NPS-EC-97-001

NAVAL POSTGRADUATE SCHOOL
Monterey, California

Classification of Ocean Acoustic Data Using AR Modeling and Wavelet Transforms

by

M.P. Fargues
R. Bennett
R.J. Barsanti

January 1997

Approved for public release; distribution is unlimited.

Prepared for: Naval Undersea Warfare Center
Code 33
Newport, RI 02841

NAVAL POSTGRADUATE SCHOOL
Monterey, California

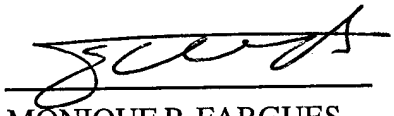
Rear Admiral M.J. Evans
Superintendent

R. Elster
Provost

This report was prepared for and funded by NUWC-Newport, Code 33.

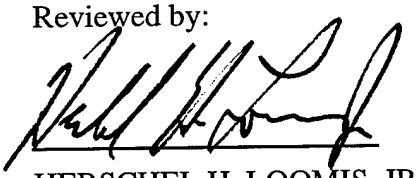
Approved for public release; distribution is unlimited.

The report was prepared by:



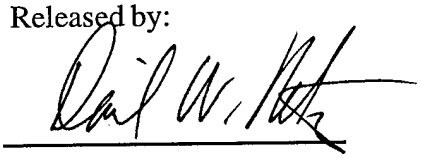
MONIQUE P. FARGUES
Associate Professor
Department of Electrical and
Computer Engineering

Reviewed by:



HERSCHEL H. LOOMIS, JR.
Chairman
Department of Electrical and
Computer Engineering

Released by:



DAVID W. NETZER
Associate Provost and
Dean of Research

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	January 1997	December 1994-June 1996

4. TITLE AND SUBTITLE
Classification of Ocean Acoustic Data Using AR Modeling and Wavelet Transforms

5. FUNDING NUMBERS
N0002495WR10820

6. AUTHOR(S)
M. P. Fargues, R. Bennett, R. J. Barsanti

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Department of Electrical and Computer Engineering
Naval Postgraduate School
Monterey, CA 93943-5000

8. PERFORMING ORGANIZATION REPORT NUMBER
NPS-EC-97-001

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Naval Undersea Warfare Center
Newport Division
Attn: Mr. Ed Jensen
Newport, RI 02841

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
The views expressed in this report are those of the author and do not reflect the official policy or position of the Department of Defense or the United States Government.

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

This study investigates the application of orthogonal, non-orthogonal wavelet-based procedures, and AR modeling as feature extraction techniques to classify several classes of underwater signals consisting of sperm whale, killer whale, gray whale, pilot whale, humpback whale, and underwater earthquake data. A two-hidden-layer back-propagation neural network is used for the classification procedure. Performance obtained using the two wavelet-based schemes are compared with those obtained using reduced-rank AR modeling tools. Results show that the non-orthogonal undecimated A-trous implementation with multiple voices leads to the highest classification rate of 96.7%.

14. SUBJECT TERMS
wavelet transform, AR modeling, classification

15. NUMBER OF PAGES
42

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	SAR

Table of Contents

List of Figures

1. Introduction

Automatic sound identification is one of the major goals of underwater acoustics. Quieting techniques have greatly reduced the principal sources of acoustic energy used for detection and classification by passive sonar. However, short duration transient signals may be used to detect and classify underwater sources. The success of any classification scheme depends to a large extent on the specific preprocessing techniques used to extract information regarding the features of the various classes of signal under study. This study explores modeling and wavelet decompositions as feature extraction techniques applied to underwater signals. The AR technique chosen uses the reduced-rank covariance method which combines the traditional covariance method with the singular value decomposition to reduce the effect of additive noise in the signal. Two implementations of the wavelet transform are considered in the study: the decimated orthonormal wavelet transform and the non-orthonormal A-Trous decomposition. Feature vectors obtained from the 3 types of decompositions considered in this study are used as inputs to a two hidden layer back- propagation network, and the resulting performances compared.

Section 2 describes the various underwater signals selected for our study. Section 3 reviews the reduced-rank AR covariance method. Section 4 presents results obtained by applying an ALE filter to denoise the data under considered. Section 5 introduces the wavelet transforms considered in this work. Classification results are presented in Section 6. Finally, Section 7 presents conclusions and suggestions for further research .

2. Signals Description

The recordings used in our study were of real, open ocean encounters from various signal collection platforms. The signals, as an artefact of the collection procedures, were all corrupted with background noises, which included sounds from ships, small boats, and other disturbances occurring in the natural environment, plus artificial noise from the means of collection. Six different classes of signals were selected for our study:

- Sperm whale,
- Killer whale,
- Humpback whale,
- Gray whale,
- Pilot whale,
- Underwater earthquake.

Each recording varied in length between fifteen to thirty seconds. Each signal was digitized on a 486 PC using a Media Vision Pro Audio Spectrum sound card with a sampling frequency equal to 8kHz, using a single channel and 8 bits per sample. Figures 2.1 to 2.7 present typical time-domain traces and spectrograms obtained from the various classes of signals considered. Four out of five classes of underwater biological signals where cuts from what is commonly known as "whale songs" and where narrowband in nature, while sperm whale and earthquake recordings of a wider range of

frequencies as compared to other types of biological data. The sperm whale recordings were of the animal' s echo ranging sonar, and constituted of very short and rapid wideband pulses.

3. Reduced-Rank AR Modeling

AutoRegressive (AR) modeling is a time-domain technique used for modeling a set of data as the output of an all-pole Linear Time-Invariant (LTI) filter. Estimation of the filter coefficients may be carried out in a least squares sense by solving the Yule-Walker equations [1,10]. Degradations due to noise may be decreased by using a truncated inverse of the data matrix defined in the Yule Walker equations to solve for the AR coefficients. Such a truncated inverse is computed using the Singular Value Decomposition (SVD), and this approach has been used extensively in signal processing applications. This section briefly reviews the concept of AR modeling and the reduced-rank AR modeling method used in the study.

3.a Autoregressive Modeling

Autoregressive (AR) modeling is based on the idea that a signal *x(n)* can be expressed as the output of an all-pole linear shift invariant filter driven by white noise. Thus, *x(n)*is given by the following expression:

$$x(n)=-\sum_{k=1}^{p} a(k)x(n-k)+b_0 w(n), \tag{3.1}$$

where **P** is the order of the predictor, b_0 is the noise standard deviation, and $(a(1),....,a(P))$ are the coefficients of the linear predictor to be determined. The resulting transfer function of the system used to generate *x(n)* from the white noise input is given by taking the Z-transform of Eq. (3.1):

$$H(z)=\frac{X(z)}{W(z)}=\frac{b_0}{1+a_1 z^{-1}+\ldots+a_P z^{-P}}. \tag{3.2}$$

The correlation function $R_x(k)$ can be obtained from *x(n)*, which leads to:

$$R_x(k)=-a_1 R_x(l-1)-a_2 R_x(l-2)-....-a_p R_x(k-P)+b_0 R_{wx}(k). \tag{3.3}$$

The cross-correlation $R_{wx}(k)$ can be expressed as the convolution of the impulse response *h(n)* of the AR system with the autocorrelation of the noise input, which leads to the following expression:

$$R_{wx}(k)=\sigma_w^2 h^*(-k). \tag{3.4}$$

Recall that *h(n)* is the impulse response of a causal filter, therefore *h(n)* is non zero for positive lags only. In addition, using the Initial Value Theorem, leads to:

$$h(0)=\lim_{z\to\infty}H(z)=b_0. \tag{3.5}$$

Thus, Eq. (3.3) becomes:

$$R_x(k)+a_1R_x(k-1)+....a_pR_x(k-P)=|b_0|^2\sigma^2{}_w\delta(k). \tag{3.6}$$

Expressing Eq. (3.5) for $k=1,...,P$ leads to the set of Yule Walker equations:

$$\begin{bmatrix} R_x(0) & R_x(-1) & \dots & R_x(-P) \\ R_x(1) & R_x(0) & \dots & R_x(-P+1) \\ \vdots & \vdots & \ddots & \vdots \\ R_x(P) & R_x(P-1) & \dots & R_x(0) \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ \vdots \\ a_P \end{bmatrix} = \begin{bmatrix} \sigma_w^2|b_0|^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{3.7}$$

The set of AR coefficients can then be derived by solving the above matrix equation. In practical applications the correlation function is estimated from the observed data, and various estimation procedures have been considered [10]. This study uses the covariance approach to estimate the correlation lags as this procedure makes no assumption about the data outside the windows of interest. Thus, the estimated correlation function is obtained by the following computation:

$$R_x(k)=\frac{1}{N-P}\sum_{n=P}^{N-1}x^*(n-k)x(n),$$

where N represents the length of the window used for the correlation lag estimates. The spectrum of the modeled signal obtained using the AR coefficients is given by:

$$S_x(e^{j\omega})=\frac{\sigma_w^2|b_0|^2}{|A(e^{j\omega})|^2}.$$

3.b Model Order Selection

Selecting the order of an AR model is a difficult task, as the best choice is usually not known, and trial and error are sometimes used. If the data is truly described by a finite order AR model, theoretically the variance should become constant once the model order is reached. In practice this is not usually true for a variety of reasons. Therefore, several criteria have been developed to address this problem. The four most well known are: Akaike's Information theoretic Criterion (AIC), Parzen's criterion of Autoregressive transfer (CAT), final prediction error (FPE), and Schwartz and Rissanen's minimum description length (MDL) criterion [11]. All four procedures estimate the "best"

model as that obtained at the minimum of the specified criterion function. The sperm whale data was used to set the AR model order because it is the data with the broadest bandwidth of the signals considered, and results are shown in Figure 3.1. Results indicate some variation in the estimated "best" model order, (best order obtained with AIC is 27, with MDL is 22, with CAT is 26, and with FPE is 26). As a result, we chose an AR model order equal to 25.

3.c Reduced-Rank Method

The main idea behind the reduced-rank method is to compute a truncated inverse of Eq. (3.7) using the Singular Value Decomposition (SVD). This process separates the contribution due to the noise only from that due to the signal-plus-noise data, thereby, improving the quality of the estimated AR coefficients, by stabilizing the inverse [10]. The reduced-rank (i.e., the rank of the truncated inverse obtained using the SVD decomposition) was chosen by selecting a visual gap in the singular values distribution of the data correlation matrix. Figures 3.2 to 3.7 illustrate typical singular value distributions obtained for the data considered in our study. The estimated reduced-rank varied between 2 and 17 for the data under study, where the smallest ranks was found for underwater earthquake and humpback data and the highest was found for sperm whale data, as listed in Table 3.1 below. Note that the sperm whale required the highest rank, as it is the signal with the broadest bandwidth among all signals considered.

Signals	average number of singular values retained
Pilot whale	12
Killer whale	10
Sperm whale	17
Gray whale	15
Humpback whale	2
Earthquake	2

Table 3.1. Typical number of singular values selected for retention for each class of signal.

4. Adaptive Noise Filtering

Some of the underwater signals under study were buried in noise, and we attempted to decrease the effect due to wideband noise by applying an adaptive line enhancement (ALE) pre-

processing step before computing the AR parameters. The ALE filter is designed to separate narrowband from wideband signals based on the Least Mean Square (LMS) algorithm [12]. However, this pre-processing step was successful only on four classes of underwater signals (those which were the most narrowband in nature), while it performed poorly when applied to the other two more wideband signals. As a result, the overall classification rates did not improve significantly when applying this pre-processing step, and it was not pursued further in this study. Several alternatives for denoising the data are possible. A follow-on study investigates the application of wavelet-based techniques to denoise the data, see [15] for further details.

5. Wavelet Transformations

5.a Introduction

Wavelet transforms have numerous applications in signal processing, such as coding, image processing, compression, and classification, and numerous references are available [2,4,5,6]. In our study we are interested in extracting a compact (i.e., "small") set of feature coefficients which can be used to classify the different signals with a high level of accuracy (i.e., over 90% recognition rate). In addition, we expect our classification procedure to be relatively non-sensitive to time synchronization issues. Figures 2.1 to 2.7 show that the signals under study are non-stationary and vary in frequency content, magnitude, and background noise. Spectrograms (which represent the magnitude of the short-time Fourier Transform) have been used extensively to extract information from such time-varying signals, as they are easy to interpret.

The Continuous Wavelet Transform (CWT) is best understood as an extension of the Short-Time Fourier Transform (STFT), where the signal is decomposed using sinusoidal basis functions. The STFT Transform of the signal x(t) is given by:

$$S(t,f)=\int_{-\infty}^{\infty} x(\tau)g^{*}(\tau-t)e^{-j2\pi f\tau}d\tau, \qquad (5.1)$$

and displays the evolution of the signal frequency over time. Many different window functions $g(t)$ may be selected, and the choice will affect the resolution of the resulting transform. In all cases, the time-frequency resolution of the STFT is limited by the uncertainty principle which states that:

$$\Delta t\Delta f \geq \frac{1}{4\pi}.$$

The choice of the time window $g(t)$ fixes Δt, and thus fixes Δf over the whole transformation. Therefore, the STFT cannot provide both good time resolution (which requires short time windows), and good frequency resolution (which requires long time windows). The CWT provides an attractive alternative to the spectrogram information as it decomposes the signal using a basis, where the time and frequency resolution vary, thereby allowing for variable time and frequency resolution, while keeping the time-frequency resolution product fixed.

The Continuous Wavelet Transform (CWT) of a signal $x(t)$ is given by:

$$W_x(t,a)=\frac{1}{\sqrt{a}}\int_{-\infty}^{\infty} x(\tau)\Psi^*(\frac{\tau-t}{a})d\tau, \tag{5.2}$$

where $\Psi(t)$ is the mother wavelet, which must satisfy specific mathematical properties [13]. The parameter t denotes translation in time, and the scale factor a denotes dilation in time. The factor $1/\sqrt{a}$ normalizes the energy of the CWT. Two important characteristics of wavelets are that; 1) the wavelet function $\Psi(t)$ be of finite duration, and 2) the wavelet function $\Psi(t)$ have zero average value (like that of Fourier sinusoids). The second characteristic requires that the basis functions oscillate above and below zero, and gives rise to the name wavelet or small wave [15]. Although there are numerous functions that meet the necessary properties to be classified a wavelet only a few classes have thus far been shown to be of general interest in signal processing. The Haar, Daubechies, Coiflet, and Symmlet are a few of the more popular classes and are shown in Figure 5.1. Various bases were originally considered in our study. However, we decided to use the Symmlet-8 and Coiflet-3 bases because they were among those readily available to use from [14], and they did not significantly fail on any of the classes of signals considered here. The scale factor a in wavelet analysis plays an analogous role to inverse frequency in Fourier analysis, and it controls the time and frequency resolution of the transform. Thus, as a decreases, the wavelet function $\Psi(t/a)$ becomes more concentrated in the time domain, and thus more expanded in the frequency domain. Similarly, as a increases, the wavelet function $\Psi(t/a)$ becomes more expanded in the time domain, and thus more concentrated in the frequency domain. Figure 5.2 illustrates this behavior when using a Symmlet-8 mother wavelet. The magnitude of the WT called the scalogram, in analogy with the spectrogram., is a representation of the signal energy in the time-scale plane. The scalogram has high time resolution at high frequency and high frequency resolution at low frequency, as illustrated in Figure 5.3. Further insight to the multiresolution capability of the CWT can be gained by comparing the influence of signals in the time - scale plane. Figure 5.4 shows a comparison of the regions of influence of the spectrogram and scalogram for two different signals. The top plots display an impulse function at $t = t_o$. Note that the scalogram permits a narrow time localization of this signal in the low scale portion of the plot. The lower plots display the regions of influence for a signal composed of two sines at frequencies f1 and f2. Note the CWT has better frequency resolution at high scales and poorer frequency resolution at low scales.

5.b The Discrete Wavelet Transform

The Discrete Wavelet Transform (DWT) is defined by discretizing the parameters t and a of the CWT. by:

$$W_x(a,b)=\sum_{n=1}^{N} \frac{1}{\sqrt{a}}\, x(n)\, \Psi^*(\frac{n-b}{a}), \tag{5.3}$$

where a, b, n are the discrete versions of a, t, *and* τ, of Eq. (5.2) respectively. The scaling factor a

is further restricted to be given by:

$$a = a_o^J \qquad J = 0,1..... \tag{5.4}$$

The choice of a_0 will govern the accuracy of the signal reconstruction via the inverse transform. It is popular to choose $a_0 = 2$ because it provides small reconstruction errors and permits for the implementation of fast algorithms [13]. Setting $a = 2^J$ produces octave bands called dyadic scales. At each scale as J increases, the analysis wavelet is stretched in the time domain, and compressed in the frequency domain by a factor of two, as shown in Figure 5.2. As a result, the DWT output at each dyadic scale J produces more precise frequency resolution and less precise time resolution. Also note that as J increases the translation term $b/2^J$ becomes smaller, and thus b must necessarily increase to cover all translations. The result is that the DWT output grows in length by a factor of two at every scale. This produces extremely large DWT vectors at the higher scales. This computational difficulty can be alleviated by realizing that at each successive octave, the DWT output contains information at half the bandwidth compared to that of the previous scale, and thus can be sampled at half the rate according to Nyquist's rule [14]. This decimation (or subsampling) is accomplished mathematically by restricting values of the shift parameter b. Letting $b = k \cdot 2^J$ where k is an integer, and replacing a by 2^J yields the decimated DWT given by:

$$W_x(2^J, k2^J) = \sum_{n=1}^{N} \frac{1}{\sqrt{a}} x(n)\ \Psi^*(2^{-J}n - k), \tag{5.6}$$

where $J = 0,.....\log_2(N)$ and $k = 1,.....N \cdot 2^{-J}$. The term $k \cdot 2^J$ in the argument of the DWT, indicates that $W_x(a,b)$ is decimated by a factor of two at each successive scale J by retaining only the even points. The resulting DWT coefficients form a [J x k] matrix where each element represents the similarity between the signal and the analysis wavelet at scale j and shift k. It is common practice therefore to rewrite Equation (5.6) explicitly in terms of the parameters j and k, leading us to the decimated DWT equation defined as:

$$W_{j,k} = \sum_n \sqrt{1/2}^J\ x(n)\ \Psi^*(2^{-J}n - k). \tag{5.7}$$

The Symmlet-8 wavelet is shown at various scales j and shifts k in Figure 5.5.

An efficient way to implement the DWT of Eq. (5.7) using filters was developed by Mallat [13,15]. This scheme uses a complementary pair of lowpass (LP) and highpass (HP) filters. These filters equally partition the frequency axis and are known as quadrature mirror filters (QMF) [13]. Since each filter output covers only half the original frequency range of the input, each can be decimated by a factor of two by retaining only the even points. The combined decimated output of the two filters is a data set which comprise the DWT coefficients at the first scale. This process is repeated on the LP filter output to produce further decomposition of the signal into LPHP and LPLP parts at the next scale. The filtering and decimating operations can be continued until the number of samples is reduced to two. At each successive iteration (scale) the frequency range of the output is reduced in half by the LP filter, and the frequency resolution is improved by the decimation. Figure 5.6 shows how a data set of 2^J samples can be decomposed to produce a maximum of j levels of

transform coefficients. Figure 5.7, displays the resulting transformed coefficients in a tree structure. Note that movement down the tree relates to lower frequency (higher scale) coefficients.

The decimated DWT described above will produce an orthogonal decomposition of the input signal only if the QMF pairs (i.e., the wavelets) are properly chosen. Such filter pairs have to possess specific mathematical properties and exhibit restrictive symmetry characteristics [13]. Although the DWT filtering operations are linear and time invariant, the decimation combined with the filtering results in a time-variant system. Recall, that a time variant system implies that shifts in the system input will not produce an equivalent shift in the system output [13]. In fact, a shift of even a few samples in the signals starting point can completely change the wavelet decomposition coefficients. This difficulty complicates the performance of signal detection, feature extraction, and classification in the wavelet transform domain [14,15], and a number of proposals have been made to deal with the time - variant nature of the wavelet transform [15].

A non-orthogonal transform was also considered in our study, as such transforms may have advantages in applications where the redundancy makes the information easier to extract [2,4]. The non-orthogonal transform considered was the undecimated A-trous implementation of the WT using a Morlet-type mother wavelet, as introduced in Shensa [4].The Morlet wavelet is given by:

$$\Psi(t)=\exp(jvt)\exp(-\beta^2 t^2/2),$$

where β and ν respectively represent the roll-off factor and center frequency of the wavelet [4]. This undecimated transform has the additional advantages of being translation invariant. In addition, the user may vary the spectral partitioning by changing the number of voices per octave, where voices can be viewed as sub-band filters defined within a given scale, while it is fixed by the choice of basis in the orthonormal decomposition [4]. Figures 5.8 to 5.11 respectively display the spectral partitioning obtained with Coiflet-3, Symmlet-8, and the A-Trous decomposition with four and five voices for a four scale decomposition, for the following Morlet wavelet parameters used in this study; rolloff parameter β=.15 and center frequency η=.85π. Note that the A-Trous decomposition allows for much narrowband frequency partitioning than the orthonormal decompositions do.

5.c Feature Extraction

We are interested in keeping the sets used to describe each of the different classes as compact as possible, and to avoid time synchronization problems between the different signals investigated. Lemer et. al. showed in a preliminary study that using energy quantities based on Daubechies wavelets of order 6 improved performances for the specific signals they considered [3]. Expanding on this idea, we defined energy-type parameters from the wavelet coefficients and used these quantities as feature parameters for the classification scheme. When using orthogonal transforms, we defined the average energy E_i computed from the wavelet coefficients obtained at a given scale I for scales 1 to 7, and the complementary average energy contained in the low-pass operation. Thus, average wavelet-based quantities E_i for scale I used as feature parameters were defined as:

$$E_i = \frac{1}{2^i}\sum_k c_{i,k}^2 , \quad i=1,...,7,$$

where $c_{i,k}$ represents the kth wavelet coefficient obtained at scale I, and the summation operation is done over all wavelet coefficients available at a given scale. This study considers scales I from 1 to 7. A similar expression is used to derive the 2nd set of coefficients from the low-pass operations in the same range of scales. As a result, a seven scale decomposition leads to a set of 14 real energy-type feature coefficients. Such a choice insures to keep the number of feature coefficients low, and avoids potential problems dealing with time-domain synchronization. Two orthonormal bases are used in the study and their performances compared; Coiflet-3 and Symmlet-8 bases [1].

When using the non-orthogonal transform, the feature parameters chosen for the classification scheme are the set of average wavelet-based coefficients $A_{i,j}$ obtained at a given scale I and voice j, where $A_{i,j}$ is defined as:

$$A_{i,j}=\frac{1}{2^i}\sum_k |c_{i,j,k}|^2, \quad i=1,...,7, \quad j=0,...,m-1.$$

The parameter $c_{i,j,k}$ represents the kth wavelet coefficient obtained at scale I and voice j, and the summation operation is conducted over the range of wavelets coefficients obtained at a given scale and voice. Thus, a seven scale decomposition using m voices leads to $7m$ input coefficients. Several configurations of the A-trous implementation are investigated in the study, using between 4 to 7 voices. Experiments were conducted using the roll-off parameter β=.15 and the center frequency η=.85π. Results showed that the best overall classification results among the various implementations considered were obtained when using six voices per scale, as illustrated in Table 6.1.

6. Classification

6.a Network Architecture

A back-propagation neural network configuration is used in the study. Back-propagation networks are multilayer feedforward networks, which learn during supervised training sessions, where input feature vectors have target outputs. The number of input and output elements in the network is usually equal to the number of different classes under investigation. Theoretically, there should be only two possible values for each output of the network; either a "1" or a "0". Therefore, the ideal output level for all outputs should be all zero except for the output corresponding to the correct class, which should be equal to 1. In practice, the actual output levels may vary between "0" and "1".

Learning actually take place when input vectors are propagated through the network in a forward direction on a layer-by-layer basis to the output layer. The output layer is compared to the target classification and the error is back-propagated through the network layer by layer, neuron by

neuron, updating the connection weights which contain the memory of the system. Once the network converges on a stopping criterion, the weights become fixed and the network can be used for testing. The NN implementation used in this study used hyperbolic tangent for transfer function, and the normalized-cumulative delta learning rule to update the learning coefficients. In addition, avoiding saturation of the transfer function is handled by scaling the input values in the range of ±2 [9]. The number of Processing Elements (PEs) in the hidden layers, and the number of hidden layers are important decisions in NN architecture. Most back-propagation networks will have one or two hidden layers with the number of PEs in the hidden layers falling in between the number of input values and the number of PEs. The number of PEs depends on the complexity of the relationships between different classes, as signals that are not easily separated may require more PEs to distinguish between them. A common rule of thumb to estimate the number of hidden layers needed by a back-propagation network is:

$$h = \frac{number\ of\ training\ files}{5(m+n)},$$

where:

his the number of PEs in the hidden layer,
m is the number of PEs in the output layer,
n is the number of PEs in the input layer. [9]

However, applying the above rule led to networks which did not converge in any reasonable time for our data. The architecture of the network which constantly converged in a reasonable time frame for this study included a first hidden layer with the number of PEs close to the number of inputs, followed by a second hidden layer with 15 PEs, and six output elements.

6.b Classification Rates

This study uses classification rates as the overall measure of performance for the network. The idea behind the classification rate is for the network to pick a winner, which is simply the output PE with the largest value. Thus, if we compare the winner with the target we have a binary yes-or-no answer for correctness in classification. The neural network software used in this study, NeuralWorks II/ Professional Plus [9] has such an instrument built in the algorithm. The NeuralWorks classification rate instrument provides a two-dimensional comparison of desired results to actual network response. In our case, it provides a 6*6 matrix, as there are six different classes under study. The output response of the network is thresholded with a 1-of-N transformation, where the winner output is valued at 1, and the others are valued at 0. The sum of the winners are divided by the number of input sets per output category, and the overall classification rate of the entire network is the average of the six classifications rate per category. However, the classification rate doesn't indicate what the output PE levels are, and what their range is. Such information is useful as it allows the designer to quantify the quality of the classification, and to adjust the threshold level above which classification to a specific class may be assigned if desired. PE output levels are presented in Tables

6.2 to 6.9.

6.c Classification Results

The study was conducted in two phases. The first phase of the study investigated the application of AR coefficients as feature parameters, while the second phase investigated the application of wavelet-type quantities as feature vectors. The same data was used for both phases, however, a smaller selection of training and testing files selected from the data was used for the first phase for all classes except the sperm whale class, which explains why the AR classification rates are based on a different number of training and testing files from those obtained for the Wavelet-based schemes. During the second phase of the study, each training class contained 87 signals, and an average of 50 sets per testing class was used for the testing phase.

Table 6.1 presents the overall classification rates obtained for the each of the feature extraction schemes considered in this study, where each network configuration is described in terms of the number of inputs/number of PEs in the first hidden layer/number of PEs in the second hidden layer/number of output nodes. Tables 6.2 to 6.8 present the detailed results obtained for each scheme. For clarity purposes, a detailed explanation of Table 6.2 is presented next, Tables 6.3 to 6.8 follow the same presentation. The first row in Table 6.2 shows the number of testing files presented to the NN for classification in each class. The first column indicates the number of testing files classified in a specific class. All rows show the mean and standard deviation (STD) obtained at each output node; the classification rate CR (in percentage), and the number of files classified in each specific signal class. Rows two to seven present individual performance results obtained for each class. For example, row number two shows that 38 files were classified as "sperm whale", and that 28 out of these 38 files were correctly classified, leading to a classification rate for that class CR=80%. Misclassified sperm whale data was classified as, either killer whale (5 files), or earthquake data (2 files). Recall that ideal output node levels should be either 1 or 0, however, in practice the levels are between 0 and 1. For example, the average output level obtained for the testing sperm whale data is 0.6887 and its standard deviation is .34. In addition, the sperm whale output node level significantly drops down to around 0-.05 when presented with other types of signals.

Classification results using AR coefficients

Classification results obtained when using AR coefficients as feature parameters are presented in Table 6.2. Results show that using AR coefficients lead to a relatively low classification rate equal to 84.76%.

Classification results using orthonormal wavelet coefficients

Classification results obtained for the orthonormal Wavelet Transforms considered are shown in Tables 6.3 and 6.4. We were unable to explain the difference in classification performances obtained while the frequency partitioning for the two bases is so similar. Further testing would be needed to explain such a difference further.

Classification results using non-orthonormal wavelet coefficients

Classification results obtained for the A-Trous Wavelet Transforms considered are shown in Tables 6.5 to 6.8 for Morlet wavelet with 4 to 7 voices. Results show that classification rates are higher than those obtained with AR and orthonormal wavelet transforms (between 93% and 96%). Note that this NN implementation may be viewed as an energy-type classifier, as the feature parameters chosen represents a measure of the average energy obtained in a given frequency range. Thus, the finer the spectral partitioning, the better the classification rates. Using multiple voices in the non-orthogonal transformation leads to a finer frequency decomposition of the signal information, which leads to a better match of the spectral information contained in the data under study.

7. Conclusions

This study compared the classification rates obtained when using Wavelet Transforms and AR modeling to select feature parameters as back-propagation NN inputs for classification purposes. Results show that the best overall classification rates are obtained when using the undecimated non-orthogonal A-trous implementation with multiple voices. These results are to be expected as the feature extraction scheme chosen for the Wavelet transforms can be viewed as an "energy-based" classifier, where the choice of the basis specifies the type of frequency partitioning used. The A-Trous implementation is well matched to the narrowband underwater data under study, as it leads to a finer frequency resolution than that obtained using orthogonal bases, which results in higher overall classification performances.

The data used in this study contained significant additive noise. An initial attempt to extract the signal form its noisy environment to increase the separability of the classes by applying a basic ALE filter. This adaptive scheme performed very poorly on wideband underwater signals (sperm whale and underwater earthquake), which contributed to an overall degradation of the classification performances. An improvement in the denoising could potentially be achieved by employing a wavelet-based denoising scheme based on work originally proposed by Donoho et. al. [5,15]. Further details regarding denoising schemes and their applications to underwater signals may be found in [15].

References

[1] M.P. Fargues, R. Cristi, and M.M. Vanderkamp, "Modeling and Classification of Biological Signals Using Least-Squares Prony-SVD AR Models," 36th Symposium on Circuits and Systems, Detroit, MI, Aug. 16-18, 1993.

[2] O. Rioul and M. Vetterli, "Wavelets and Signal Processing," ASSP Magazine, Oct. 1991, pp. 14-38.

[3] A. Lemer, J.M. Nicolas, and P. Giancome, "Identification Automatique de Bruits Impulsifs Sous-Marins," Proc. GRETSI, 1989.

[4] M.J. Shensa, "The Discrete Wavelet Transform; Wedding the A-Trous and Mallat Algorithms," IEEE Trans. on SP., Vol. 30, No. 5, Oct. 1982, pp. 734-738.

[5] D. L. Donoho, "De-Noising via Soft Thresholding," Technical Report, Department of Statistics, Stanford University, CA, 1992.

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

[7] K.L. Frack, *Improving Transient Signal Synthesis Through Noise Modeling and Noise Removal*, MSEE Thesis, Naval Postgraduate School, March 1994.

[8] M.P. Fargues and W. A. Brooks, "Comparative Study of Time-Frequency and Time-Scale Transforms for Ultra-Wideband Radar Transient Signal detection," Submitted to IEE Proceedings, Radar, Sonar and Navigation.

[9] NeuralWorks Professional II/Plus, NeuralWare, Inc., 1993.

[10] M.H. Hayes, *Statistical Digital Signal Processing and Modeling*, Wiley and Sons, 1996.

[11] C.W. Therrien, *Discrete Random Signals and Statistical Signal Processing*, Prentice Hall, 1992.

[12] S. Haykin, *Adaptive Filter Theory*, Prentice Hall, 1995.

[13] M. Vetterli and J. Kovacevic, *Wavelets and Subband Coding*, Prentice Hall, 1995.

[14] J. Buckheit, S. Chen, D. Donoho, and J. Scargle, *Wavelab .701*, URL: http://www.wavelab/playfair.stanford.edu, 1996.

[15] R.J. Barsanti, *Denoising of Ocean Acoustic Signals using Wavelet-based Techniques,* MSEE Thesis, Naval Postgraduate School, December 1996.

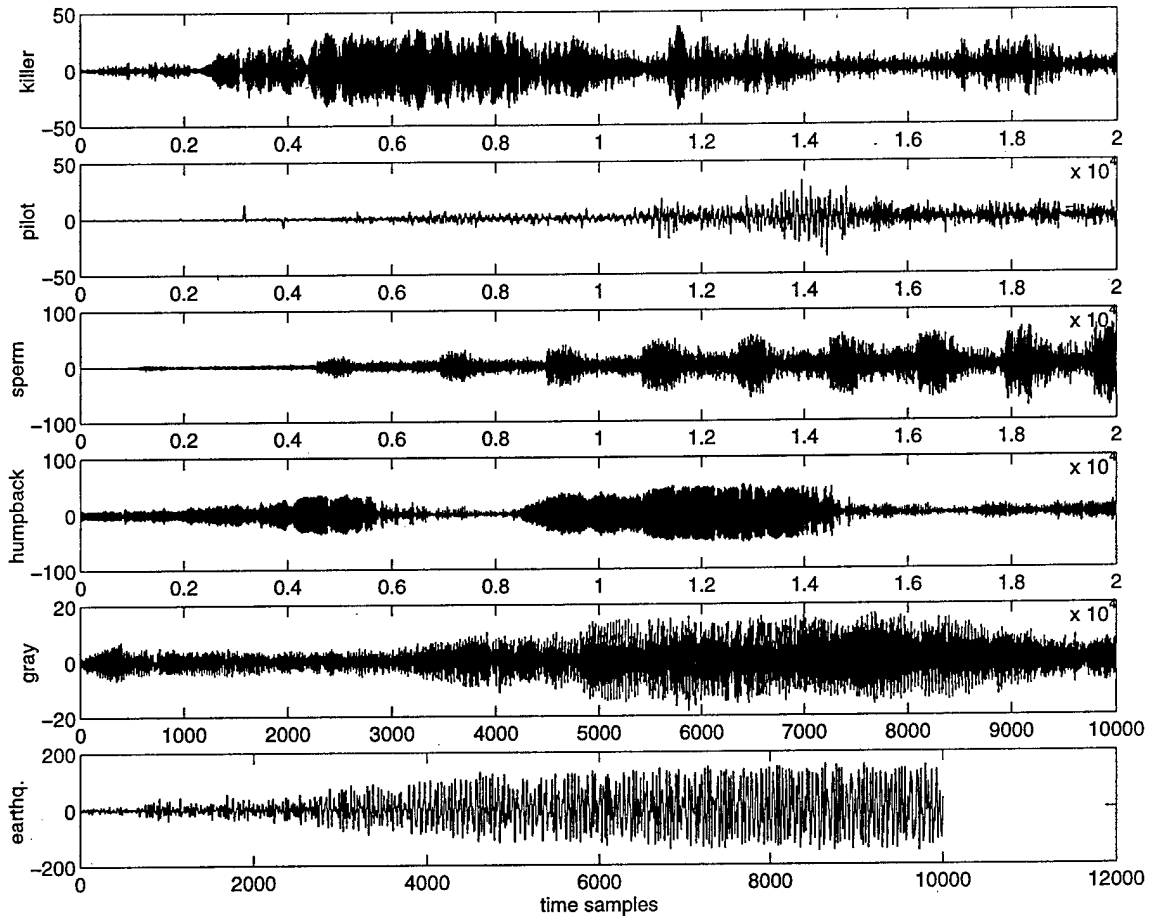


Figure 2.1. Time domain signals.

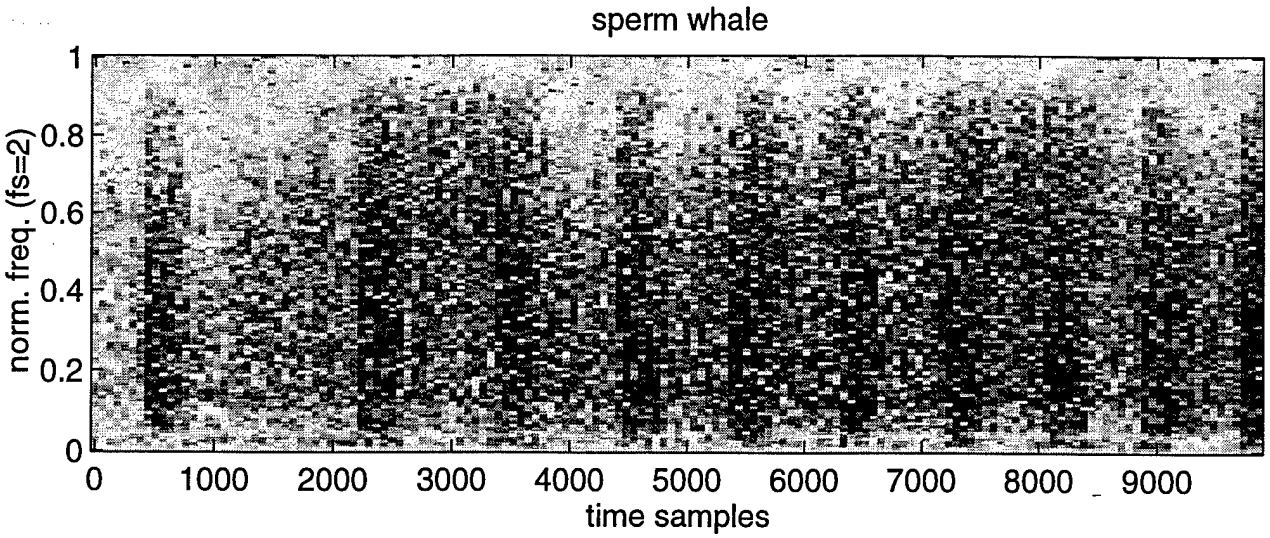


Figure 2.2. Spectrogram of sperm whale data; normalized frequency (f_s=8kHz), normalized time (number of samples).

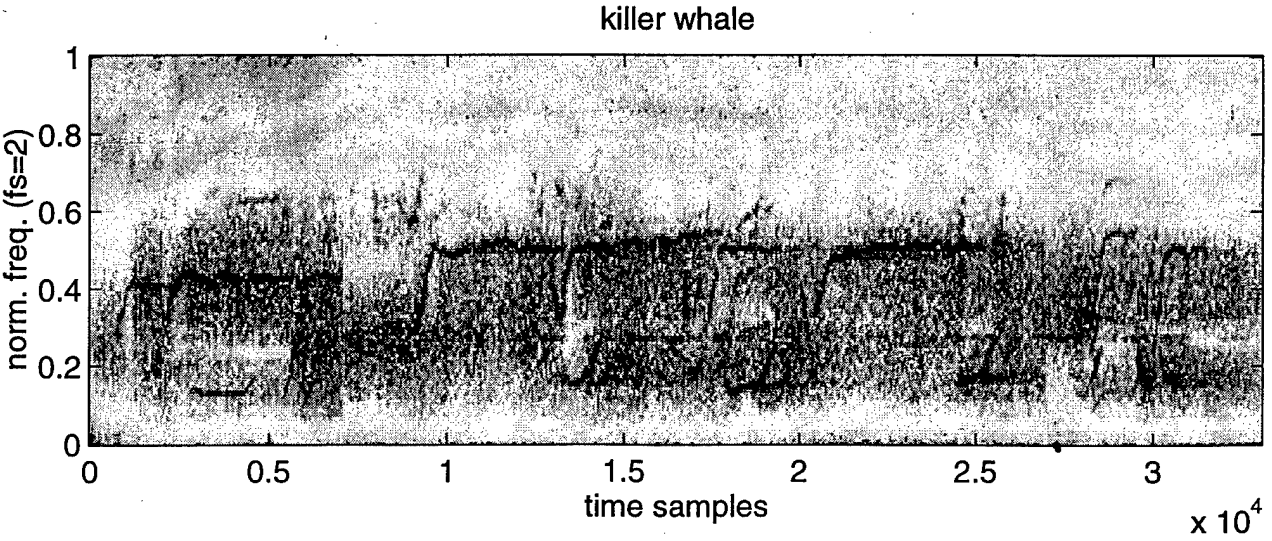


Figure 2.3. Spectrogram of killer whale data; normalized frequency (f_s=8kHz), normalized time (number of samples).

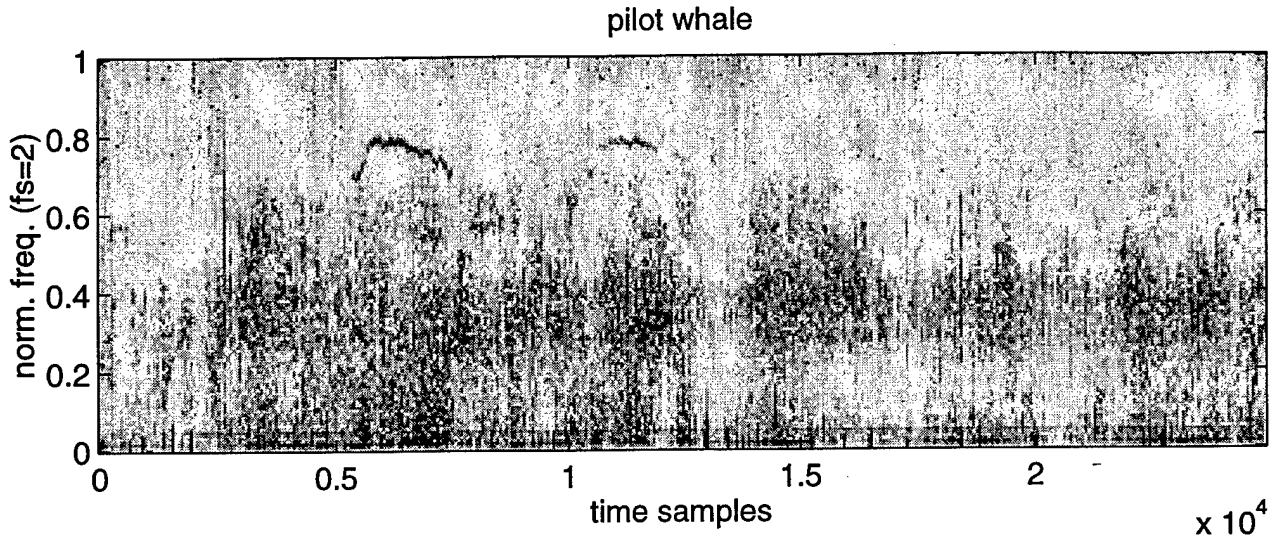


Figure 2.4. Spectrogram of pilot whale data; normalized frequency (f_s=8kHz), normalized time (number of samples).

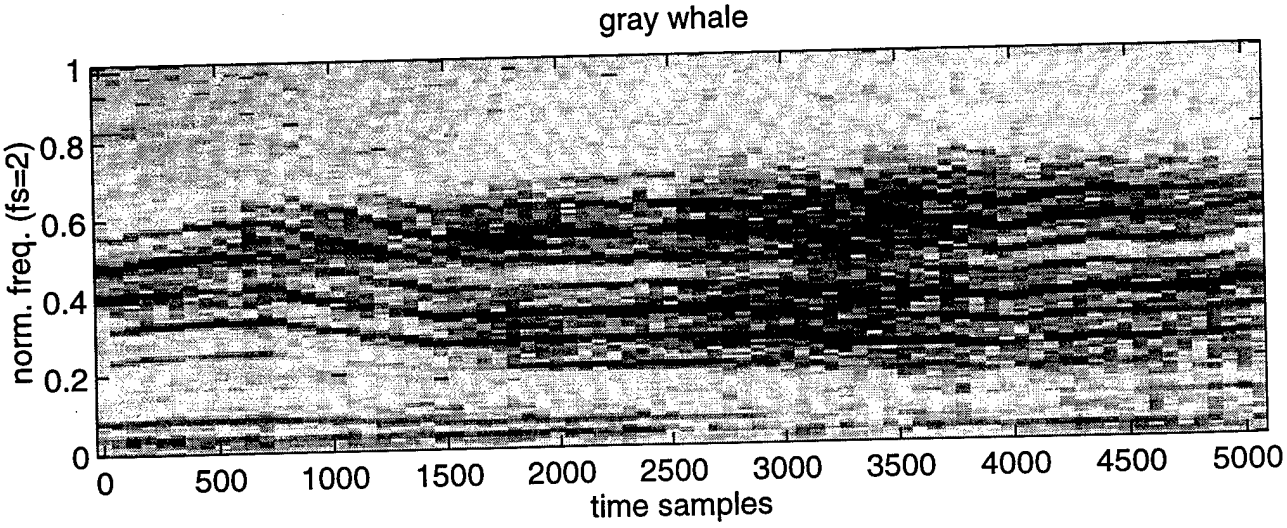


Figure 2.5. Spectrogram of gray whale data; normalized frequency (f_s=8kHz), normalized time (number of samples).

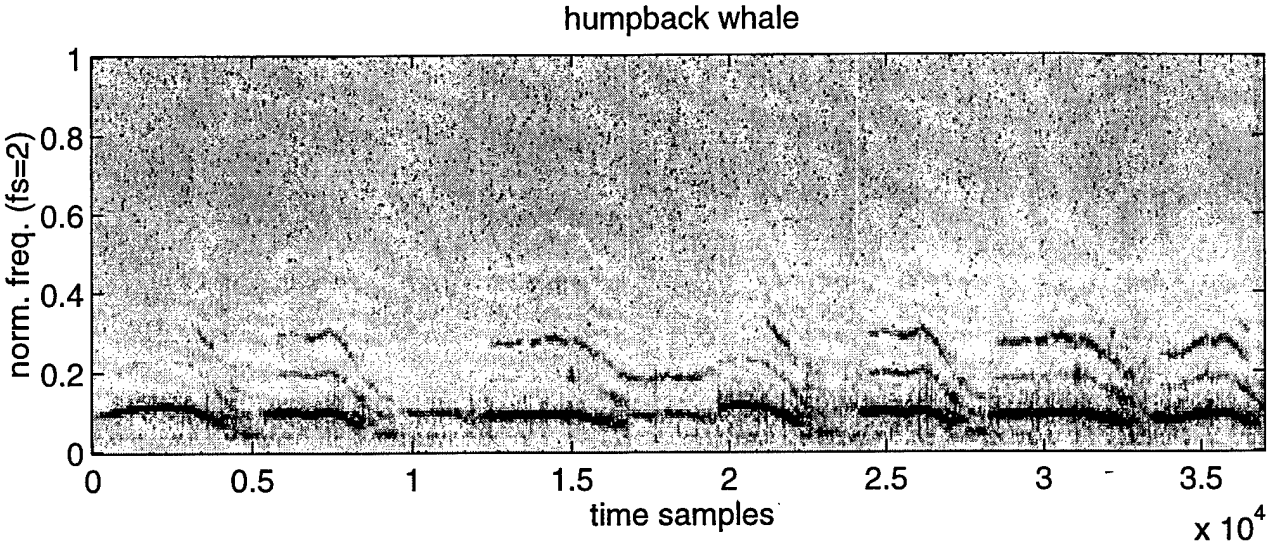


Figure 2.6. Spectrogram of humpback whale data; normalized frequency (f_s=8kHz), normalized time (number of samples).

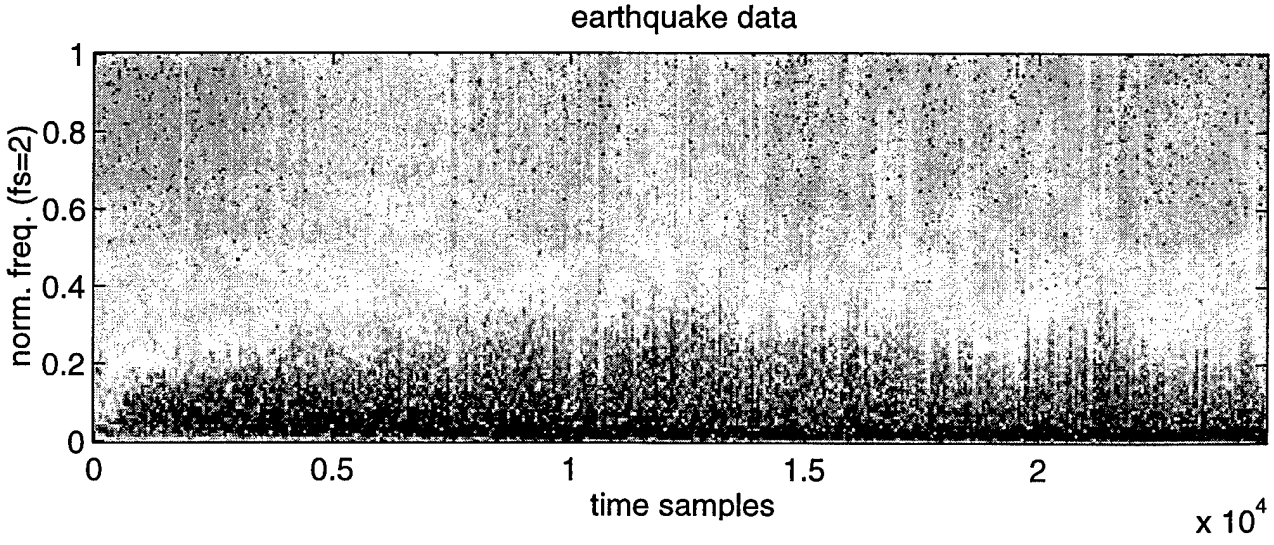


Figure 2.7. Spectrogram of underwater earthquake data; normalized frequency (f_s=8kHz), normalized time (number of samples).

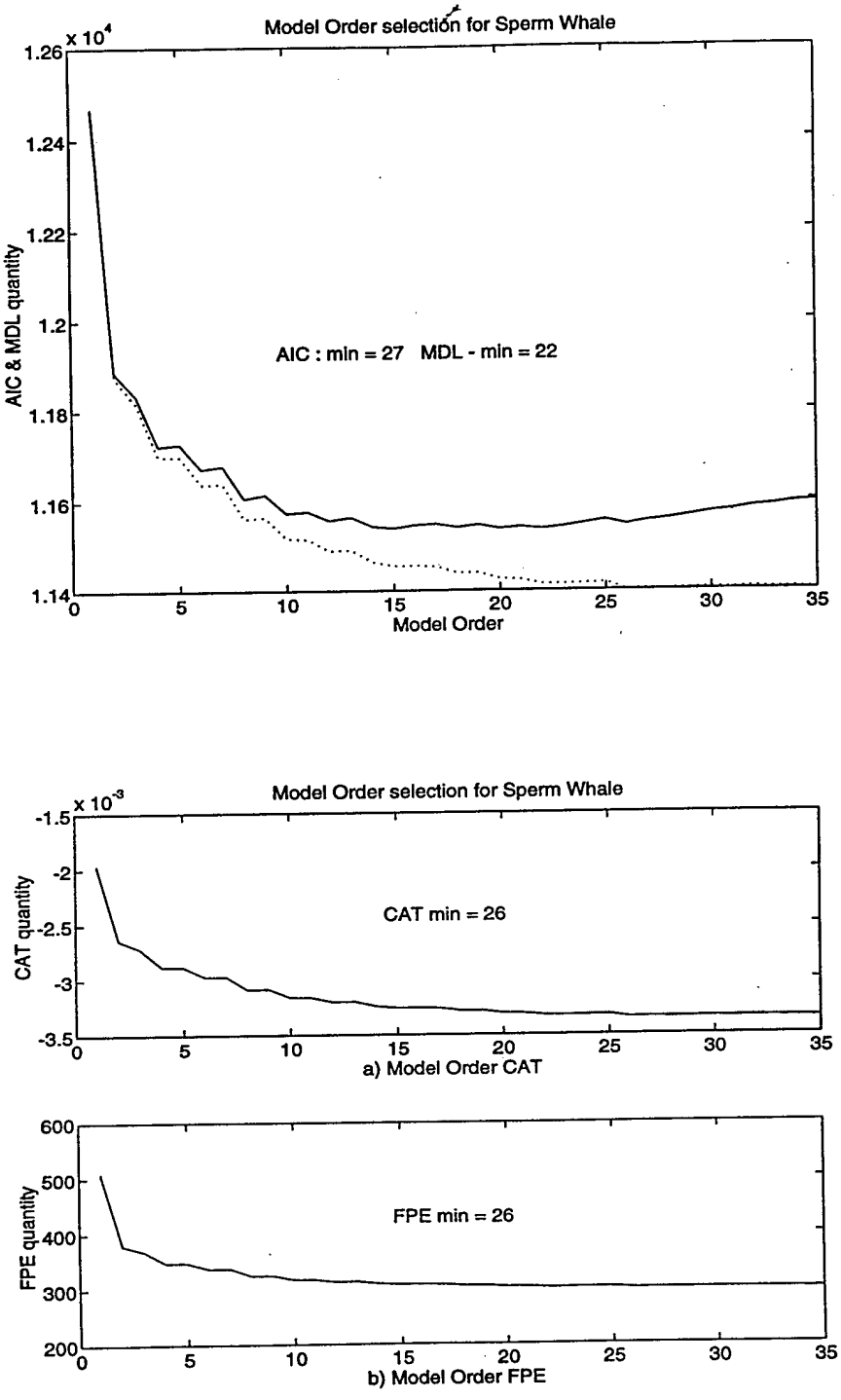


Figure 3.1. Model order selection for sperm whale using AIC, MDL, CAT and FPE criteria.

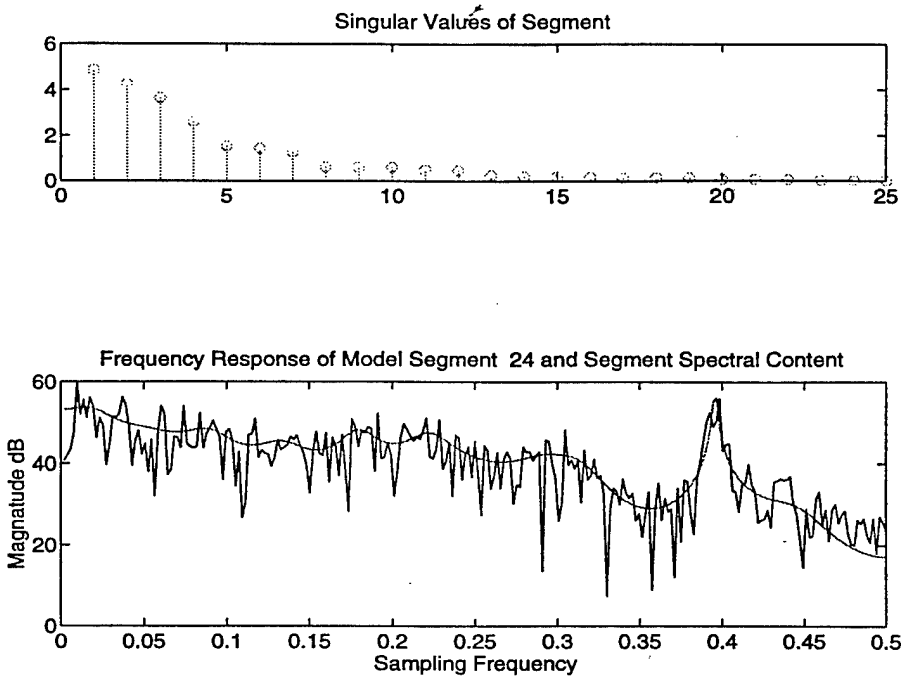


Figure 3.2. Pilot whale data; top plot: singular values of AR covariance matrix of order 25, bottom plot: typical AR and frequency spectra of data segment of length 512. Normalized frequency fs=1.

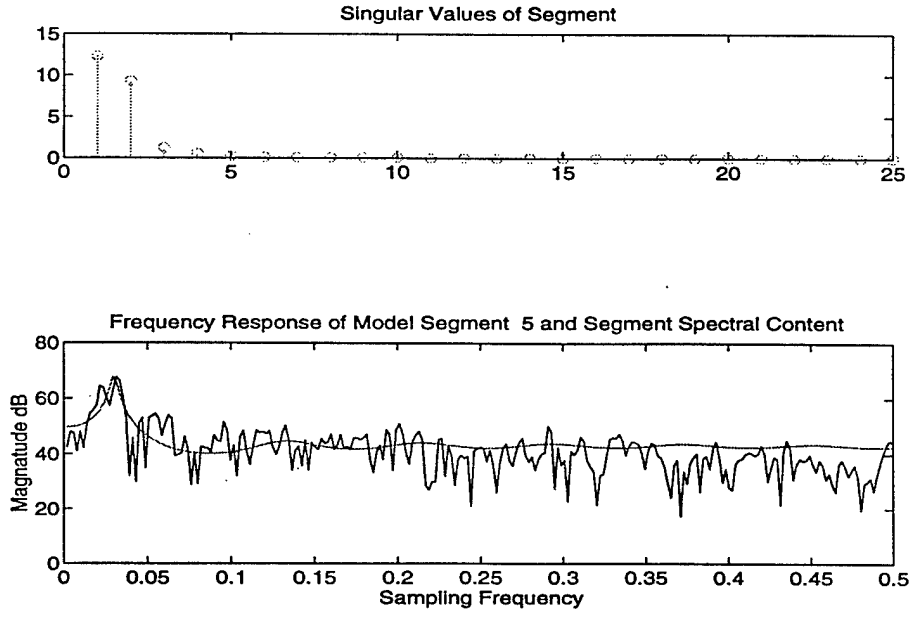


Figure 3.3. Earthquake data; top plot: singular values of AR covariance matrix of order 25, bottom plot: typical AR and frequency spectra of data segment of length 512. Normalized frequency fs=1.

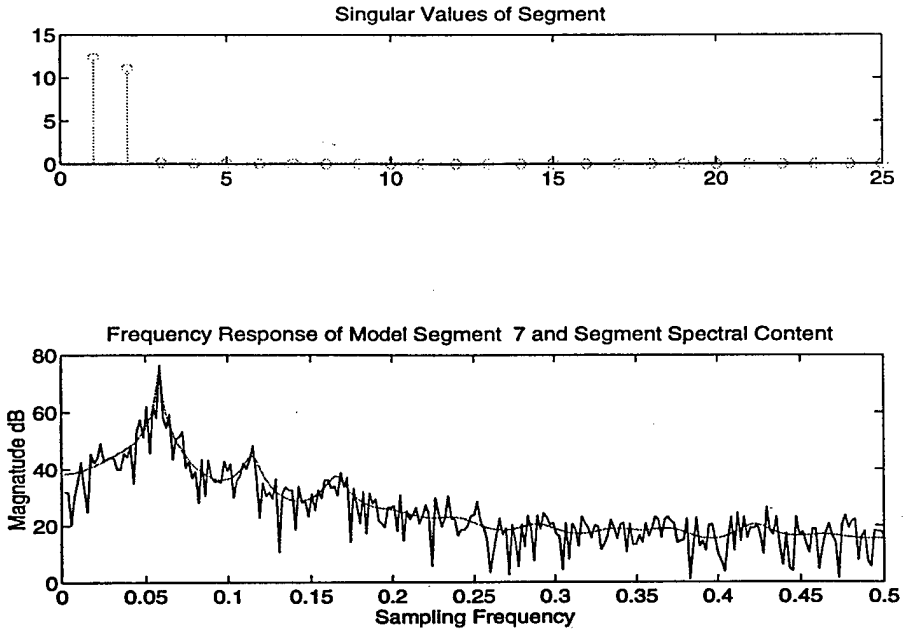


Figure 3.4. Humpback whale data; top plot: singular values of AR covariance matrix of order 25, bottom plot: typical AR and frequency spectra of data segment of length 512. Normalized frequency fs=1.

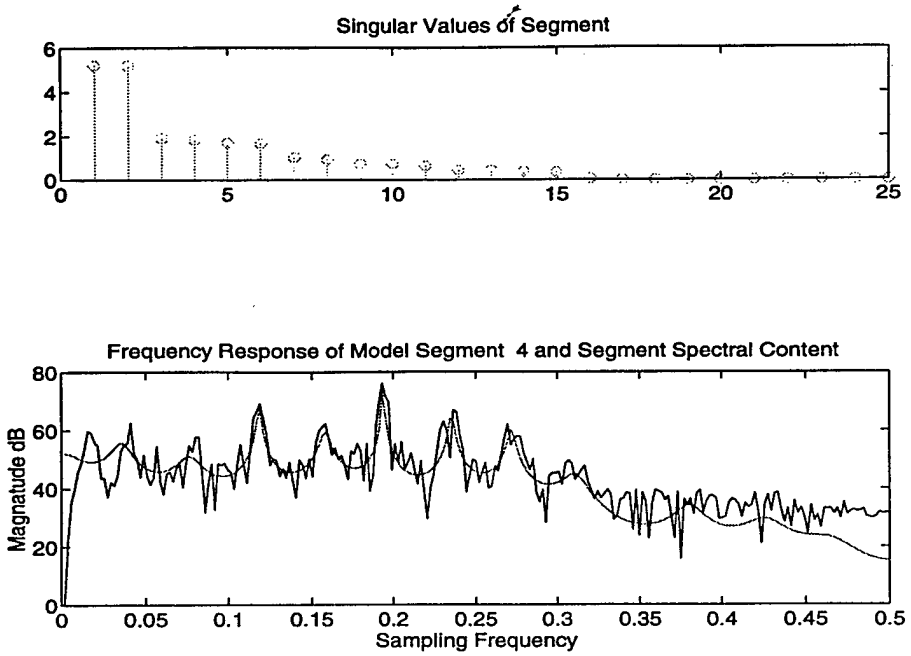


Figure 3.5. Gray whale data; top plot: singular values of AR covariance matrix of order 25, bottom plot: typical AR and frequency spectra of data segment of length 512. Normalized frequency fs=1.

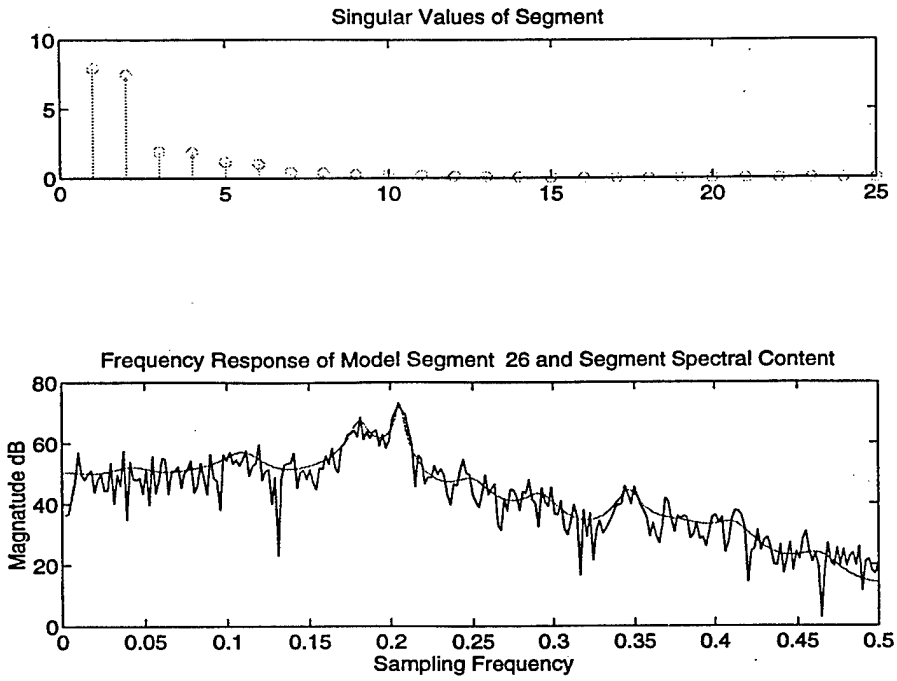


Figure 3.6. Killer whale data; top plot: singular values of AR covariance matrix of order 25, bottom plot: typical AR and frequency spectra of data segment of length 512. Normalized frequency fs=1.

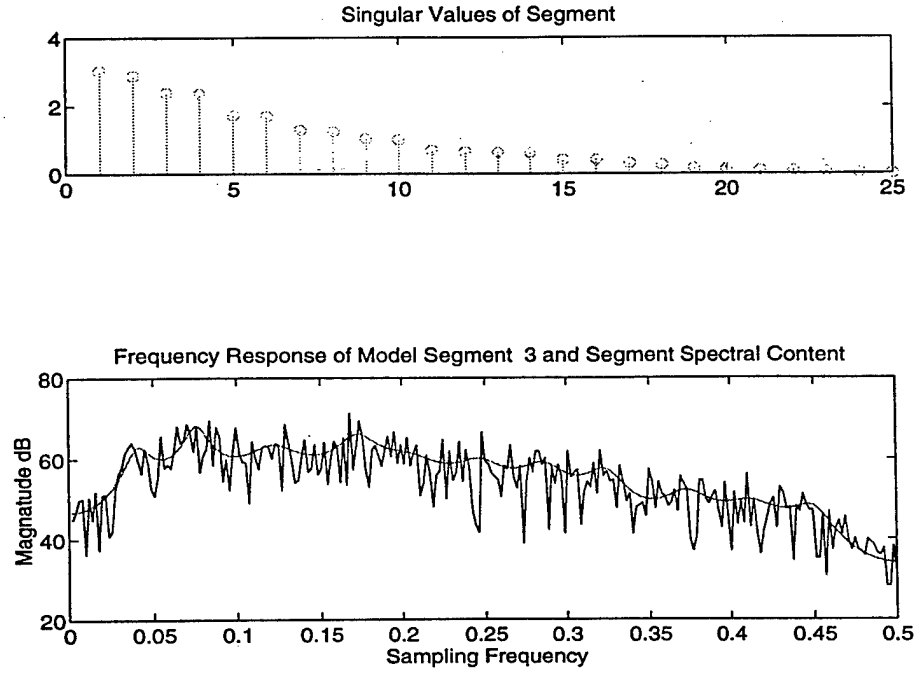


Figure 3.7. Sperm whale data; top plot: singular values of AR covariance matrix of order 25, bottom plot: typical AR and frequency spectra of data segment of length 512. Normalized frequency fs=1.

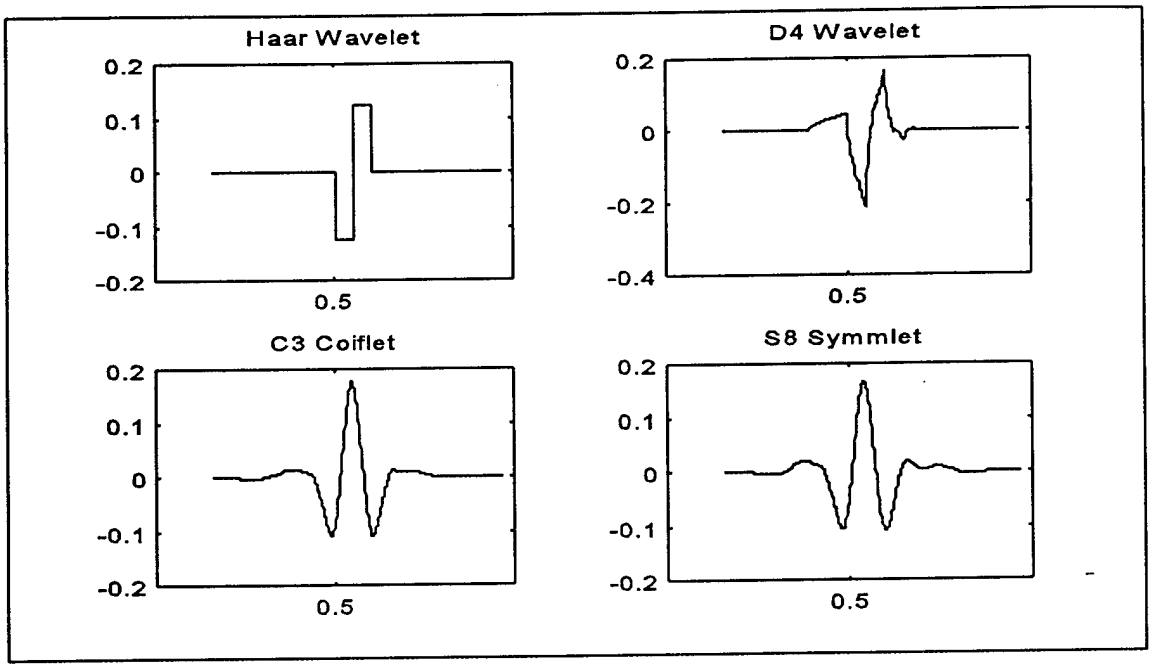


Figure 5.1. Four Wavelets in the Time Domain. From Ref. [14].

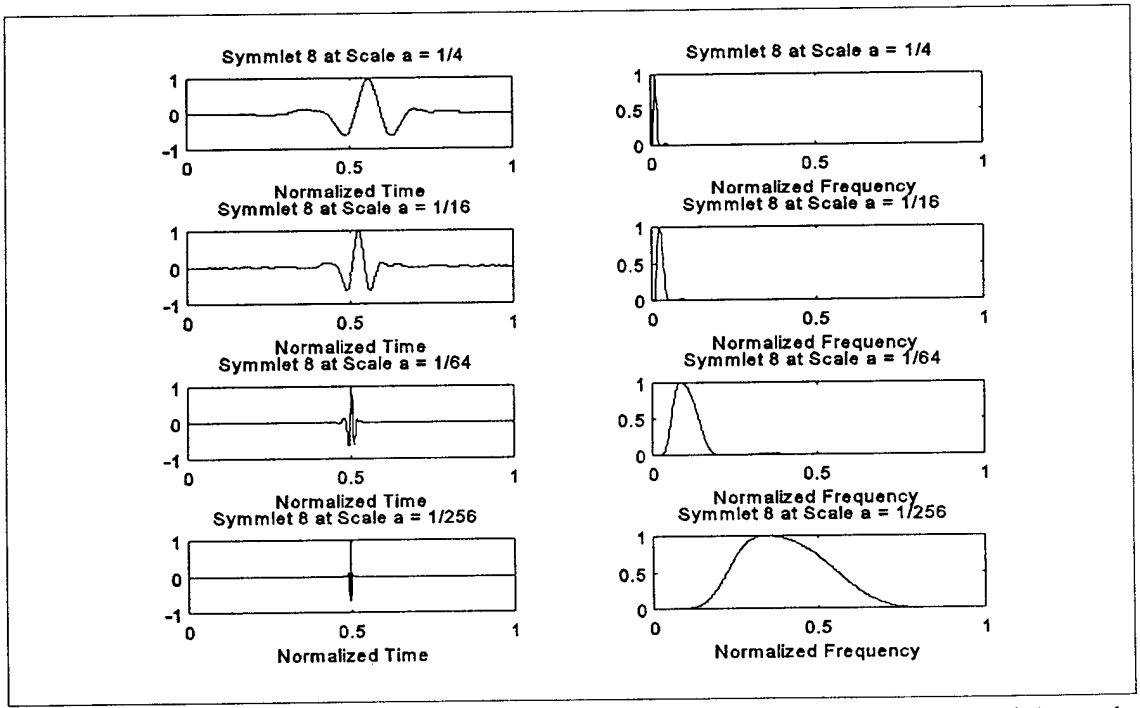
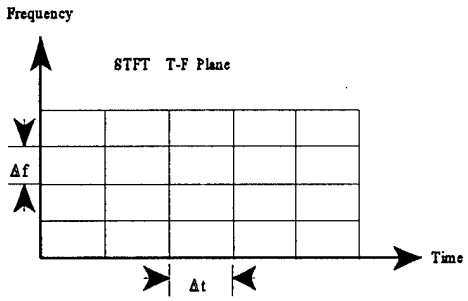


Figure 5.2. Symmlet 8 Wavelet in Time and Frequency domains as a function of the scale parameter a. The scale factor a decreases from the top to bottom plots. After Ref [14].



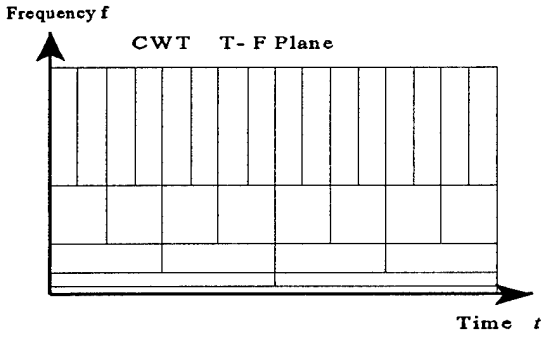


Figure 5.3. Time - Frequency plane for STFT and CWT.

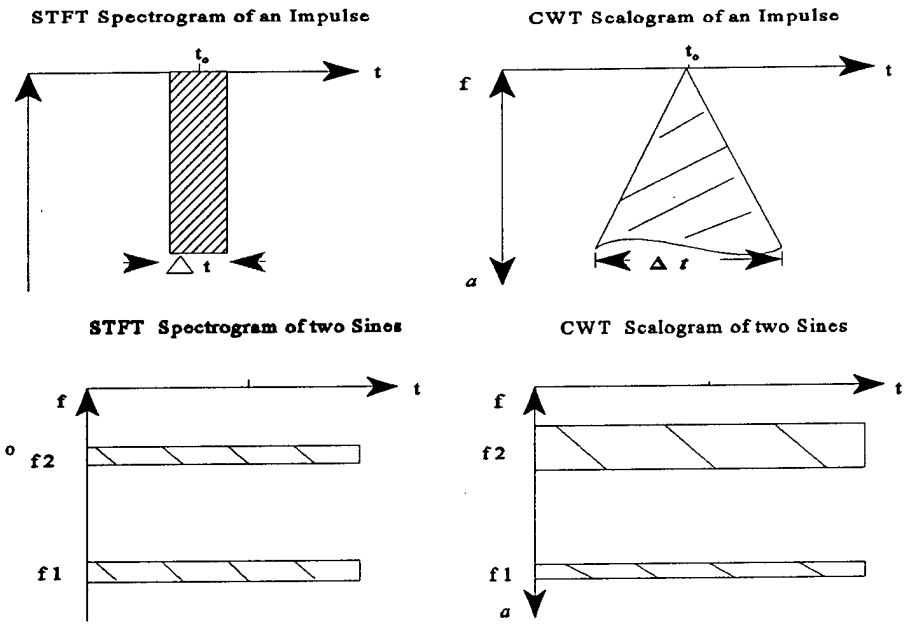


Figure 5.4. Spectrograms and Scalograms for two signals. Top plots display transforms for an impulse function. Bottom plots display transforms for two sines. After Ref [2].

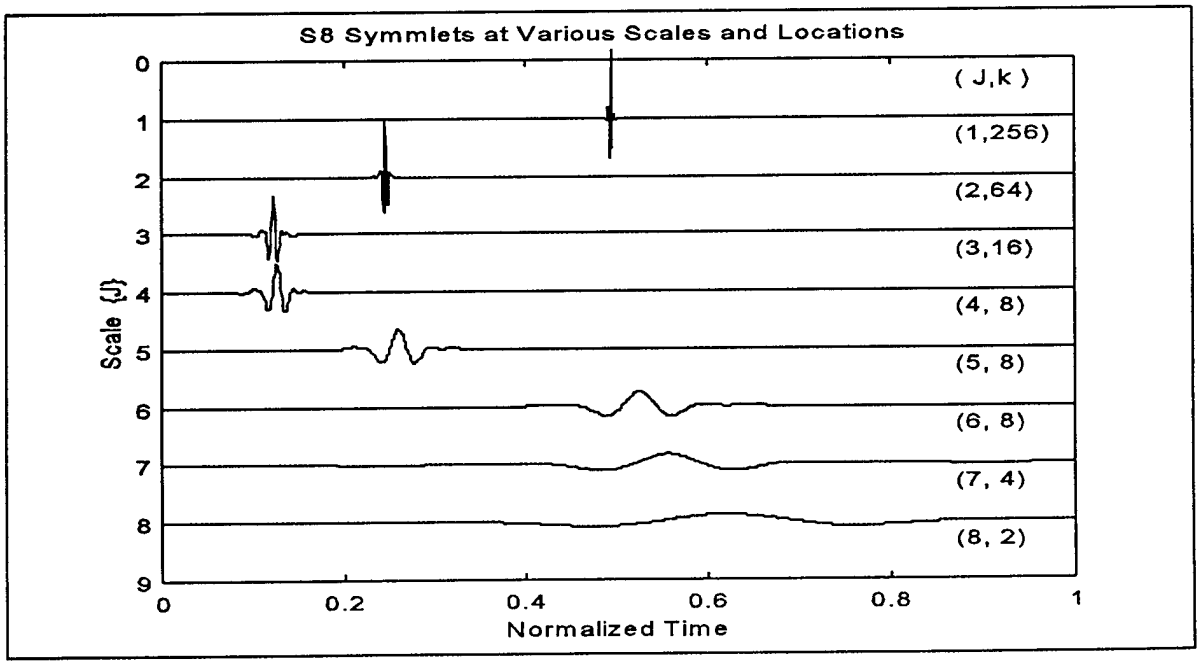


Figure 5.5. Symmlet 8 wavelet at various scales J and positions k.

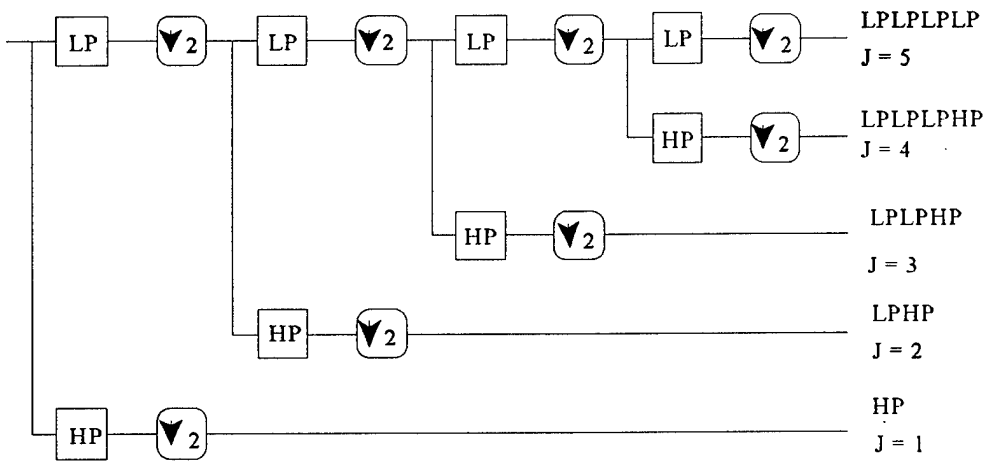


Figure 5.6. DWT implementation using filtering and down sampling operations.

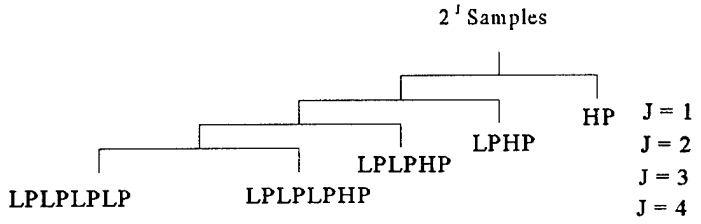


Figure 5.7. DWT tree structure.

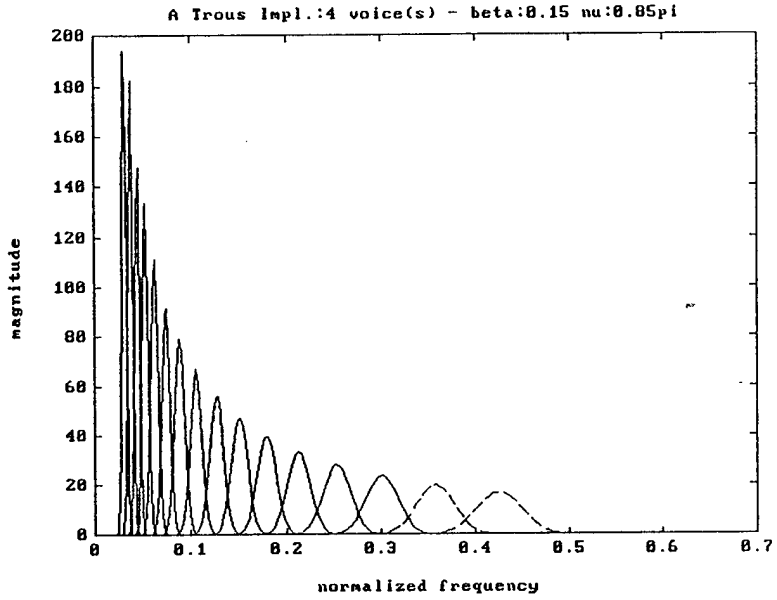


Figure 5.8.Spectral partitioning obtained for the A-Trous algorithm; 4 voices per octave; β=.15, η=.85π; 4 scales decomposition.

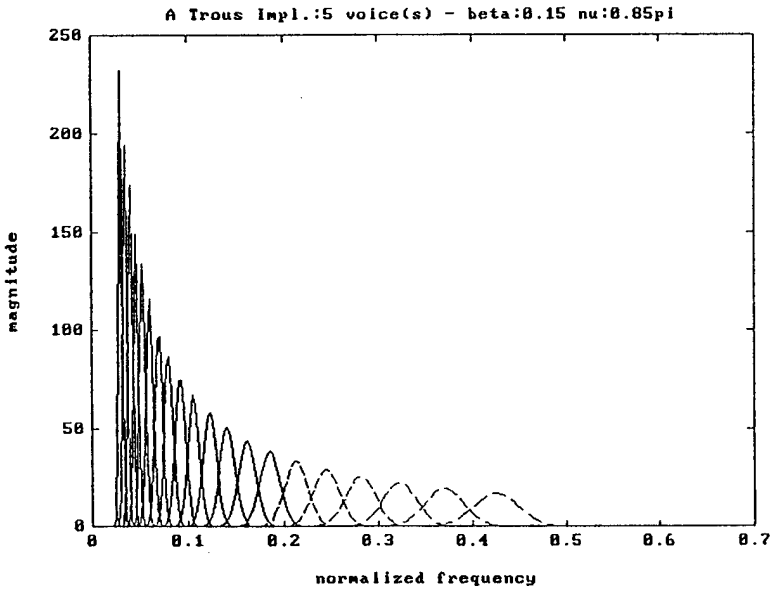


Figure 5.9. Spectral partitioning obtained for the A-Trous algorithm; 5 voices per octave; β=.15, η=.85π; 4 scales decomposition.

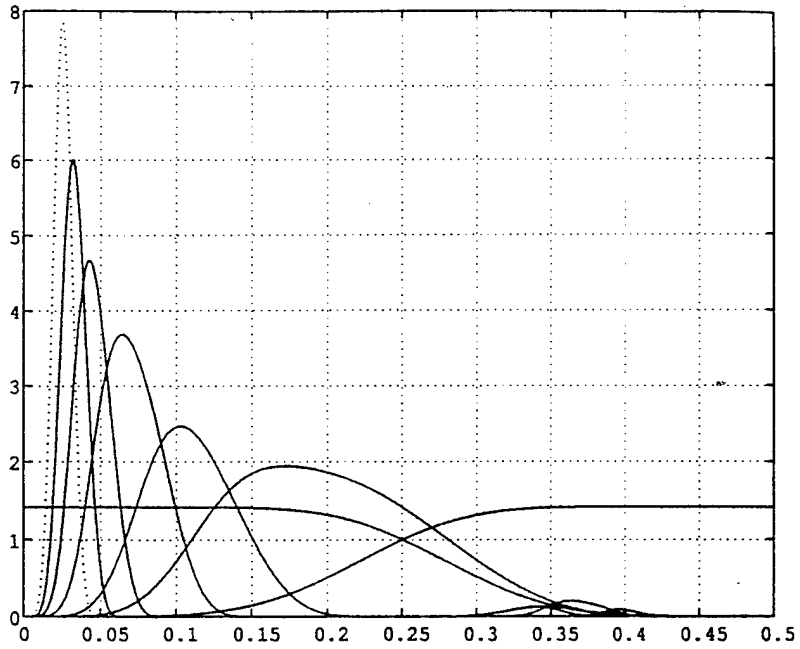


Figure 5.10. Spectral partitioning obtained for the Coiflet-3 wavelets.

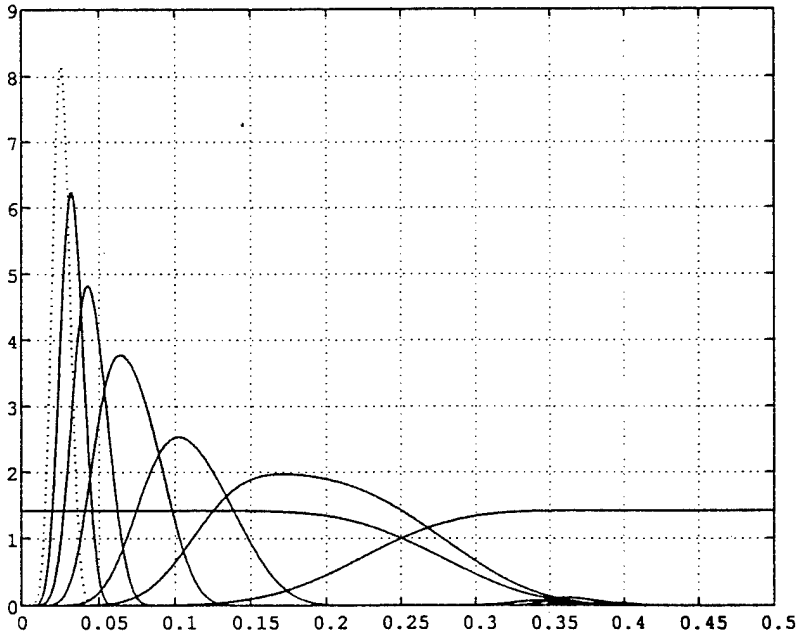


Figure 5.11. Spectral partitioning obtained for the Symmlet-8 wavelets.

Classification Technique	Neural Network input PE/ hidden layer 1/ hidden layer 2/ output	Overall Classification Rate
AR coefs	25/20/15/6	84.7%
ALE & AR coefs	25/20/15/6	83.7%
WT; Symmlet 8	14/14/10/6	84.67%
WT; Coiflet 3	14/14/10/6	78.2%
A-Trous; 4 voices	28/28/15/6	93.1%
A-Trous; 5 voices	35/20/15/6	93.4%
A-Trous; 6 voices	2/42/15/6	96.7%
A-Trous; 7 voices	49/49/15/6	95.1%

Table 6.1. Overall classification results

Table 6.2: Classification Performance Obtained Using Reduced Rank AR Coefficients

Test ⟹ Files / Classification Results ⇓	Sperm Input 35 files	Killer Input 35 files	Humpback Input 35 files	Gray Input 35 files	Pilot Input 35 files	Earthquake Input 35 files
	Mean Standard Direction (st) Classification Rate (CR%) Number of Files					
Sperm Whale Output 38 Files	0.6887 (0.3339) 80% 28 files	0.2132 (0.2791) 25.71% 9 files	0.0132 (0.1598) 0%	−0.0024 (0.1506) 0%	−0.0488 (0.1581) 2.86% 1 file	0.0627 (0.2187) 0%
Killer Whale Output 31 Files	0.0547 (0.3055) 14.29% 5 files	0.9824 (0.1273) 68.57% 24 files	0.0556 (0.1085) 0%	−0.0676 (0.0783) 5.71% 2 files	−0.0924 (0.0546) 0%	−0.0131 (0.1268) 0%
Humpback Whale Output 36 Files	0.0057 (0.0713) 0%	−0.0438 (0.1647) 0%	0.9387 (0.1959) 94.29% 33 files	−0.0796 (0.0695) 0%	0.0329 (0.0941) 8.57% 3 files	0.0286 (0.1801) 0%
Gray Whale Output 32 Files	−0.0339 (0.1194) 0%	0.0103 (0.2620) 2.86% 1 file	−0.0765 (0.0619) 0%	0.8505 (0.3716) 85.71% 30 files	0.1156 (0.2962) 2.86% 1 file	−0.0587 (0.0773) 0%
Pilot Whale Output 31 Files	0.0357 (0.2328) 0%	−0.0489 (0.1247) 0%	−0.0145 (0.2400) 0%	0.0433 (0.2217) 8.57% 3 files	0.7589 (0.3719) 80% 28 files	−0.0169 (0.1910) 0%
Earthquake Output 42 Files	−0.0988 (0.0224) 5.71% 2 files	−0.0345 (0.0204) 2.86% 1 file	0.0064 (0.0690) 5.71% 2 files	−0.0074 (0.0357) 0%	0.0198 (0.0247) 5.71% 2 files	0.9146 (0.1061) 100% 35 files

Overall classification rate: 84.67%

Table 6.3: Classification Performance Obtained Using Coiflet 3 Wavelet Basis

Test ⟹ Files / Classification Results ⇓	Sperm Input 50 files	Killer Input 50 files	Humpback Input 50 files	Gray Input 50 files	Pilot Input 50 files	Earthquake Input 50 files
	Mean Standard Direction (st) Classification Rate (CR%) Number of Files					
Sperm Whale Output 39 Files	0.8001 (0.2951) 70.73% 35 files	0.3167 (0.3426) 4.88% 2 files	−0.0118 (0.0550) 0%	−0.0220 (0.1354) 4.88% 2 files	−0.0075 (0.0622) 0%	0.0043 (0.1033) 0%
Killer Whale Output 58 Files	0.0866 (0.2688) 26.09% 13 files	0.6478 (0.2769) 75.61% 38 files	−0.0101 (0.1388) 0%	0.1137 (0.2751) 10% 5 files	−0.0179 (0.0828) 4.88% 2 files	−0.0693 (0.0513) 0%
Humpback Whale Output 48 Files	0.0181 (0.0430) 0%	−0.0020 (0.0283) 4.88% 2 files	0.9006 (0.1915) 90.24% 45 files	−0.0476 (0.0365) 0%	−0.0225 (0.0285) 2.44% 1 file	0.0535 (0.2207) 0%
Gray Whale Output 49 Files	−0.0229 (0.2280) 4.88% 2 files	0.2802 (0.1779) 16.43% 8 files	−0.0180 (0.0808) 0%	0.4671 (0.1797) 65.85% 33 files	0.2061 (0.2006) 12.2% 6 files	−0.0174 (0.0328) 0%
Pilot Whale Output 58 Files	−0.0878 (0.0630) 0%	0.0395 (0.1911) 0%	0.0464 (0.1183) 0%	0.0994 (0.2053) 19.85% 10 files	0.7962 (0.3190) 81.69% 41 files	−0.0571 (0.0561) 14.63% 7 files
Earthquake Output 48 Files	−0.0551 (0.0579) 0%	−0.0518 (0.0773) 0%	−0.1071 (0.0278) 10% 5 files	−0.0899 (0.0535) 0%	0.0662 (0.1589) 0%	0.8424 (0.3679) 85.37% 43 files

Overall classification rate: 78.24%

Table 6.4: Classification Performance Obtained Using Symmlet 8 Wavelet Basis

Test ⟹ Files / Classification Results ⇓	Sperm Input 50 files	Killer Input 50 files	Humpback Input 50 files	Gray Input 50 files	Pilot Input 50 files	Earthquake Input 50 files
	Mean Standard Direction (st) Classification Rate (CR%) Number of Files					
Sperm Whale Output 62 Files	0.8362 (0.2347) 90.2% 46 files	0.1897 (0.2012) 22% 11 files	−0.0362 (0.0334) 0%	0.0417 (0.1776) 10% 5 files	−0.0590 (0.0716) 0%	−0.0154 (0.0381) 0%
Killer Whale Output 46 Files	0.1990 (0.2662) 4% 2 files	0.6077 (0.3677) 64.71% 33 files	0.0042 (0.0842) 0%	0.0844 (0.2198) 22% 11 files	−0.0039 (0.1077) 0%	−0.0392 (0.0526) 0%
Humpback Whale Output 46 Files	−0.0330 (0.0451) 0%	0.0050 (0.1858) 0%	0.9086 (0.2236) 92% 46 files	−0.0008 (0.0342) 0%	−0.0278 (0.0757) 0%	0.0566 (0.2145) 0%
Gray Whale Output 41 Files	0.0926 (0.2294) 4% 2 files	0.2735 (0.1955) 12% 6 files	0.0056 (0.0339) 0%	0.4547 (0.2418) 62% 31 files	0.0996 (0.2053) 4% 2 files	0.0097 (0.1499) 0%
Pilot Whale Output 50 Files	−0.1071 (0.0213) 0%	−0.0576 (0.0826) 0%	−0.0124 (0.0385) 0%	0.1022 (0.1670) 4% 2 files	0.9957 (0.1474) 96% 48 files	−0.0073 (0.0776) 0%
Earthquake Output 55 Files	−0.0188 (0.0522) 0%	−0.0151 (0.0558) 0%	−0.0114 (0.0337) 8% 4 files	−0.0306 (0.0630) 2% 1 file	−0.0135 (0.0603) 0%	1.0203 (0.1275) 100% 50 files

Overall classification rate: 84.66%

Table 6.5: Classification Performance Obtained Using the A-Trous Algorithm; 4 Voices per Octave

Test ⟹ Files / Classification Results ⇓	Sperm Input 51 files	Killer Input 51 files	Humpback Input 51 files	Gray Input 51 files	Pilot Input 51 files	Earthquake Input 51 files
	Mean Standard Direction (st) Classification Rate (CR%) Number of Files					
Sperm Whale Output 50 Files	1.1055 (0.0768) 98.04% 50 files	−0.1058 (0.0803) 1.96% 1 file	0.1253 (0.2209) 0%	−0.0932 (0.1125) 0%	0.2603 (0.2989) 0%	−0.0588 (0.0481) 0%
Killer Whale Output 50 Files	−0.0279 (0.2152) 0%	0.8835 (0.2500) 90.20% 46 files	−0.0578 (0.1303) 0%	−0.0378 (0.1090) 0%	0.0317 (0.2125) 0%	−0.0599 (0.1133) 0%
Humpback Whale Output 51 Files	−0.0081 (0.0680) 0%	0.0571 (0.1544) 0%	1.0518 (0.1237) 100% 51 files	−0.0086 (0.0826) 0%	−0.0540 (0.0655) 0%	0.1096 (0.1198) 0%
Gray Whale Output 53 Files	−0.0709 (0.0672) 0%	−0.0566 (0.0845) 0%	−0.1007 (0.0424) 0%	1.0186 (0.1749) 98.04% 50 files	−0.0055 (0.1646) 5.88% 3 files	0.0018 (0.1737) 0%
Pilot Whale Output 52 Files	0.0668 (0.1091) 1.96% 1 file	−0.0250 (0.0857) 7.84% 4 files	0.0008 (0.0876) 0%	0.1808 (0.1753) 0%	0.7893 (0.2384) 92.16% 47 files	−0.0273 (0.1775) 0%
Earthquake Output 53 Files	−0.0799 (0.0371) 0%	−0.0862 (0.0699) 0%	−0.1247 (0.0009) 0%	0.0248 (0.1136) 1.96% 1 file	0.2086 (0.2744) 1.96% 1 file	1.0175 (0.2166) 100% 51 files

Overall classification rate: 96.41%

Table 6.6: Distribution of the Neural Network Classifications Obtained Using the A-Trous Implementation; 5 Voices per Scale

Test ⟹ Files / **Classification Results ⇓**	**Sperm Input 51 files**	**Killer Input 51 files**	**Humpback Input 51 files**	**Gray Input 51 files**	**Pilot Input 51 files**	**Earthquake Input 51 files**
	Mean Standard Direction (st) Classification Rate (CR%) Number of Files					
Sperm Whale Output 50 Files	1.0772 (0.1719) 98.04% 50 files	−0.0993 (0.0937) 0%	0.1176 (0.1915) 0%	−0.1176 (0.0225) 0%	0.2158 (0.2746) 0%	−0.0244 (0.0847) 0%
Killer Whale Output 50 Files	−0.0241 (0.2131) 0%	0.8539 (0.2843) 86.27% 44 files	−0.0205 (0.1391) 0%	0.0344 (0.2557) 5.88% 3 files	0.0585 (0.2295) 1.96% 1 file	−0.0349 (0.0497) 0%
Humpback Whale Output 49 Files	0.1622 (0.3230) 0%	−0.0119 (0.0822) 0%	1.0302 (0.2397) 96.08% 49 files	0.0658 (0.2254) 0%	0.0883 (0.2702) 1.96% 1 file	0.0450 (0.1794) 0%
Gray Whale Output 53 Files	−0.0877 (0.0376) 0%	0.0312 (0.2181) 11.76% 6 files	−0.0973 (0.0332) 0%	0.8953 (0.3045) 92.16% 47 files	−0.0364 (0.1775) 7.84% 4 files	0.0257 (0.1469) 0%
Pilot Whale Output 51 Files	−0.0437 (0.1527) 1.96% 1 file	−0.0075 (0.1035) 1.96% 1 file	0.0159 (0.0980) 0%	0.0592 (0.1890) 1.96% 1 file	0.7276 (0.3100) 88.24% 45 files	0.0036 (0.0991) 0%
Earthquake Output 53 Files	−0.0660 (0.0500) 0%	−0.1115 (0.0310) 0%	−0.1215 (0.0076) 3.92% 2 files	0.0119 (0.1349) 0%	0.2458 (0.3796) 0%	1.0710 (0.0951) 100% 51 files

Overall classification rate: 93.46%

Table 6.7: Classification Performance Obtained Using the A-Trous Implementation; 6 Voices per Scale

Test ⟹ Files / Classification Results ⇓	Sperm Input 51 files	Killer Input 51 files	Humpback Input 51 files	Gray Input 51 files	Pilot Input 51 files	Earthquake Input 51 files
	Mean Standard Direction (st) Classification Rate (CR%) Number of Files					
Sperm Whale Output 50 Files	1.1107 (0.0751) 98.04% 50 files	−0.0848 (0.1879) 0%	0.1438 (0.0100) 0%	−0.1224 (0.3110) 0%	0.2516 (0.0670) 0%	−0.0579 (0.0481) 0%
Killer Whale Output 51 Files	−0.0735 (0.1462) 0%	0.9539 (0.2745) 98.04% 50 files	0.0807 (0.1903) 0%	−0.0271 (0.1585) 1.96% 1 file	−0.0321 (0.0450) 0%	−0.0340 (0.1133) 0%
Humpback Whale Output 49 Files	−0.0468 (0.0856) 0%	0.2230 (0.2996) 0%	1.0638 (0.1425) 96.08% 49 files	0.0304 (0.1423) 0%	−0.0205 (0.1759) 0%	0.0399 (0.1198) 0%
Gray Whale Output 52 Files	−0.0916 (0.0480) 0%	0.0776 (0.3277) 0%	−0.1028 (0.3471) 0%	0.9014 (0.1959) 96.08% 49 files	−0.0263 (0.1633) 5.88% 3 files	0.0083 (0.1737) 0%
Pilot Whale Output 48 Files	0.0221 (0.2134) 1.96% 1 file	−0.0585 (0.0892) 1.96% 1 file	0.0136 (0.2466) 0%	0.0238 (0.3758) 0%	0.8450 (0.1460) 92.16% 46 files	−0.0198 (0.1775) 0%
Earthquake Output 53 Files	−0.1000 (0.0249) 0%	−0.1159 (0.0198) 0%	−0.1233 (0.2570) 3.92% 2 files	0.1588 (0.3106) 1.96% 1 file	0.1851 (0.1454) 1.96% 1 file	1.0276 (0.2166) 100% 51 files

Overall classification rate: 96.73%

Table 6.8: Classification Performance Obtained Using the A-Trous Implementation; 7 Voices per Scale

Test ⟹ Files / Classification Results ⇓	Sperm Input 51 files	Killer Input 51 files	Humpback Input 51 files	Gray Input 51 files	Pilot Input 51 files	Earthquake Input 51 files
	Mean Standard Direction (st) Classification Rate (CR%) Number of Files					
Sperm Whale Output 51 Files	1.1107 (0.0751) 100% 51 files	−0.0848 (0.1879) 0%	0.1438 (0.0100) 0%	−0.1224 (0.3110) 0%	0.2516 (0.0670) 0%	−0.0579 (0.0481) 0%
Killer Whale Output 57 Files	−0.0735 (0.1462) 0%	0.9539 (0.2745) 98.04% 50 files	0.0807 (0.1903) 0%	−0.0271 (0.1585) 13.73% 7 files	−0.0321 (0.0450) 0%	−0.0340 (0.1133) 0%
Humpback Whale Output 50 Files	−0.0468 (0.0856) 0%	0.2230 (0.2996) 0%	1.0638 (0.1425) 98.04% 50 files	0.0304 (0.1423) 0%	−0.0205 (0.1759) 0%	0.0399 (0.1198) 0%
Gray Whale Output 49 Files	−0.0916 (0.0480) 0%	0.0776 (0.3277) 1.96% 1 file	−0.1028 (0.3471) 0%	0.9014 (0.1959) 84.31% 43 files	−0.0263 (0.1633) 9.80% 5 files	0.0083 (0.1737) 0%
Pilot Whale Output 47 Files	0.0221 (0.2134) 0%	−0.0585 (0.0892) 0%	0.0136 (0.2466) 0%	0.0238 (0.3758) 1.96% 1 file	0.8450 (0.1460) 90.20% 46 files	−0.0198 (0.1775) 0%
Earthquake Output 52 Files	−0.1000 (0.0249) 0%	−0.1159 (0.0198) 0%	−0.1233 (0.2570) 1.96% 1 file	0.1588 (0.3106) 0%	0.1851 (0.1454) 0%	1.0276 (0.2166) 100% 51 files

Overall classification rate: 95.098%

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Technical Information Center 8725 John J. Kingman Rd, STE 0944 Ft. Belvoir, VA 22060-6218	2
2.	Dudley Knox Library, Code 52 Naval Postgraduate School Monterey, CA 93943-5100	2
3.	Chairman, Code EC Department of Electrical and Computer Engineering Naval Postgraduate School 833 Dyer Road, Room 437 Monterey, CA 93943-5121	1
4.	Prof. Monique P. Fargues, Code EC/Fa Department of Electrical and Computer Engineering Naval Postgraduate School 833 Dyer Road, Room 437 Monterey, CA 93943-5121	4
5.	LCDR R. J. Barsanti, Code EC Department of Electrical and Computer Engineering Naval Postgraduate School 833 Dyer Road, Room 437 Monterey, CA 93943-5121	1