavior, coupled with the orthogonality-to-translations requirement poses a formidable mathematical challenge. It was seriously doubted that such a function could be constructed until the 1988 finding of Ingrid Daubechies (see Figure 2.1, D4) [2]. The discovery of (2.21) was critical to the construction of compactly-supported, piecewise-continuous mother functions and the bases derived from them. A similar equation is used to construct the actual wavelets and incorporates the same coefficients used in (2.21):

$$\psi(x) = \sum_{k=0}^{W-1} (-1)^{k+1} a_{k+W-1} \phi(2x - k) \tag{2.22}$$

Comparing (2.21) and (2.22) reveals that the coefficients of (2.21) are included in the summation of (2.22) in reverse-order and with alternating signs. This particular permutation yields wavelets that are orthogonal to the mother function. Given the expressions of (2.21) and (2.22), all that remains in constructing wavelets is to find the coefficients, a_k.

Note: The term 'wavelet coefficients' is used to refer to the $c_{\phi,k}$ and $d_{j,k}$ entities and should not be confused with the coefficients used to construct the mother function and the wavelets (the a_k terms in (2.21) and (2.22)).

2.5 Computation of the Coefficients for Wavelet Construction

It is not enough to simply solve (2.21) by any means, there are several conditions that must be satisfied and are done so by appropriate selection of the coefficients, a_k. These conditions are discussed in 2.5.1, 2.5.2, and 2.5.3. The underlying ideas behind, and implications of, the following developments are important to understanding the FWT and the relationship of the wavelet transform to QMF pairs. Before proceeding, it has been found [3] that, in the case of more than six coefficients, (2.21) is solvable by numerical methods only. A discussion of the techniques for solving (2.21) is too far removed from the intent of this paper. Those interested may refer to [3] or [4]. Much of the following is taken from Newland, [3], more mathematically-oriented discussions can be found in [2], [4], and [5].

2.5.1 Condition 1 - conservation of area

The ability of wavelets to provide long, low-frequency, as well as short, high-frequency, analysis windows is established by the conservation of area condition. Mathematically, this constraint is expressed by integrating both sides of (2.21):

$$\int_{-\infty}^{\infty} \phi(x)dx = \sum_k a_k \int_{-\infty}^{\infty} \phi(2x-k)dx. \tag{2.23}$$

Substituting $y = 2x - k$ on the right hand side of the equation yields:

$$\int_{-\infty}^{\infty} \phi(x)dx = \sum_k a_k \frac{1}{2} \int_{-\infty}^{\infty} \phi(y)dy \tag{2.24}$$

from which:

$$\sum_k a_k = 2. \tag{2.25}$$

Additionally, since the mother functions are always constructed beginning with the unit box (which **is** the mother function in the Haar case), it follows that:

$$\int_{-\infty}^{\infty} \phi(x)dx = 1. \tag{2.26}$$

2.5.2 Condition 2 - accuracy

Strang [4] and [5] has introduced two additional restrictions on the coefficients which relate to: (1) the accuracy of the transform (often called Condition A) and (2) the orthogonality of the basis functions (often called Condition O). This section addresses the accuracy condition. Accuracy, in the following context, refers to a wavelet expansion's ability to provide a faithful representation of $f(x)$ when $f(x)$ can be expressed by a sum of terms: $1, x, x^2, x^3, \ldots, x^n$. Strang [4] and [5] has determined that the Fourier transform of the scaling function must be periodically zero in order for it to accurately represent a function exhibiting polynomial behavior. Given the Fourier transform of the scaling function:

$$P(\xi) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi(x) e^{-j\xi x} dx, \tag{2.27}$$

substituting from (2.21):

$$P(\xi) = \frac{1}{2\pi} \sum_k a_k \int_{-\infty}^{\infty} \phi(2x - k) e^{-j\xi x} dx, \tag{2.28}$$

and changing the variable of integration, again using $y = 2x - k$, yields:

$$\begin{aligned} P(\xi) &= \tfrac{1}{2} \textstyle\sum_k a_k e^{-j(\frac{\xi}{2})k} \tfrac{1}{2\pi} \int_{-\infty}^{\infty} \phi(y) e^{-j(\frac{\xi}{2})y} dy \\ P(\xi) &= \tfrac{1}{2} \textstyle\sum_k a_k e^{-j(\frac{\xi}{2})k} P(\tfrac{\xi}{2}). \end{aligned} \tag{2.29}$$

A more compact expression can be formed using the following definition:

$$p(\frac{\xi}{2}) = \frac{1}{2} \sum_k a_k e^{-j(\frac{\xi}{2})k}. \tag{2.30}$$

Equation (2.29) now becomes:

$$P(\xi) = p(\frac{\xi}{2}) P(\frac{\xi}{2}) \tag{2.31}$$

and using the substitution $\frac{\xi}{2}$ for ξ yields,

$$P(\frac{\xi}{2}) = p(\frac{\xi}{4}) P(\frac{\xi}{4}) \tag{2.32}$$

and finally,

$$P(\xi) = p(\frac{\xi}{2}) p(\frac{\xi}{4}) P(\frac{\xi}{4}). \tag{2.33}$$

This factoring can be continued until the term $P(0)$ is closely approximated on the right-hand side of the equation. This term can be found using (2.27):

$$P(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi(x) dx \tag{2.34}$$

then, using the result of (2.26),

$$P(0) = \frac{1}{2\pi}. \tag{2.35}$$

Now, an even more compact expression for $P(\xi)$ is possible:

$$P(\xi) = \frac{1}{2\pi} \prod_{j=1}^{\infty} p(\frac{\xi}{2^j}) \tag{2.36}$$

where,

$$p(\frac{\xi}{2^j}) = \frac{1}{2} \sum_k a_k e^{-i(\frac{\xi}{2^j})k}. \tag{2.37}$$

Strang showed that the Fourier transform of the scaling function, $P(\xi)$, must have zeros of the highest order when $\xi = 2\pi, 4\pi, 6\pi, \ldots$ - this is the accuracy condition. Consider $\xi = 2\pi$ and (2.36):

$$P(2\pi) = \frac{1}{2\pi} p(\pi) p(\frac{\pi}{2}) p(\frac{\pi}{4}) p(\frac{\pi}{8}) \ldots \tag{2.38}$$

The first of the $p(\cdot)$ factors will be zero to order n if:

$$\frac{d^m p(\xi)}{d\xi^m} = 0 \quad \text{when } \xi = \pi \quad \text{for } m = 0, 1, 2, \ldots, n-1 \tag{2.39}$$

which, from (2.37), is true when:

$$\sum_k a_k (-ik)^m e^{-i\pi k} = 0 \quad \text{for } m = 0, 1, 2, \ldots, n-1 \tag{2.40}$$

or, equivalently, with $e^{-i\pi k} = (-1)^k$:

$$\sum_k (-1)^k k^m a_k = 0 \quad \text{for } m = 0, 1, 2, \ldots, n-1 \tag{2.41}$$

In the event that this is satisfied, $P(2\pi)$ is zero to order n or higher (depending on the other terms). If the case $\xi = 4\pi$ is examined, it is seen that (from (2.36)):

$$P(4\pi) = \frac{1}{2\pi} p(2\pi) p(\pi) p(\frac{\pi}{2}) p(\frac{\pi}{4}) p(\frac{\pi}{8}) \ldots \tag{2.42}$$

where it is noted that $p(\pi)$ is a factor and thus $P(4\pi)$ is also zero to at least order n, in the event that (2.41) is satisfied. It can be shown that the term $p(\pi)$ appears in the factorization of $P(\xi)$ for $\xi = 2\pi, 4\pi, 6\pi, \ldots$. Therefore, (2.41) can be used to determine the coefficients necessary for satisfying the accuracy requirement. It is seen, in (2.41), that there is no bound on the index m as the requirement is for zeros of the highest possible order. The actual number of equations that are extracted from (2.41) depends on: 1) the number of coefficients to be used in the wavelet; 2) the conservation of area condition; and 3) orthogonality constraints.

2.5.3 Condition 3 - orthogonality

It is recognized that the orthogonality condition is not essential and that non-orthogonal wavelet bases are used in special cases. However, the FWT is based on the existence of orthogonality. Under the assumption that the FWT is the desired approach for computing the DWT, the orthogonality condition is considered necessary.

It will be shown in 3.1 that the translates of $\phi(x)$ form a basis over a particular space, V_0. As the function is dilated and translated for a given value of j, the new set of translates $\phi(2^j x - m)$ form a basis for a different space, V_j. Consider the case $j = 1$. In order to ensure that V_1 also has an orthogonal basis of translates, the inner product:

$$< \phi_1(x), \phi_1(x - m) > \tag{2.43}$$

is examined using the first iteration of (2.21),

$$\int_{-\infty}^{\infty} \phi_1(x)\phi_1(x - m)dx = \int_{\infty}^{-\infty} \sum_k a_k \phi_0(2x - k) \sum_v a_v \phi_0(2x - 2m - v)dx. \tag{2.44}$$

or,

$$\int_{-\infty}^{\infty} \phi_1(x)\phi_1(x - m)dx = \sum_k \sum_v a_k a_v \int_{\infty}^{-\infty} \phi_0(2x - k)\phi_0(2x - 2m - v)dx. \tag{2.45}$$

As the $\phi_0(x)$ are orthogonal with respect to translations, the right-hand side integral is zero except when $k = 2m + v$ or $v = k - 2m$. Therefore, the single case in which the right-hand

side integral is non-zero is represented by:

$$\int_{\infty}^{\infty} \phi_1(x)\phi_1(x-m)dx = \sum_k a_k a_{k-2m} \int_{-\infty}^{\infty} \phi_0^2(2x)dx \tag{2.46}$$

Excluding the case where $m = 0$, which corresponds to the inner product, $< \phi_1(x), \phi_1(x) >$, the orthogonality condition is obtained:

$$\sum_k a_k a_{k+2m} = 0 \text{ for all } m \text{ except } m = 0. \tag{2.47}$$

To illustrate this idea, consider the case of six coefficients, a_0 to a_5. The orthogonality requirement is satisfied by solving:

$$\begin{aligned} a_0a_2 + a_1a_3 + a_2a_4 + a_3a_5 &= 0 \quad \text{for } m = 1 \\ a_0a_4 + a_1a_5 &= 0 \quad \text{for } m = 2. \end{aligned} \tag{2.48}$$

Note that $m = 3, 4, \ldots$ does not apply for the case of six coefficients.

2.5.4 Implications of the conditions on wavelet coefficients

In summary, given that the coefficients a_k satisfy the conditions outlined above, the following relationships between the mother function, wavelets, and their respective translates hold:

$$\int_{-\infty}^{\infty} \psi(2^n x)\phi(x-m)dx = 0 \quad n \geq 0 \tag{2.49}$$

and,

$$\int_{-\infty}^{\infty} \psi(x)\psi(2^n x - m)dx = 0 \quad n \geq 0 \text{ subject to } m \neq 0 \text{ when } n = 0. \tag{2.50}$$

The mother function and wavelets also satisfy the following relationships. By setting $m = 0$ in (2.46), the following results:

$$\int_{-\infty}^{\infty} \phi_1^2(x)dx = 2\int_{-\infty}^{\infty} \phi_0^2(2x)dx \tag{2.51}$$

Given that the unit box is always the seed of (2.21), the above can be generalized:

$$\int_{-\infty}^{\infty} \phi^2(x)dx = (2)(\frac{1}{2}) = 1. \tag{2.52}$$

Using (2.22) and the accuracy condition (2.41), it can also be shown that:

$$\int_{-\infty}^{\infty} \psi^2(x)dx = 1. \tag{2.53}$$

Substituting $y = x - k$ into (2.22); keeping $m = 0$; and integrating yields:

$$\int_{-\infty}^{\infty} \psi(x)dx = \sum_k (-1)^k a_k \frac{1}{2} \int_{-\infty}^{\infty} \phi(y)dy. \tag{2.54}$$

Then, using the result of (2.41):

$$\int_{-\infty}^{\infty} \psi(x)dx = 0. \tag{2.55}$$

The conservation of area, the accuracy, and the orthogonality conditions result in 1, $W/2$, and $W/2$ equations to be solved, respectively. In other words, for the W coefficients, there are $W+1$ equations. However, it can be shown that the first of the accuracy conditions ($m = 0$) is redundant as it can be derived from the other two conditions. This is seen by examining (2.25), (2.41), and (2.47) for W coefficients. The conservation of area requirement (2.25) becomes:

$$\sum_{k=0}^{W-1} a_k = 2; \tag{2.56}$$

the accuracy condition (2.41) assumes the form:

$$\sum_{k=0}^{W-1} (-1)^k k^m \, a_k = 0 \quad \text{for } m = 0, 1, 2, \ldots, W/2 - 1; \tag{2.57}$$

and similarly with the orthogonality constraint (2.47):

$$\sum_{k=0}^{W-1} a_k a_{k+2m} = 0 \quad m \neq 0. \tag{2.58}$$

In the final expression regarding the orthogonality criterion, it is seen that only $W/2 - 1$ equations result since $m \neq 0$.

As an aside, the requirement that $\phi(x)$ be orthogonal to its translates is examined by taking the inner product of both sides of (2.21):

$$\int_{-\infty}^{\infty} \phi^2(x)dx = \sum_k a_k^2 \int_{-\infty}^{\infty} \phi^2(2x - k)dx \tag{2.59}$$

Returning to the task of finding the redundancy, the accuracy condition (2.41) is evaluated at $m = 0$:

$$\sum_{k=0}^{W-1} (-1)^k k^{(0)} a_k = a_0 - a_1 + a_2 - a_3 \ldots = 0. \tag{2.62}$$

The expansion can be rewritten:

$$\sum_{k\ even} a_k - \sum_{k\ odd} a_k = 0 \Rightarrow \sum_{k\ even} a_k = \sum_{k\ odd} a_k. \tag{2.63}$$

If the square of both summations are added, the result of (2.61) can be used to obtain:

$$\left(\sum_{k\ even} a_k \right)^2 + \left(\sum_{k\ odd} a_k \right)^2 = 2, \tag{2.64}$$

which is then expanded:

$$\sum_{k=0}^{W-1} a_k^2 + 2 \sum_{m=1}^{W/2-1} \sum_{k=0}^{W-1} a_k a_{k+2m} = 2. \tag{2.65}$$

The redundancy is revealed by applying the result of (2.61):

$$\sum_m \sum_k a_k a_{k+2m} = 0,\ m \neq 0. \tag{2.66}$$

The conditions developed in this section have important implications outside of their ability to yield an orthogonal basis of varying-scale functions. Several of the properties that surfaced in the analyses above have led to alternative perspectives on the wavelet basis. Most importantly, Mallat and Daubechies have used some of these ideas, in conjunction with multiresolution analysis (MRA) principles (3.1), to construct the FWT. The MRA and the FWT are discussed in 3.1 and 3.3, respectively.

Chapter 3

THE WAVELET BASIS AND THE DWT

3.1 Mallat's Multiresolution Analysis (MRA) and the Wavelet Basis

Mallat [6], has shown that the wavelet basis yields a multiresolution analysis of $\mathbf{L}^2(\Re)$. While this is seen clearly in the example considered earlier, a more thorough investigation of Mallat's MRA provides valuable insight into the wavelet basis and its composition. Furthermore, it underscores the recursive element that exists in the construction of the wavelets (2.21) as well as within the bases they form. This recursive property is exploited in deriving the FWT.

The Haar basis and the interval $[0, 1)$ are considered again. Let the spaces spanned by $\phi(x)$ and $\psi(x)$ be denoted V_0 and W_0, respectively. The orthogonality of these two spaces is evident from the matrix equation (2.17) (recall the first and second columns of the matrix correspond to $\phi(x)$ and $\psi(x)$). The sum of these two spaces $V_1 = V_0 \oplus W_0$ yields the space of piecewise-constant functions on the two half-intervals $[0, \frac{1}{2})$ and $[\frac{1}{2}, 1)$ within $[0, 1)$. Another basis for V_1 could be the two functions, $p(x)$ and $q(x)$, pictured in Figure 3.1.

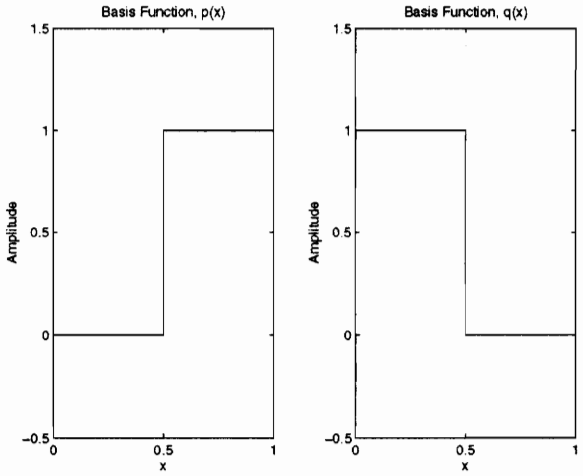


Figure 3.1: A Basis for $V_1 = V_0 \oplus W_0$ where $V_1 = [0, 1)$

Returning again to Figure 2.2 and the expressions for the Haar mother function (2.13) and wavelets (2.14), it is seen that the functions $p(x)$ and $q(x)$, (Figure 3.1), can be con-

structed from $\phi(x)$ and $\psi(x)$ using the following relationships:

$$p(x) = \frac{1}{2}(\phi(x) + \psi(x)) \tag{3.1}$$

and,

$$q(x) = \frac{1}{2}(\phi(x) - \psi(x)) \tag{3.2}$$

An important observation is that $p(x) = \phi(2x)$ and $q(x) = \phi(2x-1)$, (compare Figures 2.2 and 3.1). Therefore, (3.1) and (3.2) are equivalently expressed by:

$$\phi(2x) = \frac{1}{2}(\phi(x) + \psi(x)) \tag{3.3}$$

and,

$$\phi(2x-1) = \frac{1}{2}(\phi(x) - \psi(x)). \tag{3.4}$$

Solving for $\psi(x)$ in (3.3) and substituting into (3.4) yields:

$$\phi(x) = \phi(2x) + \phi(2x-1). \tag{3.5}$$

If the pattern is extended to the spaces V_j and W_j, both of dimension 2^j, the general definitions below can be established:

$V_j \equiv$ the space having the translates $\phi(2^j x - k)$ as a basis

$W_j \equiv$ the space having the translates $\psi(2^j x - k)$ as a basis

Thus, V_2 is spanned by: $\phi(4x)$, $\phi(4x-1)$, $\phi(4x-2)$, and $\phi(4x-3)$ with each term occupying a quarter-interval. Continuing with a development similar to that for V_0, W_0, and V_1 - V_2 can be segmented into V_1 and W_1, $V_2 = V_1 \oplus W_1$. This pattern continues and the wavelet decomposition yields a set of nested subspaces:

$$V_j \subset V_{j+1} \tag{3.6}$$

or alternatively,

$$V_{j+1} = V_j \oplus W_j = V_0 \oplus W_0 \oplus \cdots \oplus W_j \tag{3.7}$$

Thus, the wavelet basis is multiresolutional with smaller, higher-frequency windows at each level of the transform.

It is now desired to extend the ideas above to the entire line $\Re$, [4]. In this generalization, the index k is now unrestricted and the subspaces V_j and W_j are infinite-dimensional. The space of square-integrable functions, $\mathbf{L}^2(\Re)$ becomes the closure of the translates of $\phi(x)$ and $\psi(x)$ when restrictions on the indices are relaxed. From (3.7) , it is seen that one basis of $\mathbf{L}^2(\Re)$ is $\phi(x-k)$ and $\psi(2^j x - k)$ with $j \geq 0$ and $k \in \mathbf{Z}$ (note that translations of the mother function are now required, as opposed to the $[0,1)$ case). If the j index is also allowed to take negative values ($-\infty < j < \infty$), then it can be shown that another basis for $\mathbf{L}^2(\Re)$ contains $\psi_{j,k}(x)\ \forall\ j,k \in \mathbf{Z}$. Consider the Haar basis and the event of negative indices, particularly $j = -1$. In this instance, the Haar mother function and its translates, $\phi(2^{-1}x - k)$, are constant over intervals of length 2 and are thus orthogonal to the $\psi_{j,k}(x)$ functions. Therefore, the decompositions $V_{j+1} = V_j \oplus W_j$ remain valid. A more formal treatment of the ideas listed in this paragraph are beyond the scope of this paper and those interested may refer to [6] for a more detailed presentation. However, the informal development above provides some foundation for the following list of properties applicable to the sequence of closed subspaces, V_j, $-\infty < j < \infty$:

$$
\begin{aligned}
&\text{(a) } V_j \subset V_{j+1}\ ,\ \bigcap V_j = \{0\},\ \text{and } \bigcup V_j \text{ is dense in } \mathbf{L}^2(\Re);\\
&\text{(b) } f(x) \in V_j \text{ if and only if } f(2x) \in V_{j+1};\\
&\text{(c) } V_0 \text{ has an orthogonal basis of translates } \phi(x-k),\ k \in \mathbf{Z};\\
&\text{(d) } V_j \text{ is spanned by } \phi(2^j x - k);\\
&\text{(e) } W_j \text{ is the orthogonal complement of } V_j \text{ in } V_{j+1};\\
&\text{(f) } W_j \text{ is spanned by translates which are the wavelets.}
\end{aligned}
\tag{3.8}
$$

These are the elements of Mallat's MRA approach which not only characterizes the wavelet decomposition of the function space, but provides guidelines for constructing the wavelets themselves. In fact, many of these properties are implied by the dilation equation, (2.21), presented in 2.4. The components of MRA also lead to the QMF pair model perspective for examining the DWT.

3.2 QMF Pairs and the DWT

The $(j+1)$ level of the wavelet transform corresponds to the function space, $V_{j+1} = V_j \oplus W_j$. From (3.8), the basis functions for V_{j+1} are translates of the mother function ($V_j \Leftrightarrow \phi(2^j x - k)$) and wavelets ($W_j \Leftrightarrow \psi(2^j x - k)$). The multipliers for the basis functions are then calculated using inner products taken from both sides of (2.12). Denoting the multipliers by $c_{\phi,k}$ (with the ϕ implying no dilation - to be clarified later) and $d_{j,k}$ for the mother function and wavelets, respectively, and recalling orthogonality relationships among the mother function, wavelets, and their translates leads to:

$$\int_{-\infty}^{\infty} f(x)\phi(x-k)dx = c_{\phi,k}\int_{-\infty}^{\infty} \phi^2(x-k)dx. \tag{3.9}$$

Substituting $y = x - k$,

$$\int_{-\infty}^{\infty} f(x)\phi(x-k)dx = c_{\phi,k}\int_{-\infty}^{\infty} \phi^2(y)dy. \tag{3.10}$$

Then, using the result of (2.52),

$$c_{\phi,k} = \int_{-\infty}^{\infty} f(x)\phi(x-k)dx. \tag{3.11}$$

Similarly, for the wavelet multipliers, $d_{j,k}$,

$$\int_{-\infty}^{\infty} f(x)\psi(2^j x-k)dx = d_{j,k}\int_{-\infty}^{\infty} \psi^2(2^j x-k)dx. \tag{3.12}$$

Substituting $y = (2^j x - k)$,

$$d_{j,k} = \frac{\int_{-\infty}^{\infty} f(x)\psi(2^j x-k)dx}{\frac{1}{2^j}\int_{-\infty}^{\infty} \psi^2(x)dx} \tag{3.13}$$

Then, using the result of (2.53),

$$d_{j,k} = 2^j \int_{-\infty}^{\infty} f(x)\psi(2^j k-x)dx. \tag{3.14}$$

These equations establish a basis for each level of the transform based on the property $V_{j+1} = V_j \oplus W_j$; however, there is unnecessary redundancy inherent in this series of basis

functions due to the nested character of the subspaces. Recalling (3.7), it is seen that the only set of $c_{\phi,k}$ multipliers required are those at level $j = 0$ of the transform, sometimes called the ϕ level as it is the only level which directly includes weighting from the mother function. This level is also referred to as two separate levels, the ϕ and $j = 0$ levels, to distinguish the mother function term from the wavelet term and is the convention used in this paper. Additionally, it will be seen that no translations of the mother function are required. The remaining coefficients necessary to complete the basis are the $d_{j,k}$ terms.

Examining the wavelet decomposition of a signal $s(n)$, having length 2^N, it is seen that the number of elements in the j^{th} level of the decomposition is 2^j ($j = 0, 1, \ldots, N-1$). For the ϕ level there is just one term as with the $j = 0$ level. The decomposition, by level, of the sequence whose DWT was calculated in (2.12) is shown below in Figure 3.2.

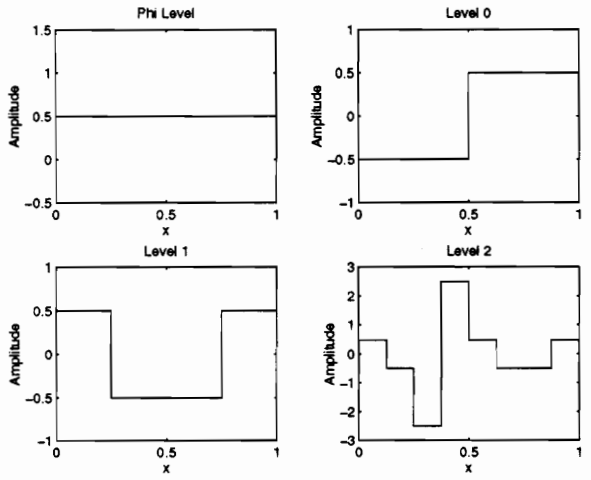


Figure 3.2: Level-By-Level Wavelet Decomposition of $\mathbf{y}$ = [1 0 -3 2 1 0 1 2]

This decimation, when considered in conjunction with the convolutions of (3.11) and (3.14) further discussed below, lead to the QMF pair model for the wavelet transform. It is from this perspective that the FWT is derived.

The accuracy condition on coefficients, (2.5.2), was defined in terms of maximum-order order zeros in the Fourier transform of $\phi(x)$. It was shown that this condition is satisfied when the discrete-time Fourier transform (DTFT) of the coefficient sequence has an n^{th}-

order zero at $\xi = \pi$. In other words, the coefficient sequence defines a low-pass filter. Referring to (2.22), the relationship between the coefficients (the a_k's used to construct the mother function and the wavelets is given by:

$$h(N-1-x) = (-1)^{N-x} g(x) \tag{3.15}$$

where $h(x)$ represents the coefficient sequence used to generate the wavelets and $g(x)$ represents the coefficient sequence used to generate the mother function.

It can be shown that, if $g(x)$ has a low-pass characteristic, then $h(x)$ will be high-pass in nature. Considered together, the sequences $h(x)$ and $g(x)$ constitute a QMF pair. This filter configuration is used in multirate signal processing systems, specifically in the subband coding of speech signals. In this scheme, a speech signal is subdivided by octave bands in order to isolate lower frequencies, where most of the speech energy is concentrated. The QMF pair is used to perform the frequency subdivisions as well as reconstruct the signals after decoding. Continuing the analogy to the wavelet decomposition, subband coding techniques incorporate decimation by a factor of two.

At this point, the QMF model is not completely clear as the convolutions of (3.11) and (3.14) do not represent explicit convolutions of the signal $s(x)$ with $g(x)$ and $h(x)$, the a_k coefficients. Rather, these convolutions are with $\phi(x)$ and dilations/translations of $\psi(x)$. Although $\phi(x)$ and $\psi(x)$ implicitly include the a_k coefficients through (2.21) and (2.22), they also depend upon dilations of $\phi(x)$ and it is thus not evident that the low-pass/high-pass characteristics still hold. However, it will be shown in the discussion of the FWT, (see 3.3), that wavelet transforms and their inverse can be computed using the a_k terms only. Thus, the QMF model is valid.

Given the implications of (3.7), the wavelet decomposition contains only wavelet terms at all levels other than $j = 0$. At these levels, the sequence under wavelet transform is both low-pass and high-pass filtered; however, only the high-pass results (the $d_{j,k}$'s) remain part of the transform. The low-pass filtered sequence is again passed through the QMF pair. The successive low-pass (averaging) filtering operations yield increasingly coarser

approximations of the original sequence while the high-pass outputs provide the detail lost in approximating V_j by V_{j-1}. As as result, the $c_{j,k}$'s are referred to as the approximation, or averaging, coefficients while the $d_{j,k}$'s are termed the detail coefficients. A block diagram representing the QMF model is shown in Figure 3.3. In the following section, the QMF model is combined with the recursion of (2.21) in developing the FWT.

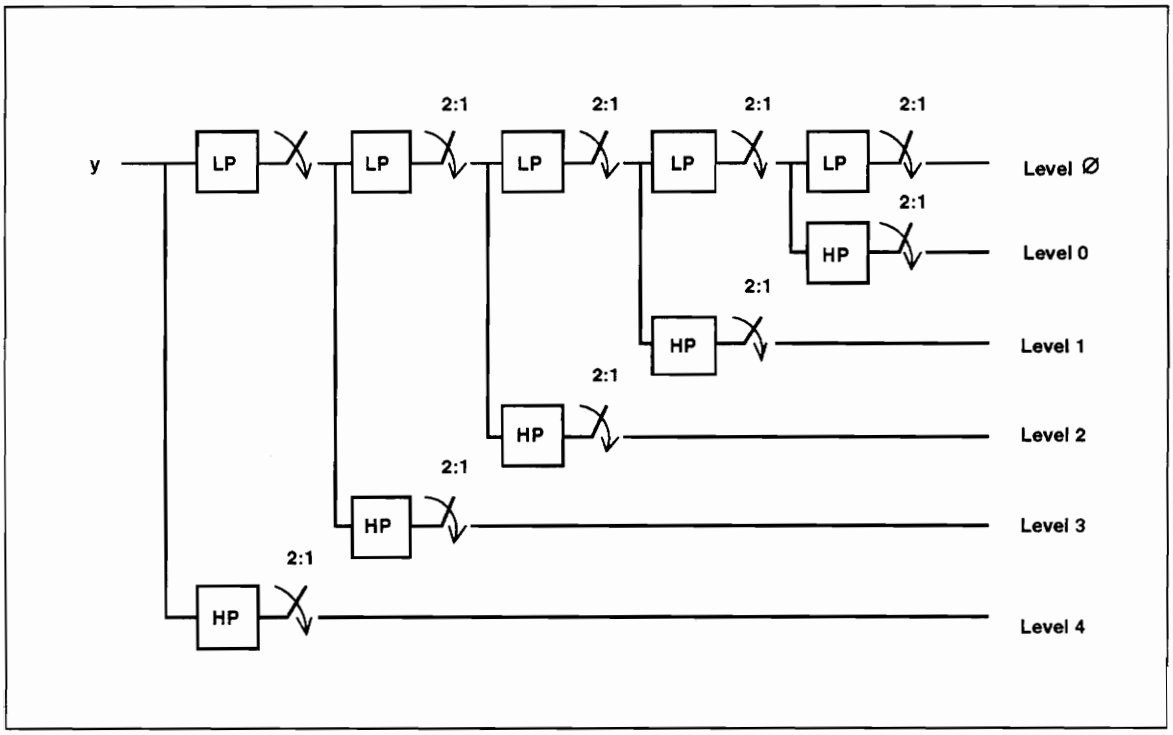


Figure 3.3: The QMF Model of the DWT for a Sequence, $\mathbf{y}$, of Length $L = 32$.

3.3 The FWT

The FWT is the wavelet equivalent of the FFT used in Fourier analysis; a comparison of these two analysis techniques is presented in Chapter 4. The following discussion provides a framework, from the wavelet perspective, for performing the comparative evaluation.

The keys to formulating the FWT lie in exploiting: (1) the fact that the iterations of the dilation equation, used to produce the mother function, begin with the unit box; (2)

the recursion found in the dilation equation; and (3) the sparseness of the transform matrix (refer to (2.17)). The FWT may be best understood by first considering its inverse. Once again, without loss of generality, the specific case involving an eight-element sequence $s(x)$ is considered [3]. The DWT of this sequence is given by:

$$\mathbf{w} = [\ c_{\phi,0} \quad \psi_{00} \quad \psi_{10} \quad \psi_{11} \quad \psi_{20} \quad \psi_{21} \quad \psi_{22} \quad \psi_{23}\ ].$$

The first term of the expansion above multiplies the mother function in the decomposition. As mentioned above, it is known that (2.21) is initiated with the unit box. Extending this idea, the first term of the wavelet decomposition can be expressed in terms of (2.21) by replacing the unit box by the box whose width occupies the interval [0,1) and has height, $c_{\phi,0}$. Using (2.21), this box may represented by:

$$c_{\phi,0} \cdot \phi(x) = c_{\phi,0} \sum_k a_k \phi(2x - k). \tag{3.16}$$

Considering the case of wavelets with four coefficients, one iteration leads to a stair-like function that now covers the interval [0,2). This new function is denoted $\phi_1(x)$. If this expression is then substituted into (2.12), a 'wraparound' effect becomes apparent. First, recall that $s(x)$ is first mapped to $f(x)$ over [0,1) through (2.15). In the translations of (2.12), the portions of $\phi_1(x)$ that actually contribute to $f(x)$ are the two halves of $\phi(x)$ the first of which extends over [0,1) and the second over [1,2). In other words, $\phi_1(x)$ can be said to 'wrap around' the interval [0,1) which is reflected in the following where ϕ_1 represents the vector products.

$$\phi_1 = \begin{bmatrix} a_0 + a_2 \\ a_1 + a_3 \end{bmatrix} [c_{\phi,0}]. \tag{3.17}$$

Accepting the QMF model is more reasonable at this point. Given the unit box, the expression above is equivalent to an expression that explicitly contains $\phi(2x-k)$ terms. Thus, the wavelet transform and its inverse can be computed in terms of the wavelet coefficients alone and can be viewed in terms of low-pass and high-pass filtering operations.

In order to obtain a one-to-one correspondence between the contribution of the mother function to $f(x)$ and the number of terms in $f(x)$, the iteration is repeated two more times.

Denoting the result of this second iteration by the term ϕ_2 and neglecting the effects of wraparound, the following matrix representation of the second iteration is obtained:

$$\phi_2 = \begin{bmatrix} a_0 & \\ a_1 & \\ a_2 & a_0 \\ a_3 & a_1 \\ & a_2 \\ & a_3 \end{bmatrix} \begin{bmatrix} a_0 + a_2 \\ a_1 + a_3 \end{bmatrix} [c_{\phi,0}]. \tag{3.18}$$

The matrix reflects translations and dilations which yield a somewhat sparse matrix. The matrix can be rewritten as shown below to reflect the 'wraparound' effect:

$$\phi_2 = \begin{bmatrix} a_0 & a_2 \\ a_1 & a_3 \\ a_2 & a_0 \\ a_3 & a_1 \end{bmatrix} \begin{bmatrix} a_0 + a_2 \\ a_1 + a_3 \end{bmatrix} [c_{\phi,0}]. \tag{3.19}$$

A general, but not analytical, technique for representing the effects for 'wraparound' is taken from the matrix expressions for the first and second iterations. For iteration i, extract a submatrix of size 2^i x i from the matrix which represents the expansion of (2.21). The first row and first column of this matrix should contain the coefficient a_0. Transpose any of the coefficients in the i^{th} column, that fall below this matrix, into the beginning empty rows of column i. Consider the matrix of (3.18), generated by the second iteration, $i = 2$. Extract the 4 x 2 matrix with first row and column taken from row 3 and column 2 of the matrix in (3.18). The coefficients a_2 and a_3 fall below this 4 x 2 matrix and are thus transposed into the first two rows of the second column thus realizing the 'wraparound' effect in matrix representation.

The process is repeated for the third iteration and, following the procedure for incorporating 'wraparound,' the following is obtained:

$$\phi_3 = \begin{bmatrix} a_0 & & & a_2 \\ a_1 & & & a_3 \\ a_2 & a_0 & & \\ a_3 & a_1 & & \\ & a_2 & a_0 & \\ & a_3 & a_1 & \\ & & a_2 & a_0 \\ & & a_3 & a_1 \end{bmatrix} \begin{bmatrix} a_0 & a_2 \\ a_1 & a_3 \\ a_2 & a_0 \\ a_3 & a_1 \end{bmatrix} \begin{bmatrix} a_0 + a_2 \\ a_1 + a_3 \end{bmatrix} [c_{\phi,0}]. \tag{3.20}$$

This product yields the 8 x 1 vector $\boldsymbol{f}^{\phi}(x)$, which represents the contribution of $c_{\phi,0} \cdot \phi(x)$ to the function $f(x)$. Equivalently, $\boldsymbol{f}^{\phi}(x)$ is the ϕ level of the transform. The matrices of (3.20) are denoted $\boldsymbol{M}_3$, $\boldsymbol{M}_2$, and $\boldsymbol{M}_1$ as they are associated with the first, second, and third iterations.

The procedure for determining the contribution of the undilated/untranslated wavelet, $w_1 \cdot \psi(x)$ to $f(x)$ is much the same as that used in the $\phi(x)$ case. The only difference is that the vector used to express the wavelet in terms of the mother function is (from (2.22)):

$$\begin{bmatrix} -a_3 \\ a_2 \\ -a_1 \\ a_0 \end{bmatrix} \tag{3.21}$$

which becomes, after the 'wraparound' effect is considered,

$$\boldsymbol{G}_1 = \begin{bmatrix} -a_3 - a_1 \\ a_2 + a_0 \end{bmatrix}. \tag{3.22}$$

Then, the contribution of $\psi(x)$ to $f(x)$, $\boldsymbol{f}^{(0)}(x)$ (the $j = 0$ level of the transform) is found from:

$$\boldsymbol{f}^{(0)}(x) = \boldsymbol{M}_3 \boldsymbol{M}_2 \boldsymbol{G}_1 d_{0,0}. \tag{3.23}$$

The method outlined above is repeated to obtain the contributions from $d_{1,0}\cdot\psi(2x)$ and $d_{1,1}\cdot\psi(2x-1)$, as well as those from $d_{2,0}\cdot\psi(4x)$, $d_{2,1}\cdot\psi(4x-1)$, $d_{2,2}\cdot\psi(4x-2)$, and $d_{2,3}\cdot\psi(4x-3)$. The reduced matrix equation that combines the contributions of $d_{1,0}\cdot\psi(2x)$ and $d_{1,1}\cdot\psi(2x-1)$, is given by:

$$\boldsymbol{f}^{(1)}(x) = \boldsymbol{M}_3\boldsymbol{G}_2 \begin{bmatrix} d_{1,0} \\ d_{1,1} \end{bmatrix} \tag{3.24}$$

where,

$$\boldsymbol{G}_2 = \begin{bmatrix} -a_3 & -a_1 \\ a_2 & a_0 \\ -a_1 & -a_3 \\ a_0 & a_2 \end{bmatrix}. \tag{3.25}$$

Similarly, the calculation of $\boldsymbol{f}^{(2)}(x)$ assumes the form:

$$\boldsymbol{f}^{(2)}(x) = \boldsymbol{G}_3 \begin{bmatrix} d_{2,0} \\ d_{2,1} \\ d_{2,2} \\ d_{2,3} \end{bmatrix} \tag{3.26}$$

where,

$$\boldsymbol{G}_3 = \begin{bmatrix} -a_3 & & & -a_1 \\ a_2 & & & a_0 \\ -a_1 & -a_3 & & \\ a_0 & a_2 & & \\ & -a_1 & -a_3 & \\ & a_0 & a_2 & \\ & & -a_1 & -a_3 \\ & & a_0 & a_2 \end{bmatrix}. \tag{3.27}$$

Having developed the orthogonality constraints on the wavelet coefficients in detail,

certain properties exist with regard to particular matrix products:

$$\tfrac{1}{2}\boldsymbol{M}_r^T\boldsymbol{M}_r = \boldsymbol{I} \quad \boldsymbol{M}_r^T\boldsymbol{G}_r = \boldsymbol{0}$$
$$\boldsymbol{G}_r^T\boldsymbol{M}_r = \boldsymbol{0} \quad \tfrac{1}{2}\boldsymbol{G}_r^T\boldsymbol{G}_r = \boldsymbol{I}$$

The QMF model of the DWT, the FWT, can now be completed by defining new variables:

$$\boldsymbol{H} = \boldsymbol{G}^T \quad \text{and} \quad \boldsymbol{L} = \boldsymbol{M}^T$$

and using the properties listed above to arrive at the following:

$$\tfrac{1}{2}\boldsymbol{L}_r\boldsymbol{M}_r = \boldsymbol{I} \qquad \boldsymbol{L}_r\boldsymbol{G}_r = \boldsymbol{0}$$
$$\boldsymbol{H}_r\boldsymbol{M}_r = \boldsymbol{0} \quad \tfrac{1}{2}\boldsymbol{H}_r\boldsymbol{G}_r = \boldsymbol{I}$$

where the variables $\boldsymbol{H}$ and $\boldsymbol{L}$ are used to represent high-pass and low-pass filtering, respectively. Mathematically, the wavelet decomposition is found using the properties above and simply reversing the operations of the inverse transform expressed in (3.20), (3.23), (3.24), (3.26). The wavelet transform is thus represented by the expressions below:

$$\begin{aligned}
c_{\phi,0} &= \tfrac{1}{2}\boldsymbol{L}_1\tfrac{1}{2}\boldsymbol{L}_2\tfrac{1}{2}\boldsymbol{L}_3\boldsymbol{f} \\
d_{0,0} &= \tfrac{1}{2}\boldsymbol{H}_1\tfrac{1}{2}\boldsymbol{L}_2\tfrac{1}{2}\boldsymbol{L}_3\boldsymbol{f} \\
\begin{bmatrix} d_{1,0} \\ d_{1,1} \end{bmatrix} &= \tfrac{1}{2}\boldsymbol{H}_2\tfrac{1}{2}\boldsymbol{L}_3\boldsymbol{f} \\
\begin{bmatrix} d_{2,0} \\ d_{2,1} \\ d_{2,2} \\ d_{2,3} \end{bmatrix} &= \tfrac{1}{2}\boldsymbol{H}_3\boldsymbol{f}
\end{aligned} \tag{3.28}$$

These expressions should be compared with Figure 3.3, inserting the $\frac{1}{2}$ gain terms where appropriate. The computational effort required in computing the FWT is best seen through example. Consider the case above of four ($W = 4$) wavelet coefficients and an eight-element vector ($L = 8$). Computing the terms $d_{2,0}$, $d_{2,1}$, $d_{2,2}$, and $d_{2,3}$ requires $\frac{1}{2}L\text{x}W$ multiplications by the high-pass filtering with $\boldsymbol{H}_3$, see (3.28). From (3.28), the sequence, $\boldsymbol{f}$ is

also filtered by L_3, again requiring $\frac{1}{2}LxW$ multiplications. Therefore, LxW multiplications are required at the first QMF pair. The decimated sequences are passed through the QMF pair again and the number of products, by comparison with the previous set of filtering operations, is halved ($\frac{1}{2}LxW$). Finally, $\frac{1}{4}LxW$ multiplications are computed at the last QMF pair. The total number of product terms in a wavelet transform is thus proportional to $2LW$.

The DWT is only one of a family of orthogonal bases distinguished by different subband intervals. This family, known as the wavelet packet library, and the associated transformations are the subject of the next section.

3.4 The Wavelet Packet Transform (WPT)

The subspaces created by the DWT are characterized by (3.8). Combining these properties with the QMF pair model reveals that the subspaces of the wavelet transform constitute a disjoint cover of the frequency space occupied by the original dataset. Coifman, [7] and [9], has shown that any set of subspaces which forms a disjoint cover of the frequency space is itself an orthogonal basis. The family of orthogonal bases whose members are distinguished by different subband intervals is known as the wavelet packet library. Graphically, the wavelet packet library is depicted by a binary tree. Table 3.1 is constructed for the vector, $\mathbf{x}$, of length $L = 8$, where the averaging and detail coefficients are denoted as p_k and q_k, respectively, at each level of the transform. The tree is presented in tabular form as this representation is useful in discussing the wavelet packet transform.

Table 3.1: The Wavelet Packet Library

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
p_1	p_2	p_3	p_4	q_1	q_2	q_3	q_4
$(pp)_1$	$(pp)_2$	$(qp)_1$	$(qp)_2$	$(pq)_1$	$(pq)_2$	$(qq)_1$	$(qq)_2$
$(ppp)_1$	$(qpp)_1$	$(pqp)_1$	$(qqp)_1$	$(ppq)_1$	$(qpq)_1$	$(pqq)_1$	$(qqq)_1$

The second row of Table 3.1 corresponds to the $j = 2$ level of the transform. The third

and fourth rows represent the $j = 1$ and $j = 0$ levels, respectively. The columns within a given row indicate the subbands that result from successive application of the QMF pair to the averaging coefficients. In other words, at the $j = 0$ level there have been three successive applications of the QMF pair. At this point, only low-frequency (undilated) basis functions remain to approximate the original signal. This approximation is coarse (see Figure 3.2). The high-pass component of the QMF produces the q_k coefficients while the low-pass filtering yields the p_k terms. Table 3.1 mirrors the level-by-level dynamics of the QMF pair filter bank.

Additional orthogonal bases can be constructed through application of the QMF pairs to either the averaging or detail coefficients at each level. The location of the subband intervals, for a given level, are determined by how the QMF pair is applied in the previous stage. The wavelet packet transform (WPT) yields a subset, subject to the disjoint cover condition, of the entities found in the wavelet packet library. A basis taken from the wavelet packet library is known as a wavelet packet basis. Disjoint covers are manifest in the binary tree as graphs. From Table 3.1, a graph is constructed by selecting any group of whole boxes such that each column in the grid contains exactly one element.

Several wavelet packet bases are shown in Tables 3.2, 3.3, and 3.4. Table 3.2 represents the DWT while Table 3.3 depicts the 'opposite' of the DWT. Table 3.4 is the subband basis taken at the $j = 1$ level.

Table 3.2: The Basis of the DWT

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
				q_1	q_2	q_3	q_4
		$(qp)_1$	$(qp)_2$				
$(ppp)_1$	$(qpp)_1$						

In the DWT, the cascade of QMF pairs produces good frequency localization at low frequencies. The WPT , combined with the different types of wavelet functions available, is a much more flexible analysis technique. The WPT can be used to isolate a given spectral

Table 3.3: The Basis of the 'Opposite' DWT

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
p_1	p_2	p_3	p_4				
				$(pq)_1$	$(pq)_2$		
						$(pqq)_1$	$(qqq)_1$

Table 3.4: A Subband Basis

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
$(pp)_1$	$(pp)_2$	$(qp)_1$	$(qp)_2$	$(pq)_1$	$(pq)_2$	$(qq)_1$	$(qq)_2$

region by applying the QMF pairs to the appropriate sets of detail and approximation coefficients. As an example, the 'opposite' of the DWT (Table 3.3), localizes the high frequency region. Choosing the optimal basis for a given signal is the subject of 5.2.2.

Chapter 4

FOURIER & WAVELET ANALYSIS - DIFFERENT PERSPECTIVES

Fourier and wavelet transforms provide techniques for decomposing a signal in order to evaluate its characteristics in another domain. Due to the differences of these decompositions, as determined by the basis functions, the resultant analyses are given different names. Fourier transforms lead to a *time-frequency analysis* while the DWT/WPT are said to yield a *time-scale analysis* . Therefore, paramount to understanding the difference between Fourier and wavelet analysis is discerning between the concepts of scale and frequency [8].

The time-scaling of a function $s(t) \rightarrow s(at)$ with $a > 0$ results in either contraction ($a > 1$) or expansion ($a < 1$). Therefore, referring to the expression for the CWT (2.3), it is seen that for increasing values of the scale factor, wavelet basis functions (dependent on $1/a$) expand. As a result, only long term behavior in $s(t)$ is captured. A change of variable in the CWT, $t = at$, leads to the equivalent expression:

$$S(b,a) = |a|^{\frac{1}{2}} \int_{-\infty}^{\infty} s(at)\psi(t - \frac{b}{a})dt \tag{4.1}$$

Therefore, an alternative view of wavelet analysis involves a signal that is contracting as the scale factor increases. However, resolution ability is held constant - large scales correspond to global views with local perspectives corresponding to smaller scales.

Scale, however, cannot be regarded as a parameter independent of frequency due to the inherent relationship between time and frequency. The term scale, aside from the meaning assigned above, is also used to avoid confusion with frequency as considered in more traditional Fourier analysis. When scaling is performed on a wavelet, the local frequency of the wavelet changes. However, this frequency is not related to the oscillations that characterize the Fourier basis functions.

Fourier and wavelet analysis are similar in that the FFT and FWT are both linear and can be expressed as matrix operations. The inverse transforms are computed using the inverse of the original transform matrix. The DWT/WPT of a signal of length 2^N can be stopped at an arbitrary level $N - k$, $k > 0$. This partial transform is also a valid transform for another sequence of length 2^{k+1}. QMFs are thus considered the equivalent

of the butterfly operation used in FFT algorithms, both serving as orthogonal transform kernels. Additionally, the basis functions used in Fourier and wavelet analysis are well-localized in frequency, providing the opportunity to perform frequency-based analysis.

In terms of the differences between the two techniques, the primary discrepancy lies in the time-domain localization found in wavelets as opposed to the infinite-duration sinusoids found in the Fourier basis. This property allows wavelets to approximate signals with sharp peaks/discontinuities more efficiently. Time-frequency resolution can be quantified as a product of the time-domain duration of the analyzing function, Δt, with the bandwidth of the analyzing function, Δf, [10]. This product remains constant for all analyzing functions of a given transform. The simultaneous nature of time and frequency suggests a limit on this resolution which is found from Heisenberg's uncertainty principle and expressed by:

$$\Delta t \cdot \Delta f \leq \frac{1}{4\pi}. \tag{4.2}$$

Since wavelets are dilated and translated at each level of the DWT, they provide long, low-frequency and short, high-frequency observation windows. In the case of the WPT, varying time-frequency resolutions are allocated based on the nature of the signal under analysis. These properties make it possible to isolate time-domain discontinuities (short windows) as well as perform detailed frequency analysis (long windows) [11]. This is in contrast to the basis found in Fourier analysis where the sinusoids are all windowed by the same-duration square wave. The coverage of the time-frequency plane is often used in wavelet literature to illustrate the differences between wavelet and Fourier analysis. The different tilings of the time-frequency space are shown in Figure 4.1, taken from [11]. Viewed in terms of (4.2), the quantities Δf and Δt are constants in Fourier analysis. Under wavelet analysis, Δt and Δf change at each scale of the transform.

In Fourier analysis, scale is constant (Figure 4.1). However, the term 'time-scale' does not imply constant frequency; rather, it is the ratio $\delta f/f$ which remains invariant at a given scale (level) of wavelet analysis. The quantity $\delta f/f$ refers to the ratio of a wavelet's bandwidth to it's center frequency as seen in the Fourier transform of the wavelet. This type

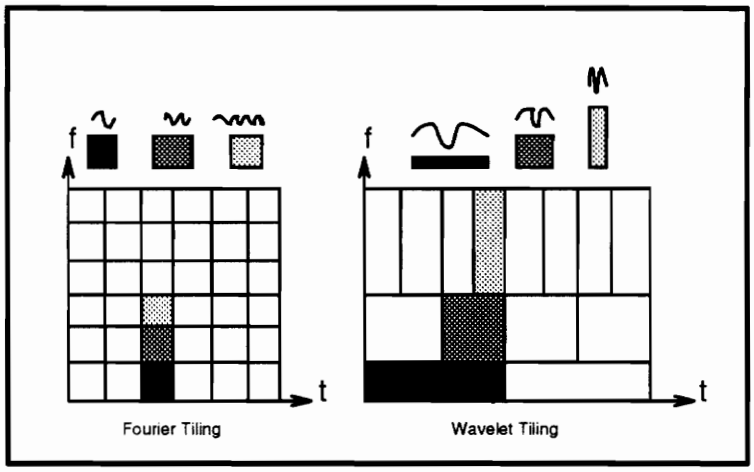


Figure 4.1: A Comparison of Time-Frequency Coverage Offered by Fourier and Wavelet Analysis.

of behavior is also seen in the human auditory system. Modeling the auditory system as a bank of filters, it has been shown that these filters have constant $\delta f/f$ or Q at frequencies above 800 Hz. These bandwidths and center frequencies are related to the psychoacoustic concept of critical bands [12]. This relationship suggests that wavelet analysis is an attractive alternative to conventional audio signal processing techniques. This is the motivation for the application discussed in Chapter 6.

Considering implementation, 3.3 revealed that the number of multiplications per transform is equal to $2LW$; this is compared to the $L\log_2 L$ value associated with the FFT. The computational intensity required in computing the FWT and FFT can be compared in an effort to find the value of W for which the FWT is less computationally intensive:

$$\begin{aligned} 2LW &\leq L\log_2 L \\ 2W &\leq \log_2 L \\ W &\leq \tfrac{1}{2}\log_2 L \end{aligned} \tag{4.3}$$

Table 4.1 shows the value of L, (L_T) above which the FWT requires less effort, given wavelets with W coefficients. It is seen that, given wavelets constructed from more than six coefficients, the FWT is more efficient than the FFT for large sequences only.

Table 4.1: Length of Transform, L_T, at which the FWT, Given Wavelets of W Coefficients, is Less Computationally Intensive than the FFT.

Number of Wavelet Coefficients (W)	Threshold Length of Transform (L_T)
2	16
4	256
6	4096
8	65536

Wavelets are viewed as either a breakthrough or as a synthesis of the ideas already surfaced in constant-Q filter analysis, QMF applications, MRA decomposition, and other, more mathematically-oriented endeavors. In [8], Rioul and Vetterli state that the merit of wavelet theory lies in unifying all of the above efforts under a common framework. Current wavelet research involves the combined efforts of pure mathematicians, image and speech processing engineers, general signal processing engineers, physicists, and applied mathematicians examining fractal behavior. Under examination by such a diverse community, the merits and shortcomings of both wavelet and Fourier analysis tools are certain to become more evident as research continues.

Chapter 5

THE PRACTICAL APPLICATION OF WAVELETS

5.1 Examples of Wavelet Applications

Currently, most research into the practical applications of wavelets is focused on data compression and noise reduction. More isolated efforts involve detecting self-similarity in a time-series and the development of a music synthesizer. These ventures are briefly discussed in the following sections. The specific application of noise reduction, with audio-specific considerations, is discussed in more detail in Chapter 6.

5.1.1 Image compression and noise reduction

The principles of data compression and noise reduction are similar, regardless of the type of signal, under wavelet analysis. The processes for implementing these principles are covered in 5.2. However, in the advent of concentrated wavelet research, David Marr proposed that the human retinal system implements a hierarchical processing scheme. Each of these processing stages performs edge detection at varying degrees of spatial tuning. This idea led to a multiscale model of the visual system and an analyzing device that is now known as the Marr wavelet. An approach for noise reduction in images, based on Marr's ideas, is oulined in [13]. One of the first and, as of this date, few practical applications of wavelet analysis has been in the compression of fingerprint images for the FBI. In this application, a 20:1 ratio has been attained without noticeable degradation.

5.1.2 Detection of self-similarity

A plot of the average of the squared wavelet coefficients against scale is known as a scalegram in wavelet analysis. This device is used to examine the relationships existing in a signal at varying degrees of resolution. Using this property, Scargle and others at the NASA-Ames Research Center have investigated the luminosity associated with one member of a binary star system. Examining self-similarity, across different scales of analysis, in both quasi-periodic oscillations and very low-frequency noise provided the foundation for characterizing the source of fluctuation in luminosity.

5.1.3 Music synthesis

Wickerhauser has suggested that wavelet packets may be used to realize a synthesizer. After a musical note has been decomposed into wavelet coefficients, playback would involve loading the coefficients into a wavelet packet generator. Attack and delay could be implemented through envelope generators. Various combinations of wavelet packets could be selected experimentally by the musician. This idea suggests a potential library of sounds more extensive than any found on contemporary music synthesizers.

5.2 Noise Reduction and Data Compression

Wavelet-based noise reduction and data compression applications are implemented through a technique termed by Donoho as 'wavelet shrinkage.' This technique is discussed in detail in [14], [15], and [16] and, as introduced by Donoho and Johnstone in [14], consists of the following steps:

(1) Compute the DWT of the noisy data.

(2) Determine/apply a threshold to the noisy wavelet coefficients ('Wavelet Shrinkage').

(4) Compute the inverse DWT.

The technique outlined above accepts, as a given, that an analyzing wavelet has been selected. Additionally, the procedure assumes the DWT; however, it has been discussed in 3.4 that the WPT is a better, more general, analysis tool. Assuming that the WPT is to be used, the technique outlined above is modified and the following approach proposed:

(1) Select an analyzing wavelet.

(2) Select a wavelet packet basis.

(3) Compute the WPT of the noisy data.

(4) Determine and apply a threshold to the noisy wavelet packet coefficients.

(5) Compute the inverse WPT.

The remainder of this section examines steps (1), (2), and (4) of the above algorithm. Steps (3) and (5) are performed as discussed previously.

5.2.1 Selecting an analyzing wavelet

Current literature largely does not address the task of selecting an analyzing wavelet. A procedure for choosing the optimal, in a mean-square error (MSE) sense, wavelet is presented in [17]. This presentation, however, assumes a noiseless environment and is best-suited to coding/data compression applications. A *super-wavelet* has been proposed by Xie and Beex in [18]. The *super-wavelet* is constructed of linear combinations of elemental wavelets already known to satisfy the necessary mathematical constraints. The elemental wavelets are constructed and combined based on AR model estimation of spectral components within each octave band. This technique was motivated, in part, by the restrictions of the orthogonal DWT that are apparent when comparing with the WPT; namely, octave-band coverage of the entire signal space. This rigid octave-band segmentation may not be appropriate for any or all portions of a given signal. This is the same difficulty that the WPT seeks to overcome. A combination of these two approaches is a possibility.

However, given that wavelets are localized in frequency, it is reasonable to select a wavelet that bears some time-domain resemblance (visual or, if possible, functional) to the signal of interest. An obvious example would be selecting the Haar basis for step-like functions. The fractal structure of the Daubechies wavelets (Figure 2.1) makes them good candidates for some naturally-occuring phenomena known to exhibit self-similarity. However, the sharp discontinuities found in the Daubechies functions suggests that other wavelets are more appropriate for smooth signals. The importance of employing a well-correlated analyzing wavelet is underscored in Chapter 6.

5.2.2 The 'best basis' algorithm

Data compression is often a primary consideration in digital signal processing systems. In wavelet-based signal processing, compression is performed not only for the sake of more efficient transmission and coding schemes. Noise reduction techniques under wavelet processing are based on compression and the notion of coherence. Coherence [19] is a measure

of the degree to which a signal, s, and a particular basis are related. Coherence is determined by the number and magnitude of coefficients with a small number of large-amplitude coefficients indicating high coherence. A wavelet packet basis to which $s(n)$ is maximally coherent is desired.

The most efficient representation of s in the wavelet packet library is found using the 'best basis' algorithm of Coifman and Wickerhauser [7] and [9]. The algorithm selects, from the wavelet packet library, the basis which minimizes some additive measure of information, M. The concept of coherence, discussed above, is related to the additive characteristic of the measure.

Several measures of information used in the 'best basis' algorithm include [9]: entropy; the number of values above a threshold; and bit counts. The bit-count measure involves counting the number of bits, above a threshold η, which are required for transmission to accuracy ϵ. The noise reduction application discussed in Chapter 6 was examined using both the entropy and values-above-a-threshold measures. The entropy measure, more specifically the Shannon-Weaver entropy, of a sequence x is given below:

$$H(s) = -\sum_k p_k \log p_k, \tag{5.1}$$

where,

$$p_j = \frac{|s_k|^2}{\| s \|^2}.$$

This measure is made additive by substituting the expression:

$$\lambda(s) = -\sum_k |s_k|^2 \log |s_k|^2$$

into (5.1) which results in:

$$H(s) = \| s \|^{-2} \lambda(s) + \log \| s \|^2 . \tag{5.2}$$

The number-above-a-threshold measure is found by counting the number of elements in a sequence (the coefficients of the wavelet decomposition) whose value exceeds the threshold, t_n. This threshold may be fixed or selected adaptively. An adaptive technique for determining t_n is discussed in 5.2.6.

Once an additive information measure is defined, a basis is selected by searching the binary tree representing the wavelet packet library. The wavelet packet basis for which s has the minimal M is the best basis.

5.2.3 Wavelet shrinkage

The WPT is computed after finding the best basis. Once wavelet packet coefficients are obtained, wavelet shrinkage is applied. Wavelet shrinkage then exploits coherence and the orthogonality of the wavelet packet basis in a very simple algorithm. Due to coherence, the wavelet packet coefficients that correspond to the signal of interest should be few in number with relatively large magnitudes (discussed further below). As a consequence of basis orthogonality, noise (assumed to be white and Gaussian), is transformed into white, Gaussian noise in the wavelet domain. Thus, a threshold is applied and the noise-reduced signal is reconstructed through the inverse WPT.

Thresholding is the key operation in wavelet-based data compression and noise reduction as it serves as the compressor and noise-removing filter. Wavelet coefficients lying above the threshold are either shrunk toward the origin (soft thresholding) or left unaltered (hard thresholding). The hard and soft thresholding functions are illustrated below in Figure 5.1. The soft threshold operation, performed on a signal w, is given by:

$$\eta_s(w,\lambda) = sgn(w)(\| w \| - \lambda)_+ \tag{5.3}$$

while hard thresholding on the same signal is expressed as:

$$\eta_h(w,\lambda) = \begin{cases} w & , \quad \| w \| > \lambda \\ 0 & , \quad \text{otherwise} \end{cases} \tag{5.4}$$

In the remainder of this section, wavelet shrinkage is justified and methods for determining the associated threshold levels are discussed.

Donoho, [14] and [15], proposes three shrinkage schemes differing in the threshold level implemented and are termed as such: (1) RiskShrink, (2) VisuShrink, and (3) SUREShrink.

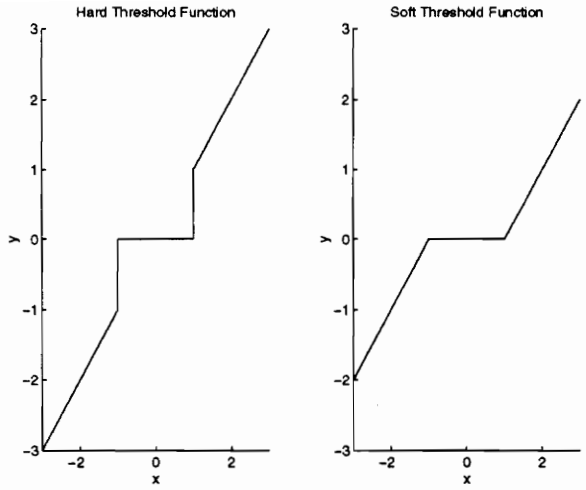


Figure 5.1: Soft and Hard Thresholding Functions Used in Wavelet Shrinkage.

These shrinkage techniques are derived from multivariate normal decision theory and are based on spatial adaptation principles. RiskShrink and VisuShrink are variations on the same fundamental concept while SUREShrink incorporates a different minimizing criteria, distinguishing it from the other procedures.

As the name suggests, spatial adaptation estimation incorporates spatially-adaptive estimators, such as a family of translated and dilated wavelets. Spatial noise reduction techniques are possible due to the localized nature of wavelets and the orthogonality present in the wavelet packet basis. Briefly, the idea is that if a function undergoes signficant change at time t, which is aligned with a particular dilation and translation of the wavelet, then the coefficient of that wavelet will be large. The orthogonally-transformed noise source is spatially homogeneous relative to the signal of interest $s(x)$ which generally undergoes substantial change in a small temporal region as indexed by the dilation and translation parameters, j and k. As a result, there will be a few clusters of large wavelet coefficients representing $s(x)$. Noisy wavelet coefficients are scattered throughout the wavelet domain and the rationale for thresholding becomes apparent.

Consider the situation of a signal θ corrupted with additive, white Gaussian noise $g \sim$

$N(0,\sigma)$ This can be expressed as shown below where, $z \sim N(0,1)$:

$$w_i = \theta_i + \sigma z_i \;\; i = 1, \ldots, n. \tag{5.5}$$

In spatial adaptation methods, diagonal linear projectors are examined. These operators perform the function:

$$T_{DP}(w, \delta) = (\delta_i \cdot w_i)_{i=1}^{n} \text{ where } \delta_i \in \{0, 1\} \tag{5.6}$$

In ideal spatial adaptation, an oracle $\hat{\theta}_{IDEAL}$ is present. The oracle would suggest that estimation be performed only for those θ_i which are larger that σ. This idea leads to the ideal risk, the MSE measure:

$$R(\hat{\theta}_{IDEAL}, \theta) = \sum_i \min(\theta_i^2, \sigma^2). \tag{5.7}$$

Practically, this oracle does not exist and the aim is to approach ideal risk performance through a data-based approximation to the oracle. This is the goal of the RiskShrink, VisuShrink, and SUREShrink processes.

5.2.4 RiskShrink

Considering the soft threshold estimator, pictured in Figure 5.1 and expressed in (5.3), Donoho [14] proposes the estimator:

$$\hat{\theta}_i^u = \eta_s(w_i, \sigma(2\log n)^{\frac{1}{2}}) \;\; i = 1, \ldots, n \tag{5.8}$$

which satisfies,

$$E\{\| \hat{\theta}^u - \theta \|_{2,n}^2\} \leq (2\log n + 1)\{\sigma^2 + \sum_{i=1}^{n} \min(\theta_i^2, \sigma^2)\} \text{ for all } \theta \in \Re^n. \tag{5.9}$$

Comparing (5.9) and (5.7), the proposed estimator approximates the ideal risk, plus the parameter σ^2, to within a factor of essentially $2\log n$. It is now natural to attempt finding a better threshold, λ^*, which yields a smaller constant Λ^* than the $2\log n + 1$ factor

of (5.9). In [14], Donoho develops this estimator and shows it to be the soft threshold operator incorporating the minimax threshold, λ^*,

$$\hat{\theta}^* = \eta_s(w_i, \lambda_n^* \sigma) \tag{5.10}$$

with,

$$\lambda_n^* \equiv \text{ the largest value attaining } \Lambda_n^*$$

where,

$$\Lambda_n^* \equiv \inf_{\lambda} \sup_{\mu} \frac{\rho_{ST}(\lambda, \mu)}{(n^{-1} + \min(\mu^2, 1))}.$$

The function $\rho_{ST}(\lambda, \mu)$ is related to the soft threshold function by:

$$\rho_{ST}(\lambda, \mu) = E\{(\eta(Y, \lambda) - \mu)^2\} \tag{5.11}$$

where,

$$Y \sim N(\mu, 1).$$

It is shown in [14] that Λ_n^* is much lower than the $(2\log n + 1)$ quantity for small n (on the order of a few hundred) but for large n, the $(2\log n + 1)$ quantity is asymptotically optimal. For 2000-element sequences on which Donoho experimented, the MSE using the λ^* threshold was roughly half of that seen in the $2\log n$ case, but both were very small. The estimation procedure based on wavelet shrinkage under the threshold λ_n^* is termed RiskShrink, where 'Risk' refers to the fact that the estimator minimizes the MSE, a measure of risk.

5.2.5 VisuShrink

As an alternative to the minimax threshold, λ_n^*, the $2\log n$ quantity is revisited. It can be said that the estimator based on this value is better, in a visual sense, than that incorporating λ_n^* as n becomes large. This threshold is less restrictive than λ_n^* and thus the risk of suppressing signal components is lower, possibly yielding better visual results. This 'visual' estimator relies on:

$$p\{\max_i \parallel z_i \parallel > (2\log n)^{\frac{1}{2}} \to 0 \text{ as } n \to \infty. \tag{5.12}$$

Thus, given $z_i \sim N(0,1)$ and large n, this estimator is expected to perform better than the λ_n^*-based estimator. The estimation procedure based on wavelet shrinkage under the 'visual' threshold, $2\log n$, is termed VisuShrink.

As a final note, the discussions above are developed within the context of a soft threshold. However, the hard threshold operator exhibits the same asymptotic behavior limit of $(2\log n)^{\frac{1}{2}}$, [14].

5.2.6 SUREShrink

In [15], a new estimator based on an adaptive threshold is proposed. Again, the form of the estimator is the soft threshold operator and the optimization criteria is a MSE measure. Given an n-dimensional vector representing the desired signal, θ, and n noisy observations $w_i \sim N(\theta_i, 1)$ $i = 1, \ldots, n$, the estimator $\hat{\theta} = \theta$ is considered. Stein [20], has shown that the loss $\| \hat{\theta} - \theta \|^2$ may be estimated unbiasedly for most estimators.

The estimator is written as $\hat{\theta}(\mathbf{w}) = \mathbf{w} + \mathbf{g}(\mathbf{w})$ where $\mathbf{g}(\mathbf{w})$ is the soft threshold limiter. Stein's result is that, if $\mathbf{g}(\mathbf{w})$ is weakly differentiable, then:

$$E_\theta\{\| \hat{\theta}(\mathbf{w}) - \theta \|^2\} = \mathrm{n} + \mathrm{E}_\theta\{\| \mathbf{g}(\mathbf{w}) \|^2 + 2 \cdot \nabla \cdot \mathbf{g}(\mathbf{w})\}. \tag{5.13}$$

Substituting $\eta_s(w_i, t)$ for $\mathbf{g}(\mathbf{w})$, Stein's Unbiased Estimate of Risk (SURE) is obtained:

$$SURE(t; \mathbf{w}) = \mathrm{n} - 2 \cdot \#\{\mathrm{i} : |\mathrm{x_i}| \leq \mathrm{t}\} + \sum_{\mathrm{i}=1}^{\mathrm{n}} (|\mathrm{x_i}| \wedge \mathrm{t})^2 \tag{5.14}$$

and is, in turn, used to determine a threshold value,

$$t_{SURE} = \arg\min_{t \geq 0} SURE(t; \mathbf{w}). \tag{5.15}$$

As is the case with the minimax and visual thresholds, t_{SURE} is asymptotically optimal as n becomes large. The development outlined above is the foundation for SUREShrink.

The SURE-based technique is problematic when the wavelet coefficients are *extremely* sparse. Extreme sparsity results not from a large degree of coherence between θ and the wavelet packet basis, but when θ has a large number of zero-valued elements. Consider

the second term of (5.14), $-2 \cdot \#\{i : |w_i| \leq t\}$, in the situation of extreme sparsity. This quantity will contribute significantly to the SURE measure leading to poor results. Donoho proposes a hybrid threshold to compensate.

In the hybrid method, the terms $\beta_n = \log_2(n)^{\frac{3}{2}}$ and $s_n^2 = n^{-1} \sum_i (w_i^2 - 1)$ are defined. Comparing these two values gives an indication of the degree of sparsity in the sequence. In the event of extreme sparsity, $s_n^2 \leq \beta_n$, the visual threshold is selected. The hybrid threshold is given by:

$$t_{hybrid} = \begin{cases} t_{SURE} & s_n^2 > \beta_n / \sqrt{(n)} \\ t_{visual} & s_n^2 \leq \beta_n / \sqrt{(n)} \end{cases} \tag{5.16}$$

The remaining components of the shrinkage algorithm are executed as outlined in 5.2. As a final comment, Donoho [15] proves that SURE estimation achieves the optimal rate of convergence to the large-n MSE bound. Since the SURE technique assumes no knowledge of the object under estimation, this estimator is termed optimally smoothness-adaptive.

5.2.7 Local trigonometric analysis

Generally, information-bearing audio signals are periodic to some degree. This property, from a spectrally-based perspective, motivates the search for a Fourier-like approach to analyzing these signals. As discussed earlier, the Fourier transform, while orthogonal and frequency-localized, suffers from infinite-duration support in the time domain. This is in contrast to the signals under consideration which are correlated over temporally local regions. The discrete-sine and discrete-cosine transforms (DST and DCT) are an attempt at recovering this time-domain localized behavior. The DST and DCT analyze segments of the signal separately and achieve better temporal resolution. However, this segmentation and improved resolution lead to some unfortunate side-effects as outlined in [21]:

- As a result of their periodic nature, the DST and DCT do not perform well with segments of a signal that are not periodic.

- The DST and DCT may perform too well for a given segment. If this 'good' performance is repeated for a neighboring segment, the two entities become distinct and inter-segment continuity is lost (the blocking effect in image processing terms).

- Arbitrary segmentation discounts correlation among regions.

Coifman and Meyer [7] propose the local sine and local cosine transforms. Denoted LST and LCT to distinguish them from the DST and DCT, these bases are not subject to the first two concerns listed above. Briefly, the idea involves splitting the signal under analysis through smooth windowing functions. Overlapping portions of these segments are folded in such a way as to be suited to the *smoothed* trigonometric basis. This new *smoothed* trigonometric basis is obtained by multiplying the local sine and cosine functions by a bell-like function. Additionally, the presentation in [7] is general and the 'best basis' algorithm can be used to search the local trigonometric basis much like in the case of the wavelet packet library. This minimizes the adverse effect of segmentation (the third concern). The bases used in the LST and LCT expansion are given below:

$$\begin{aligned}
(LST-I)(-;-) &\Rightarrow \sqrt{\tfrac{2}{|I|}}b_I(x)\sin[k\tfrac{\pi}{|I|}(x-\alpha)];\ \ k=1,2,\ldots \\
(LST-III)(-;+) &\Rightarrow \sqrt{\tfrac{2}{|I|}}b_I(x)\sin[\tfrac{2k+1}{2}\tfrac{\pi}{|I|}(x-\alpha)];\ \ k=0,1,2,\ldots \\
(LCT-I)(+;+) &\Rightarrow \tfrac{1}{\sqrt{|I|}}b_I(x)\sqrt{\tfrac{2}{|I|}}b_I(x)\cos[k\tfrac{\pi}{|I|}(x-\alpha)];\ \ k=1,2,\ldots \\
(LCT-III)(+;-) &\Rightarrow \sqrt{\tfrac{2}{|I|}}b_I(x)\cos[\tfrac{2k+1}{2}\tfrac{\pi}{|I|}(x-\alpha)];\ \ k=0,1,2,\ldots
\end{aligned} \tag{5.17}$$

The '+' and '-' symbols contained parenthetically on the left-hand side of the above expressions indicate polarity. Polarity is defined within the context of the smooth window operator used to segment the signal. This windowing is viewed as a projection, where dilations and translations of the smooth window multiply translations and dilations of the signal. A positive polarity results when these products are added; negative polarity refers to the projection obtained through cumulative differencing of the product terms. The numerical suffix is also related to polarity and serves as an index to the proper basis for a given projection. A more complete discussion of these smoothed local trigonometric bases

and the mechanics of the LST and LCT algorithms, is presented in [22]. As a clarification in terminology, when the LST and LCT are combined with the best basis algorithms, the result is referred to as the cosine packet transform (CPT), analogous to the WPT.

The similarities between local trigonometric, and wavelet analysis are evident. Localization, a versatile library of bases, and translations/dilations in the analyzing kernel characterize both techniques. It is not, therefore, surprising that the basis functions of the LCT and LST are also closely related to a class of wavelets. It can be shown that the Fourier transforms of the Shannon wavelet,

$$\psi_{Shannon}(x) = \frac{\sin x}{x} \tag{5.18}$$

and its translations/dilations, form a basis for $\mathbf{L}^2(\Re)$. If this basis is smoothed by a window that is infinitely continuously differentiable, the Meyer wavelet is obtained [2]. The Meyer wavelet is quite different from those of Figure 2.1 derived through the dilation equation, and has its own transform. Kolaczyk discusses implementing the Meyer transform in [22].

Finally, it is noted that since localization and orthogonality exist in the LST and LCT, the principles of spatial adaptation are applicable to these analysis tools as well. As a result, the shrinkage techniques discussed in 5.2.4, 5.2.5, and 5.2.6 may be integrated into the local trigonometric evaluation framework. The procedure of 5.2 is modified to accommodate this new approach by simply replacing the WPT with the CPT.

Wavelet shrinkage is seen to be a relatively simple approach to noise reduction and data compression, an important distinction when compared to Fourier-based methods. An additional distinguishing factor between the techniques of both paradigms results from the nearly ideal diagonal projector operators offered by wavelet shrinkage. Smoothing under wavelet shrinkage can lead to more appealing visual results since features are not broadened as in Fourier-based approaches. Now that the general approach for noise reduction has been introduced, attention is turned toward a specific application. The following section examines shrinkage techniques as applied to audio signals where visual appearance is of little importance.

Chapter 6

NOISE REDUCTION IN AUDIO SIGNALS

6.1 Description of Signals

Five noise-degraded signals are chosen for evaluating the shrinkage algorithms. The signals are selected to provide a range of environments for examining the noise reduction technique. A brief description of each is given below.

- *Enrico Caruso, 'La Bella'* - A short segment (the phrase 'La Bella') from a 1903 recording of Enrico Caruso singing an aria from Puccini's Tosca. The primary sources of corruption are the familiar scratches and pops associated with analog vinyl pressings. However, in this recording these types of disturbances are exaggerations of those normally found.

- *The Grateful Dead, 'Viola Lee Blues'* - A short segment of The Grateful Dead performing Viola Lee Blues, recorded in 1968. The recording is relatively clean without large noise transients but suffers from ambient low-level degradation and is interesting in that the segment is gradually 'faded' in and out.

- *The Wizard of Oz, 'I Am Oz'* - A quote from the 1939 movie, 'The Wizard of Oz.' This speech segment also suffers from low-level background noise but is relatively clean of unwanted transients. This segment is interesting in that during this passage, the 'Oz' character is surrounded by machines that produce noiselike sounds. This segment is selected as it tests the algorithm's ability to distinguish between types of noise sources, some of which may not be undesirable.

- *William McKinley, 'My fellow citizens, . . .'* - A portion of one of McKinley's 'front porch' speeches delivered sometime during his term (1897-1901). This recording is degraded not only by low-level background noise but some form of distortion that has 'warped' certain passages.

- *Wolf howling in 'wind'* - A clean recording of a wolf howl is corrupted by additive white, Gaussian noise to simulate wind.

Using these signals, shrinkage techniques are applied as described in 5.2 and the associated subsections in an attempt to clean the signals. Specifics regarding the implementation are discussed in more detail in the following section.

6.2 Implementation Procedure

Important elements and considerations in implementing the noise reduction procedures presented in this paper include: the software environment; adaptation of the shrinkage technique to audio signals; and the general approach for varying algorithmic parameters. These issues are the subject of this section.

6.2.1 The Software Environment - WaveLab

All processing described in this paper was implemented using MATLAB. In addition, WaveLab routines, or modifications thereof, were an integral component of the application. WaveLab is a MATLAB toolbox developed by Buckheit, Chen, Donoho, Gao, Johnstone, Kolaczyk, and Scargle most of which are associated with the Statistics Department of Stanford University. WaveLab provides a library of routines for wavelet, wavelet packet, cosine packet and other related types of analysis and is available through anonymous ftp at **playfair.stanford.edu** in directory **\pub\wavelab**.

6.2.2 Special Implementation Issues

Shrinkage approaches assume a $N(0,1)$ corruption. As a result, an estimate of the noise signature is required so that the signal may be scaled accordingly. In [15], a procedure for obtaining the noise power is suggested which considers the high-pass output from the first QMF pair. The mean absolute deviation (MAD), robust to the existence of high-frequency components, of this filtered sequence is used as the noise standard deviation estimate. Denoting the noise estimate by $\hat{\sigma}$,

$$\hat{\sigma} = MAD(\nu)/0.6745; \tag{6.1}$$

where,

$$MAD(\boldsymbol{\nu}) = Median\{\sum_{i} |\nu_i|\}. \tag{6.2}$$

It is noted that this procedure yielded better results than estimates taken from 'quiet' passages, when available, in the signals under evaluation.

Given the pre-processed, noise-normalized, signal, the shrinkage algorithm was implemented. Referring to Figure 2.1, the Symmlet wavelet was selected for all wavelet packet realizations due to its smooth, oscillatory-like behavior. The local trigonometric, cosine packet, implementations incorporated a discrete-cosine transform.

The application of shrinkage techniques to audio signals is not new; specifically, Berger, Coifman, and Goldberg [23] have performed an extensive evaluation of these procedures. Their investigation was performed using two signals: a 1889 recording of Johannes Brahms and the Caruso pressing mentioned earlier. A key element of their approach that was employed in the evaluation discussed in this paper concerns segmentation.

The signals range in length from approximately 11,000 to 270,000 elements. In order to accurately capture the desirable dynamics in a nonstationary signal, small observation windows are necessary. There is a trade-off. Windows that are too small will not provide enough data from which to extract only short-term stationarity, thereby overemphasizing the noise contribution. Excessively long windows contain too many transition regions and an averaging effect results. Regarding segmentation, in [23], the authors caution against edge effects that can appear when the individual, independently-processed intervals are simply pasted back together. The edge effects that appear in the implementation presented in this paper are minimal or undetectable and require no further attention.

Finally, given the number of parameters that can be manipulated, an approach for evaluation is necessary and is given as a four stage process. Consider the three shrinkage variations outlined in 5.2.4, 5.2.5, and 5.2.6; it is assumed that the SUREShrink method holds the most promise. The RiskShrink and VisuShrink processes do not have the 'adaptability-to-unknown-smoothness' feature that distinguishes the SUREShrink method. The hybrid shrinkage technique is also considered even though it is believed that the extreme sparsity

condition will be inapplicable. Furthermore, it is assumed that given the class of orthogonal wavelets considered, local trigonometric analysis will be superior. This idea is discussed in [23]. Essentially, the audio signals of interest are composed of distinct frequencies which can be approximated by local trigonometric functions better than wavelets which are localized over octave intervals.

Therefore, in the first stage, the expected worst case analysis tool (the wavelet) is examined under each of the shrinkage approaches. Given the SUREShrink variation provides the best results, the wavelet method will be compared with the local trigonometric approach (incorporating SUREShrink thresholding) in the second stage of the evaluation. If local trigonometric analysis proves superior, the method of choosing the best basis algorithm will be examined in the third stage. Specifically, an entropy-based discriminator will be compared with the SURE threshold technique. In the fourth stage, the effects of segmenting by different frame sizes are considered under what is, roughly speaking, estimated to be the: best wavelet coefficient threshold (stage 1); optimal analysis packet (stage 2); and superior best basis determination method (stage 3). Clearly, there are many other parameter combinations that could/should be evaluated. However, it is seen that the above approach does yield meaningful data although it cannot be said to provide a comprehensive foundation for obtaining many absolute conclusions. The following section presents the results of the evaluation performed under this outline.

6.2.3 Results

The results of applying shrinkage as a noise reduction technique are evaluated in two ways: visually and sonically. Although some quantitative measures can be applied, these are most revealing as a sort of after-the-fact 'debugging' tool, an idea clarified later. Visual evaluation serves as a quick and easy indicator of the general effectiveness of the procedure, in terms of noise reduction. By simply comparing pre- and post-processed signals, one can check the 'quiet' passages to see what remains following shrinkage. In the final analysis, however, only auditory assessments are meaningful. As will be demonstrated later, visual

evaluation is a poor predictor of subjective overall audio quality.

As a result of the many different parameters and associated permutations, combined with the number of signals considered, over 50 separate noise reduction efforts are presented. Most of the plots used to visually assess performance do not vary greatly across these trials. Therefore, plots are included only: to illustrate that a visual method is viable as an indication of noise reduction and in instances where they aid in interpreting subjective auditory evaluation.

6.2.4 Stage 1 - Wavelet Packet Noise Reduction

Holding the frame length and approach for finding the best basis constant, wavelet packet shrinkage is applied using different wavelet coefficient thresholds. These results are compared in order to find the best of these thresholds. A summary of the subjective auditory evaluation performed for each of the signals processed in stage 1 is summarized in Table 6.2.4.

Reviewing Table 6.2.4, results indicate that the background noise has been reduced in every trial. Visually, this is seen in reviewing Figures 6.1, 6.2, 6.3, 6.4, and 6.5, which compare the pre- and post-processed signals for each of the signals, given a SURE soft threshold as the nonlinear wavelet coefficient filter (the third set of entries in Table 6.2.4).

The plots reflect that, in general, noise is reduced. However, it is just as important to note that shrinkage has introduced a 'warbling' type of distortion (see Table 6.2.4), similar to that of sound propagating through water. These results hold for most of the signals examined in stage 1. Two interesting exceptions are: (1) the Grateful Dead recording is not enhanced by any of the above techniques and is harmed quite a bit by MINMAX and visual thresholding and (2) any difference in the McKinley passage is not audibly discernible. It also noted that the SURE and Hybrid thresholds yield nearly equivalent results, in an auditory sense. This is somewhat expected since the Hybrid threshold is designed to compensate for extreme sparsity (many non-zero values), a nonexistent condition in the signals considered. As a result, either of these filters seems to be a reasonable choice for

Table 6.1: Subjective Auditory Evaluation of Shrinkage Noise Reduction Efforts - Stage 1

SIGNAL	THRESHOLD APPLIED	COMMENTS
Caruso	MINMAX	all background noise reduced, including transients; warbling (underwater effect)
Dead	MINMAX	lots of static-like noise; not better by any measure
McKinley	MINMAX	no obvious difference
Oz	MINMAX	background noise reduced, to include the machinery; warbling (underwater effect)
Wolf	MINMAX	background noise eliminated; warbling (underwater effect)
Caruso	Visual	background noise reduced; heavy warbling
Dead	Visual	even worse than in the MINMAX case
McKinley	Visual	no obvious difference
Oz	Visual	background noise nearly eliminated, to include the machinery; warbling
Wolf	Visual	background noise eliminated; heavy warbling
Caruso	SURE	background noise reduced, better attenuation on transients; warbling
Dead	SURE	very little static-like degradation, but not improved over original
McKinley	SURE	no obvious difference
Oz	SURE	background noise reduced, machinery preserved better than with MINMAX and Visual thresholds; warbling
Wolf	SURE	background noise reduced but not as much as with MINMAX and Visual thresholds, but less warbling than with those thresholds; warbling decreased
Caruso	Hybrid	very similar to SURE case
Dead	Hybrid	very similar to SURE case
McKinley	Hybrid	very similar to SURE case
Oz	Hybrid	very similar to SURE case
Wolf	Hybrid	very similar to SURE case

Wavelet (Symmlet 8) used as the analysis function.
SURE thresholding used to determine best basis.
512-element frames used to segment signal.

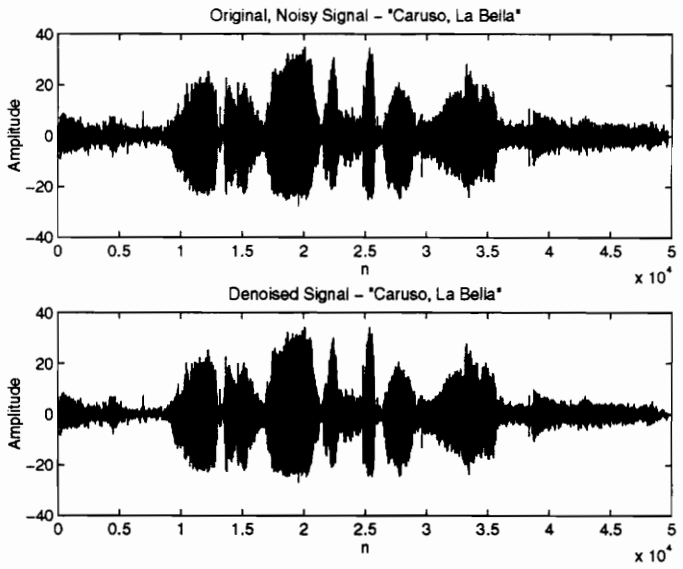


Figure 6.1: Caruso, 'La Bella' - Before and After Wavelet Shrinkage.

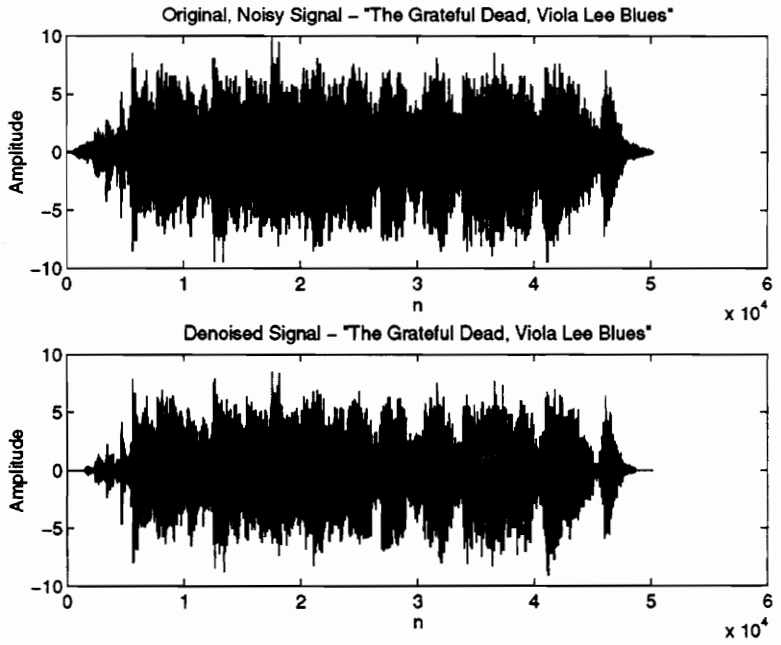


Figure 6.2: The Grateful Dead, 'Viola Lee Blues' - Before and After Wavelet Shrinkage.

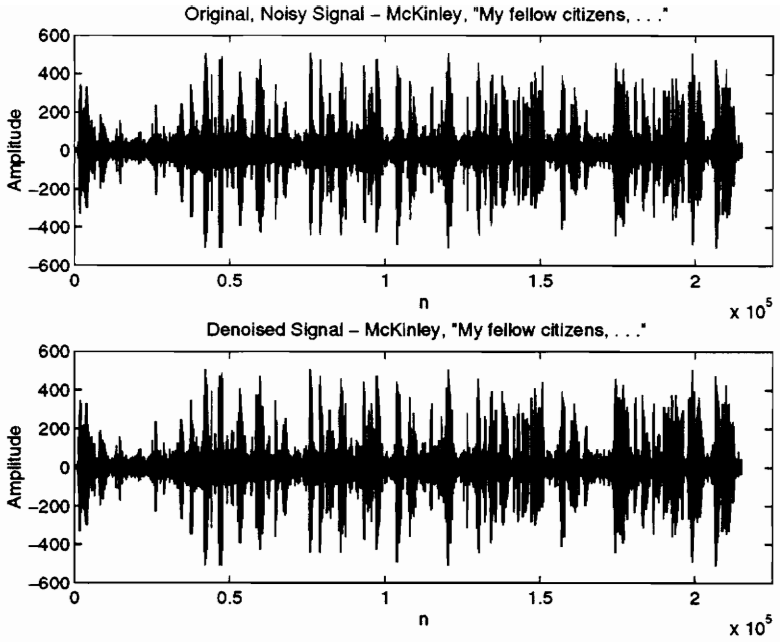


Figure 6.3: McKinley, 'My fellow citizens, . . .' - Before and After Wavelet Shrinkage.

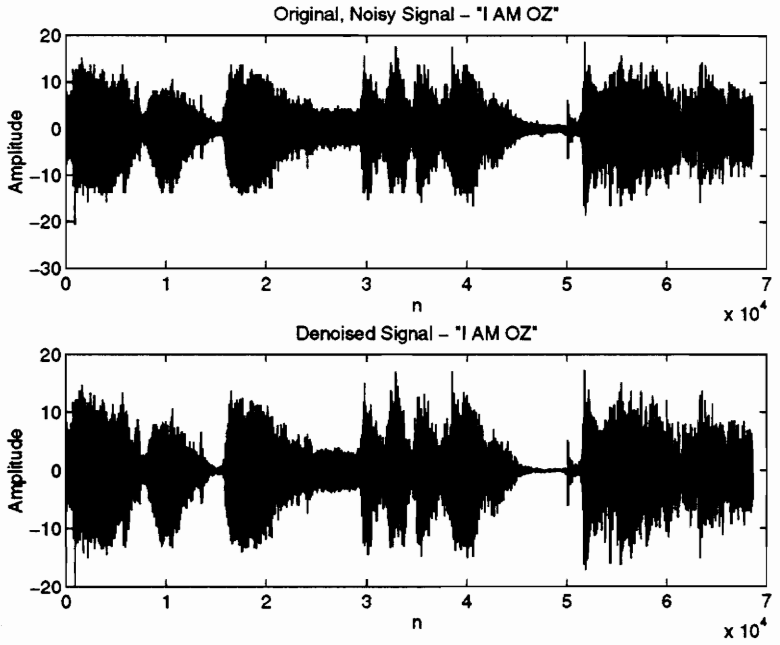


Figure 6.4: The Wizard of Oz, 'I Am Oz' - Before and After Wavelet Shrinkage.

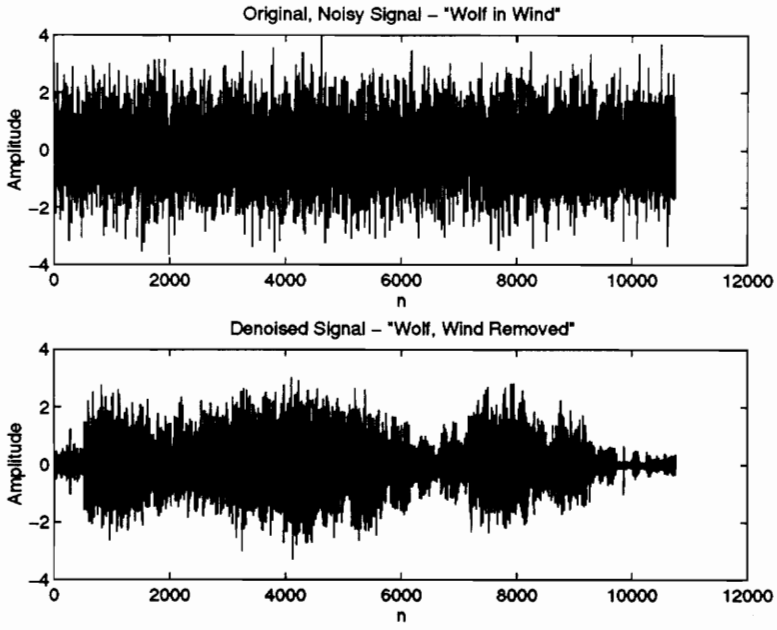


Figure 6.5: Wolf Howl - Before and After Wavelet Shrinkage.

further analysis. The SURE threshold is selected for comparing the wavelet packet and cosine packet approaches.

6.2.5 Stage 2 - Comparison of Wavelet and Cosine Packet Approaches

The parameters corresponding to the best overall case found in wavelet packet shrinkage (SURE best basis, SURE coefficient thresholding, and a 512-element frame) are maintained and a local trigonometric analyzing function substituted for the Symmlet. The results from wavelet and local trigonometric processing are then compared. A summary of the subjective auditory comparison is presented in Table 6.2.

First, the two music signals are considered. In the Caruso signal, a trade-off results. Wavelet packet shrinkage does a better job at removing the noise, particularly the transients (clicks and pops). However, this technique is more damaging to the signal in terms of the warbling distortion introduced. In the case of the Grateful Dead recording, the music is more dynamic and warbling is less evident than in the wavelet packet case.

Table 6.2: Subjective Auditory Evaluation of Shrinkage Noise Reduction Efforts - Stage 2

SIGNAL	PACKET TYPE	COMMENTS
Caruso	Wavelet (Symmlet 8)	background noise reduced, good attenuation on transients; warbling
Dead	Wavelet (Symmlet 8)	small bit of static-like degradation; no noticeable improvement over original
McKinley	Wavelet (Symmlet 8)	no obvious difference
Oz	Wavelet (Symmlet 8)	background noise reduced, machinery preserved well; warbling
Wolf	Wavelet (Symmlet 8)	background noise reduced but not as much as in other wavelet methods; warbling
Caruso	Cosine	background noise reduced and transients attenuated, but not as well as in wavelet case; warbling reduced and better than the wavelet case
Dead	Cosine	generally has removed the ambient 'dullness', some warbling but reduced
McKinley	Cosine	no obvious difference
Oz	Cosine	background noise reduced, machinery preserved better than wavelet packet case; warbling not as bad as in wavelet packet case
Wolf	Cosine	background noise more evident than wavelet packet case; warbling less evident than in wavelet packet case
SURE thresholding used to filter wavelet coefficients. SURE thresholding used to determine best basis. 512-element frames used to segment signal.		

Examining the speech processing results, it is seen that 'Oz' benefits from the local trigonometric basis. Not only is the voice more vibrant with the reduction of background noise, but the noise-like machinery is less affected in comparison with the wavelet packet processing. The McKinley speech segment remains relatively immune to shrinkage under the new basis. As in the case of the Caruso signal, the wolf howl is less distorted under cosine packet shrinkage but more noise remains after filtering.

As an illustration of the caution that should be exercised in interpreting graphical results, refer to Figure 6.6. Comparing the two post-processed signals, there is clearly more noise in the signal that has been filtered under the cosine packet basis. However, it also might be tempting to assume that the wolf howl transformed into the wavelet packet basis and subsequently filtered will sound better as it visually appears more full and closer to the original. This is not the case as the howl under cosine packet shrinkage sounds better, exhibiting less warble. It is concluded that warbling does not manifest itself in an visually obvious way.

It can be argued that the wavelet techniques perform a bit better, given that the elimination of wideband noise is the goal. However, this approach invariably produces greater distortion in the signal. It is desirable to eliminate background noise but better subjective quality is the more lofty ambition. Therefore, the cosine packet basis is selected for further efforts as it appears to be less threatening to the underlying audio, although not quite as effective in removing noise.

6.2.6 Stage 3 - Evaluating Best Basis Selection Methods

The technique used to determine the best basis has remained constant to this point. The aim now is to evaluate the entropy-based method for deciding upon the best basis and compare it to the SURE thresholding scheme used in stages 1 and 2. Given SURE thresholding of the wavelet coefficients, cosine packet analysis, and a frame size of 512, a comparison of best basis approaches is made. The results are shown in Table 6.3.

Generally speaking, the entropy-based method performs better than the SURE thresh-

Table 6.3: Subjective Auditory Evaluation of Shrinkage Noise Reduction Efforts - Stage 3

SIGNAL	BEST BASIS METHOD	COMMENTS
Caruso	SURE	background noise reduced, good attenuation on transients; some warbling
Dead	SURE	ambient 'dullness' reduced, some warbling
McKinley	SURE	no obvious difference
Oz	SURE	background noise reduced, machinery preserved well; slight amount of warbling
Wolf	SURE	background noise reduced but not as much as in other wavelet methods; warbling decreased
Caruso	Entropy	better background noise suppression than in SURE case, good attenuation on transients; same amount of warbling
Dead	Entropy	ambient 'dullness' reduced, slightly more warbling than in the SURE case
McKinley	Entropy	no obvious difference
Oz	Entropy	slightly less background noise than in SURE case; machinery preserved as well as in SURE case; same, slight amount of warbling but a more vibrant 'Oz'
Wolf	Entropy	less background noise and slightly less warbling than in SURE case

SURE thresholding used to filter wavelet coefficients.
Local cosine functions used as the analysis function.
512-element frames used to segment signal.

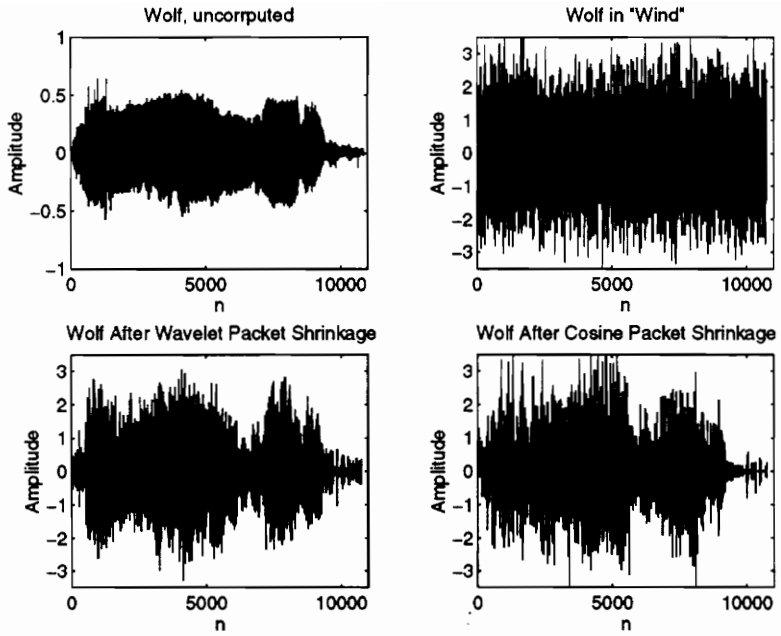


Figure 6.6: Wolf Howls (clockwise from upper left) - Original Howl, Wolf in Wind, Wolf After Wavelet Shrinkage, and Wolf after Cosine Shrinkage

olding method for selecting the best basis. This is seen in either reduced background noise, leading to more vibrant sounding audio, or less distortion. An exception is the Grateful Dead recording which is equally lively for both basis selection approaches, but suffers greater distortion under the entropy-determined basis. As a result, the SURE threshold will be employed throughout the remainder of the evaluation when considering this signal. The entropy minimization scheme will be applied to all other signals.

6.2.7 Stage 4 - Determining the Effects of Frame Size

The parameters of the algorithm have been chosen according to what appears to be the best fit for the signals analyzed, providing a foundation for assessing the effects of varying frame lengths. The comparative analysis is summarized in Table 6.4. Comments associated with the 512-element frame are made relative to the original, noisy signal. Comments regarding the other frame sizes are relative to the 512-element case.

Table 6.4: Subjective Auditory Evaluation of Shrinkage Noise Reduction Efforts - Stage 4

SIGNAL	FRAME SIZE	COMMENTS
Caruso	128	transient noise not suppressed as well as in the baseline and more warbling
Dead	128	scratchy, more warbling than in the baseline
McKinley	128	no obvious difference from the baseline
Oz	128	scratchy, not as clean as in the baseline
Wolf	128	scratchy, noisier than in the baseline
Caruso	256	no obvious difference from the baseline
Dead	256	no obvious difference from the baseline
McKinley	256	no obvious difference from the baseline
Oz	256	no obvious difference from the baseline
Wolf	256	no obvious difference from the baseline
Caruso	512	lower background noise, good attenuation on transients; small amount of warbling
Dead	512	ambient dullness reduced; warbling noticed mainly during the 'fades'
McKinley	512	no obvious difference from original
Oz	512	lower background noise, machinery well-preserved better small amount of warbling
Wolf	512	lower background noise; obvious warbling
Caruso	1024	no obvious difference from the baseline
Dead	1024	less warble than the baseline; slightly better
McKinley	1024	no obvious difference from the baseline
Oz	1024	no obvious difference from the baseline
Wolf	1024	no obvious difference from the baseline
Caruso	2048	no obvious difference from the baseline
Dead	2048	dullness reappears; better during 'fades'
McKinley	2048	no obvious difference from the baseline
Oz	2048	no obvious difference from the baseline
Wolf	2048	slightly more background noise than in baseline but a bit more realistic (less warbling)
Caruso	4096	no obvious difference from the baseline
Dead	4096	dullness more pronounced; better during 'fades'
McKinley	4096	no obvious difference from the baseline
Oz	4096	no obvious difference from the baseline
Wolf	4096	increase in background noise but possibly more realistic wolf

Considering the smallest frame case, the processed signals are all noisier than the reference (512-element frames). This result was predicted in [23], the dynamics of the desired signals cannot be fully captured within this brief interval. At 256- and 1024-element frames, little difference is noted between the associated signals and those processed in 512-element segments. As the frame size continues to increase, it is seen that the wolf howl is a little better given 2048-element processing intervals. However, for the Grateful Dead recording, the long-term averaging effect becomes significant and the dull ambience resurfaces although warbling during the 'fades' is less noticeable. In the remaining recordings, the increased frame size has sonically imperceptible effects. For 4096-element segments, the averaging effect further degrades the Grateful Dead recording and begins to adversely alter the wolf howl as well. The remaining signals continue to be essentially unaffected.

6.3 Conclusions and Extensions

Prior to discussing more general results, the McKinley speech is considered. No improvement nor damage was found in all of the attempts at noise reduction. This would seemingly suggest an underestimation of the noise in the pre-processing stage. However, subsequent efforts in which the noise was estimated from quiet passages yielded similar results or lack thereof. It may, therefore, also be possible that the signal is not as noisy as it is distorted. This conclusion is supported by the Figure 6.7 which shows the estimated noise source. This particular noise estimate was taken from the situation corresponding to cosine packet processing under entropy-driven best basis selection (see Table 6.3). However, it is noted that all noise estimates corresponding to the McKinley speech were similar in power, regardless of the type of processing.

Comparing Figure 6.7 with Figure 6.3, it is seen that the noise power is a small fraction of that in the signal. Thus noise reduction is possible, though not audibly consequential. This result suggests that the degradation in the McKinley speech may be the result of strongly-correlated distortion, possibly modulation by the original recording equipment.

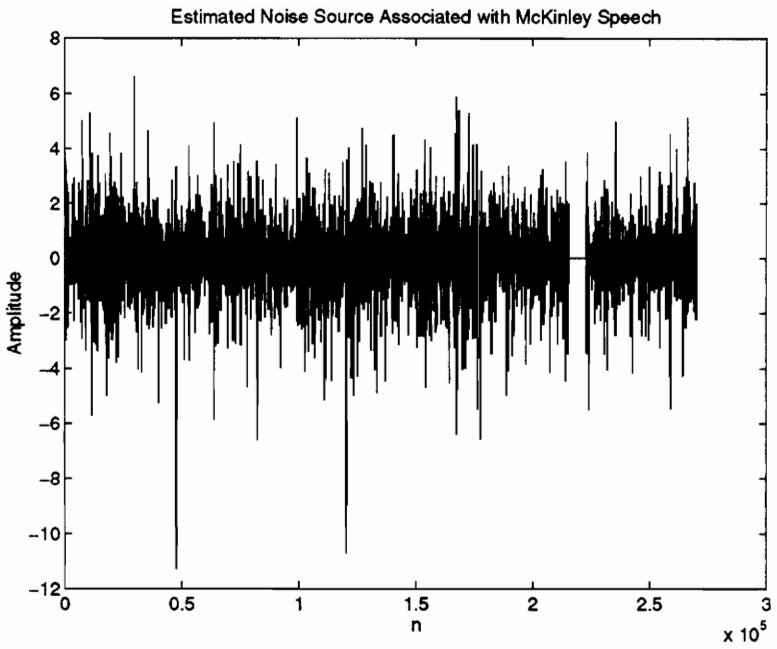


Figure 6.7: Estimate of the Noise Corrupting the McKinley Speech - 'My fellow citizens, . . . '

Although no improvement results from applying shrinkage to this signal, it may be to the credit of the technique, illustrating its ability to 'leave well enough alone.'

A closer look at the results pertaining to the Grateful Dead recording, presented in 6.2.7 above, may provide insight into the mechanics of the algorithm. In addition to the return of an overall sonic haze, as the frame size increases, there is less warble heard during the faded sections. Consider that these fading portions of the signal begin or end (depending on whether a 'fade in' or 'fade out' is under consideration) as very low-level passages. As the signal is divided into longer segments, these fades become less significant. As discussed earlier, the dynamics of any given small-duration region of the signal contribute less as the best basis algorithm is forced to estimate over many transition regions and periods of short-term stationarity. The result is something less than a series of localized best bases and consequently, reduced coherence between the signal and the packet basis.

It is emphasized that the warbling and background noise are two *very different* enti-

ties. The background noise being the original problem which is quite effectively remedied by shrinkage techniques. The warbling is an undesirable artifact introduced by shrinkage. Continuing with observations on this distortion, it is less obvious during periods of high signal energy. It is most apparent during periods of low-level signal activity, but not necessarily silence. Ironically, the better the algorithm performs the fundamental noise reduction task, the more pronounced the adverse side effect.

An attempt to explain this effect is based on the QMF pair. QMF pairs are often referred to as 'perfect reconstruction filters.' It is true that a signal can be decomposed using the subband coding approach and then reassembled by passing the filtered signal through the set of filters comprising the QMF inverse. Since the QMF pairs cannot practically meet the ideal brickwall filter specifications, nonideal realizations which overlap and intersect at $\pi/2$ are used. This leads to aliasing in the transform but even this is not a problem as it can be shown that after passing through the inverse filters, the aliasing terms have cancelled perfectly. This assumes that all of the elements of the decompostion are kept intact; in other words, have not been thresholded. The nonlinear filtering cannot remove only the noise, a bit of the signal and the alias-cancelling terms may be eliminated as well. This remaining distortion can appear as warbling, whistles, etc. This presents a serious obstacle to audio applications if quality is the ultimate goal. If intelligibility is the concern, this drawback isn't quite as serious. Despite this shortcoming, the preliminary results are promising and extensions to the method should be examined; several proposals are presented shortly.

First, however, some of the ideas discussed so far can be illustrated by viewing the autocorrelation sequence for the estimated noise signal associated with two different noise reduction approaches applied to the same signal, each yielding similar background noise levels and noticeably different degrees of warbling. In the case of lesser warbling, the autocorrelation sequence for the estimated noise signal ($signal_{\text{pre-shrinkage}} - signal_{\text{post-shrinkage}}$) should be more like that of white noise (with smaller correlations for non-zero lag). Consider the speech segment 'I Am Oz, . . .,' and the comparison of stage 2 (refer to Table 6.2. The noise environment was perceived to be similar for both cases, but the denoising performed

under cosine packet processing introduced less warble. Autocorrelations for noise estimates taken at various intervals are provided in Figures 6.8 and 6.9 which provide *weak* support for the suggestion that greater warbling results from increased aliasing. The term weak is used since there is no means of quantifying the actual noise reduction and the possibility exists that the more noiselike autocorrelations are the result of greater noise reduciton.

Improving the shrinkage technique appears to be possible. A better analyzing function is an obvious starting point. Recently, Newland [24] has proposed harmonic wavelets and an extension thereof termed musical wavelets. In short, these wavelets allow octave subdivisions into semitones aligned with the musical scale. It is strongly believed that greater coherence should result with these functions.

In terms of filtering the wavelet coefficients, it should be possible to develop a threshold that incorporates psychoacoustic principles. The results of the best basis search provides information regarding the frequency content of the signal segment under analysis. This information could be combined with critical band parameters to define segment-specific thresholds, accounting for hearing sensitivity. Ultimately, extensive listening-based adjustments would be required.

The aliasing problem mentioned earlier must also be addressed if it is indeed the source of warbling. While not mentioned above, there also exists the possibility that the primary cause of this warbling is the removal of too much harmonic content from the desired signal. It is believed with some confidence, however, that aliasing is the area of concern since high-energy regions do not sound as distorted as low-energy regions. This problem obviously requires further research. If aliasing is found to be the source of distortion, it may be possible to identify the unwanted components spectrally. Assuming that the corruption is tonal or harmonic, which is yet another area for examination, the frequency-domain representations of the pre- and post-filtered signal could be compared. Given that the original audio is not too noisy, it may be possible to determine, based on comparison, undesired contributions.

In summary, the approach examined is more than adequate in terms of noise reduction. The audible results, particularly in signals having consistently high-energy regions revealed

undeniable improvement with the audio more vibrant and dynamic. Given the many areas that are still open to examination using recent developments and the promise offered by current results, further investigation of this procedure is justified.

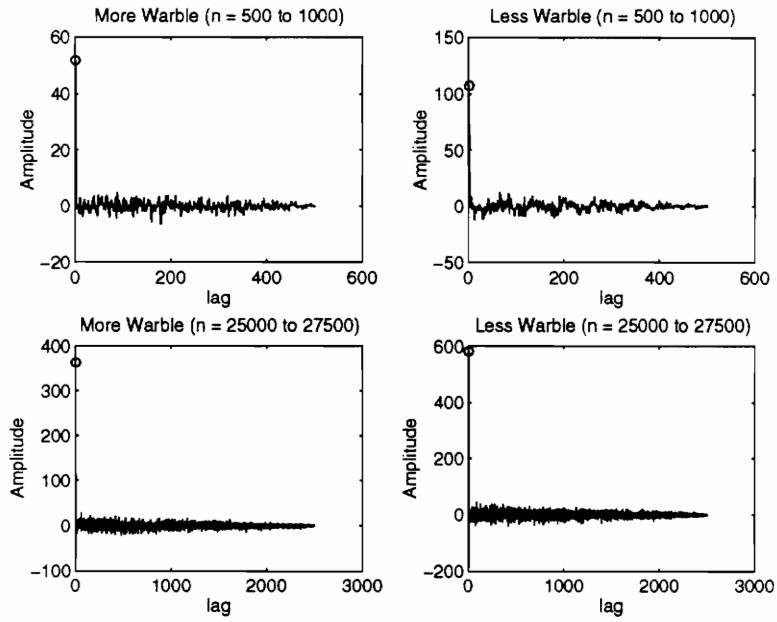


Figure 6.8: Autocorrelation Sequences of the Estimated Noise Signals Associated with Wavelet Packet (More Warble) and Cosine Packet (Less Warble) Shrinkage of 'I Am Oz, . . .'

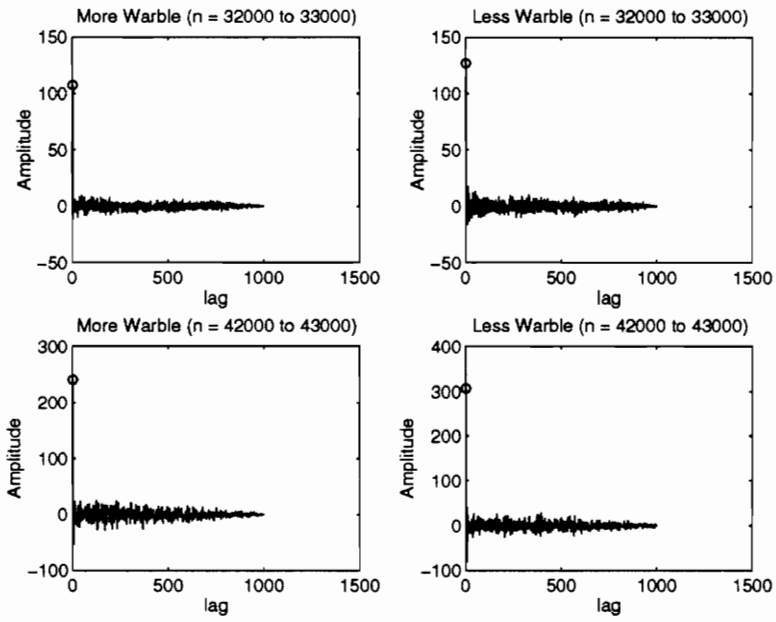


Figure 6.9: More Autocorrelation Sequences of the Estimated Noise Signals Associated with Wavelet Packet (More Warble) and Cosine Packet (Less Warble) Shrinkage of 'I Am Oz, . . .'

REFERENCES

[1] Vidakovic, B., and Muller, P., "Wavelets for Kids, A Tutorial Introduction," 1994, unpublished. Available by FTP at **ftp.isds.duke.edu** in directory **pub/Users/brani/papers/wav4kids[A,B].ps.Z** .

[2] Daubechies, I., *Ten Lectures On Wavelets*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1992.

[3] Newland, D., *Random Vibrations, Spectral & Wavelet Analysis*, Third Edition, Longman Scientific and Technical: Essex, England and John Wiley & Sons: New York, NY, 1993.

[4] Strang, G.,"Wavelet Transforms Versus Fourier Transforms," *Bulletin of the American Mathematical Society*, Vol. 28, No. 2, April 1993, pp. 288-305.

[5] Strang, G., "Wavelets and Dilation Equations: A Brief Introduction," *Society for Industrial and Applied Mathematics Review*, Vol. 31, No. 4, December 1989, pp. 614-627.

[6] Mallat, S., "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 11, 1989, pp. 674-693.

[7] Coifman, R., and Wickerhauser, V., "Entropy-Based Algorithms for Best Basis Selection," *IEEE Transactions on Information Theory - Special Issue on Wavelet Transforms and Multiresolution Signal Analysis*, Vol. 38, No. 2, March 1992, pp. 713-718.

[8] Rioul, O., and Vetterli, M., "Wavelets and Signal Processing," *IEEE Signal Processing Magazine*, October, 1991, pp. 14-37.

[9] Coifman, R., Meyer, Y., Quake, S., and Wickerhauser, V., " Signal Processing and Compression with Wavelet Packets," *Wavelet and Their Applications: Proceedings of the NATO Advanced Study Institute on Wavelets and Their Applications, 16-29 August 1992*, edited by: Byrnes, J.S., Byrnes J.L., Hargreaves, K., and Berry, K., Kluwer Academic Publishers, 1994, pp. 363-379.

[10] Dorize, C., Gram-Ilansen, K.,"On the Choice of Parameters for Time-Frequency Analysis," *Wavelets and Applications: Proceedings of the International Conference, May 1989*, edited by Meyer, Yves, Masson/Springer-Verlag, 1991, pp. 86-91.

[11] Gaps, A., "An Introduction to Wavelets," *IEEE Computational Science and Engineering, Summer 1995*, Vol. 2, No. 2, 1995.

[12] Zwicker, E., and Fastl, H., *Psychoacoustics: Facts and Models*, Springer-Verlag, 1990, pp. 133-134.

[13] Lu, J., Weaver, J., Healy, D., and Xu, Y., "Noise Reduction with Multiscale Edge Representation and Perceptual Criteria," *Proceedings of IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, October, 1992.*

[14] Donoho, D. and Johnstone, I., "Ideal Spatial Adaptation by Wavelet Shrinkage," April, 1993, unpublished. Available by FTP at **playfair.stanford.edu** in directory **/pub/donoho/isaws.ps.Z** .

[15] Donoho, D. and Johnstone, I., "Adapting to Unknown Smoothness via Wavelet Shrinkage," July 20, 1994, unpublished. Available by FTP at **playfair.stanford.edu** in directory **/pub/donoho/ausws.ps.Z** .

[16] Donoho, D., "Nonlinear Wavelet Methods for Recovery of Signals, Densities, and Spectra from Indirect and Noisy Data," *Proceedings of Symposia in Applied Mathematics*, Volume 00, 1993, pp. 173-205.

[17] Tewfik, A., Sinha, D., and Jorgensen, P., "On the Optimal Choice of a Wavelet for Signal Representation," *IEEE Transactions on Information Theory*, Vol. 38, No. 2, March, 1992, pp. 747-761.

[18] Xie, M., and Beex, A. A., "Optimal Wavelet Design via Multiresolution Parametric Spectral Estimation," *The International Society for Optical Engineering (SPIE) 1995 Symposium on OE/Aerospace Sensing and Dual Use Photonics, 17-21 April 1995, Vol. 2491-59 Wavelet Applications*, edited by: Szu, H., April 1995, pp. 627-638.

[19] Teolis, A., and Bendetto, J., "Local Frames," *Mathematical Imaging: Wavelet Applications in Signal and Image Processing: Proceedings of the International Society for Optical Engineering, July 15-16, 1993*, Vol. 2034, 1993, pp. 310-321.

[20] Stein, C., "Estimation of the Mean of a Multivariate Normal Distribution," *Annals of Statistics*, Vol. 9, No. 6, 1981, pp. 1135-1151.

[21] Jawerth, B., Liu, Y., and Sweldens, W., "Signal Compression with Smooth Local Trigonometric Bases," undated, unpublished. Available by FTP at **ftp.math.scarolina.edu** in directory **/pub/wavelets/varia/opteng.ps.gz** .

[22] Kolaczyk, E., "Wavelet Methods for the Inversion of Certain Homogeneous Linear Operators in the Presence of Noisy Data - Ph.D. dissertation submitted to the Department of Statistics, Stanford University," October 1994. Available by FTP at **galton.uchicago.edu** in directory **/pub/kolaczyk/Thesis/KolaczykThesis_Text.ps.Z** .

[23] Berger, J., Coifman, R., and Goldberg, M., "Removing Noise from Music Using Local Trignonmetric Bases and Wavelet Packets," *Journal of the Audio Engineering Society*, Vol. 42, No. 10, October 1994, pp. 808-818.

[24] Newland, D., "Harmonic and Musical Wavelets," *Proceedings of the Royal Society of London*, Volume 444, 1994, pp. 605-620.

VITA

K. ALLEN COMER

1776 Liberty Lane
Apartment C-26
Blacksburg, VA 24060

Phone: (540) 961-6138
email: comerk@birch.ee.vt.edu

EDUCATION

M.S., Electrical Engineering
Virginia Polytechnic Institute and State University ,
Blacksburg, VA (expected May 1996)
Thesis - 'A Wavelet-Based Technique for Reducing Noise in Audio Signals'
GPA: 3.7/4.0

B.S., Electrical Engineering, 1987
University of Virginia, Charlottesville, VA
Undergraduate thesis - 'The Digital Audio System'

RELATED COURSES & PROJECTS

Digital Signal Processing (DSP)	Speech Processing	Spectral Estimation
Microprocessor System Design	Pattern Recognition	C++
Stochastic Processes I & II	Linear Systems	Communication Circuits

- Examined popular techniques and algorithms used in speech signal processing such as hidden Markov models, dynamic time warping, and neural networks. Employed autoregressive models for formant detection and implemented other techniques (such as cepstral analysis) for pitch estimation.

- Designed lattice filters, multirate processing systems, and a homomorphic filtering system. Additionally, performed an extensive analysis of finite wordlength effects in DSP systems.

- Designed and implemented digital state-space realizations of Wiener and Kalman filters.

- Worked with the Texas Instruments TMS32010 and Motorola 56002 DSP chips. Specifically, implemented digital filters and fast Fourier transforms (FFTs) using these devices.

- Designed a music notation system using the Motorola MC68HC11 by interfacing the microcontroller to a MIDI keyboard. The system detected and converted raw MIDI data into a displayable, user-recognizable representation.

- Designed identification and prediction systems using various artificial neural network structures.

HARDWARE AND SOFTWARE

Hardware

DSP and Microcontrollers - Motorola 56002, MC68HC11, TI TMS32010, Intel 8085
Computers - IBM PCs, Macintosh PCs, Sun Sparcstations, and HP Apollo Workstations

HARDWARE AND SOFTWARE (continued)

Software and Programming Languages

C/C++	68HC11 assembly	8085 assembly	MATLAB
DFDP3/plus	UNIX	DOS/MS-Windows	LaTeX

* various DSP design packages for general PC use
* various office automation packages (MS-Word, Excel, Parado x)

WORK EXPERIENCE

Graduate Teaching Assistant - Virginia Tech, Blacksburg, current

- Responsible for grading papers and assisting professors teaching an undergraduate communication systems course.
- Responsible for assisting students in understanding the material presented in the aforementioned course.
- Responsible for generating simulations to demonstrate the behavior of systems covered in lecture material.

Electrical Engineer - Computer Sciences Corporation, Falls Church, VA, 8/93 - 5/94

- Prepared Sections B, C, L, and M of an RFP used to solicit bidders for a federal Government software development contract.
- Prepared technical documentation justifying various aspects of the hardware procurement component of the solicitation.
- Performed a failure rate analysis and assisted in the development of a maintenance schedule for computer equipment.

Electrical Engineer - EER Systems, Vienna, VA, 9/88 - 8/93

- Developed a test plan and administered a user test on software ta rgeted for national use.
- Prepared a research paper outlining procedures for developing battlefield-survivable medical equipment.
- Performed various computer programming tasks and hardware evaluations.

Systems Analyst/Project Manager - Integrated Research, Inc., Clifton, VA, 4/88 - 9/88.

- Co-founded company and co-wrote proposal used to win a contract with the U.S. Department of Transportation (DOT).
- Served as project manager and co-wrote software used to track grants administered by the Urban Mass Transit Authority of the DOT.
- Developed all user documentation.

PUBLICATIONS

"Nuclear and NBC Contamination Survivability of Medical Materiel, " Comer, K. A., Wright, W., and Zajac, A., March 1990, DTIC Reference No. 88-219-869.

REFERENCES

Available upon request.