

A Toolset For Classifying Reflectors In Acoustic Images

by


Dean M. Addison
B.Sc., UVic, 1991

A Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of


MASTER OF APPLIED SCIENCE

in the Department of Mechanical Engineering.


We accept this thesis as conforming
to the required standard




Dr. D. R. Topham, Supervisor (Department of Mechanical Engineering)



Dr. S. Dost, Departmental Member (Department of Mechanical Engineering)



Dr. D. Olesky, Outside Member (Department of Computer Science)



Dr. A. Zielinski, External Examiner (Department of Electrical and Computer Engineering)

© DEAN M. ADDISON, 1998
University of Victoria


All rights reserved. This thesis may not be reproduced in whole or in part, by photocopy or other means, without permission of the author.

Supervisor: Dr. David R. Topham

Abstract

The focus of this work is a set of computational tools with which to classify and manipulate reflectors in underwater acoustic images. Acoustic images are generated by recording the reflections of repeated pulses of sound projected through some medium. The reflections from each pulse characterise the objects from which they are reflected. A multifaceted examination of the constituents of reflected signals in the time/space and amplitude domains can be used to classify the reflectors in the medium. Three methods: Spatial deconvolution, wavelet analysis, and particle tracking will, in many cases, provide a useful tool-set to classify biological reflectors in hydro-acoustic images.

Examiners:


Dr. D. R. Topham, Supervisor (Department of Mechanical Engineering)


Dr. S. Dost, Departmental Member (Department of Mechanical Engineering)


Dr. D. Olesky, Outside Member (Department of Computer Science)



Dr. A. Zielinski, External Examiner (Department of Electrical and Computer Engineering)

Table of Contents

Abstract	i
Table of Contents	ii
List of Tables	iv
List of Figures	v
Acknowledgements	vii
Dedication	viii
1 Introduction	1
1.1 Background	1
1.1.1 Underwater acoustics	3
1.1.2 Acoustic domain	6
1.2 Specifics	7
1.2.1 Solution of pulsating sphere	7
1.2.2 Physical medium	11
1.2.3 Digital medium	15
1.2.4 Equipment	17
2 Investigation	22
2.1 Detection	23
2.1.1 Image normalisation	23
2.1.2 Peak detection	29
2.1.3 Deconvolution	30
2.1.4 Wavelet detection	41
2.2 Classification	49
2.2.1 Transient properties	49
2.2.2 Persistent properties	55
2.3 Other techniques	64
2.3.1 Image compression	64
2.4 Synthetic data	66

TABLE OF CONTENTS

iii

2.4.1	Particle Response Model	66
2.4.2	Model Construction	69
3	Experimental Results	71
3.1	Synthetic data	73
3.1.1	Detection	74
3.1.2	Classification	85
3.2	Sampled data	91
3.2.1	Detection	95
3.2.2	Classification	105
4	Discussion and Conclusion	118
	References	123
	Suggested Reading	125

List of Tables

2.1	List of parameters for the synthetic image model.	70
3.1	Basic synthetic image, S1, with four reflectors.	73
3.2	Synthetic image S2.	74
3.3	Parameters for the two sampled images 3B6 and KL930.	92

List of Figures

1.1	Acoustic wave generated by a radially oscillating sphere.	8
1.2	A comparison of the physical and digital representation of the reflector. . .	15
1.3	200 kHz modulated carrier with positive and negative envelopes.	18
1.4	Sample image vector, a) without and b) with logarithmic compression. . .	20
2.1	dLTM normalisation example. The original image vector, a), is normalised with dLTM to yield b).	27
2.2	edLTM normalisation example. The original image vector, a), is normalised with edLTM to yield b).	28
2.3	Three distinct sets of response functions from the 3B6 acoustic image. The set of objects in the foreground were the deepest, those in the background, the shallowest.	35
2.4	Each response function is the average of a set of extracted objects from 3B6 a) normalised to have an integral of 1.0 and b) normalised to have a peak value of 1.0.	36
2.5	a) This response function is constructed from a linear attack and an exponential decay. b) A number of singular poles are evident in the Fourier domain.	38
2.6	a) The response function with an exponential attack. b) The representation of a) in the Fourier domain.	39
2.7	a) This response function is constructed from an exponential decay only. b) The Fourier domain of the response.	40
2.8	Daubechies wavelets: a) DAUB4 and b) DAUB12.	43
2.9	Example of the discrete wavelet transform using the DAUB4 wavelets. The graph in b) is the DWT of a).	47
2.10	Particle properties for non-moving and moving reflectors. Reflector images appear in a) $z^k = \text{constant}$ and c) $z^k = f(t)$. Amplitude profiles are in b) and d), respectively.	56
3.1	The data-flow diagramme for the evaluation of the detection, classification, and extraction algorithms. The separate levels of the diagramme represent the separate stages of the analysis.	72
3.2	The synthetic images a) S1 and b) S2.	75

3.3	The result of the time-domain peak detection algorithm, dPeak, on images a) S1 and b) S2, respectively.	77
3.4	This a) response function and, its associated b) Fourier domain representation is extracted from the synthetic image S1.	78
3.5	The image a) S1 and b) S2, deconvolved with the PRM response function: $a = 0.05$, $b_1 = -4.5$, $b_2 = -3.0$, $b_3 = -1.0$	80
3.6	The result of the application of a threshold to the deconvolved images a) S1 and b) S2.	81
3.7	The wavelet domain representation of a) S1, b) S2, and c) an enhanced S1. In the last image the separate hierarchies are clearly visible.	82
3.8	The result of the wavelet detection of images a) S1 and b) S2.	84
3.9	An example of the granularity of the wavelet detection.	86
3.10	The particle tracks found in image S1. The Streaks classification used the results of the a) dPeak, b) Deconv-dPeak, and c) Wavelet-dPeak algorithms.	87
3.11	The particle tracks found in image S2. The Streaks classification used the results of the a) dPeak, b) Deconv-dPeak, and c) Wavelet-dPeak algorithms.	88
3.12	The inverse wavelet transform of hierarchies a) 6 and b) 9 from S1.	90
3.13	The image 3B6.	93
3.14	The image KL930.	94
3.15	The image 3B6, normalised with the calculated LTM.	96
3.16	The image 3B6, normalised with the exponential fitted LTM.	97
3.17	The image KL930, normalised with the edLTM.	98
3.18	The dPeak detection of image 3B6 showing the peak detection in the time domain.	100
3.19	The dPeak detection of image KL930 showing the peak detection in the time domain.	101
3.20	a) The response function extracted from image 3B6. b) The magnitude and its reciprocal of the sample response function in the Fourier domain.	102
3.21	The result of the deconvolution applied to 3B6.	103
3.22	The result of Deconv-Thresh applied to 3B6.	104
3.23	The wavelet transform of image 3B6.	106
3.24	The wavelet transform of image KL930.	107
3.25	The Wavelet-dPeak detection of image 3B6 showing only detected peaks.	108
3.26	The Wavelet-dPeak detection of image KL930 showing only detected peaks.	109
3.27	The hSelect classification of image 3B6 using hierarchies 8–9.	110
3.28	The hSelect classification of image KL930 using hierarchies 7–8.	111
3.29	The dPeak-Streaks classification of image 3B6. All objects in the image belong to at least one particle track.	113
3.30	The Deconv-dPeak-Streaks classification of image 3B6.	114
3.31	The Wavelet-dPeak-Streaks classification of image 3B6.	115
3.32	The Wavelet-dPeak-Streaks classification of image KL930.	116

Acknowledgements

There are a number of people to whom I am indebted for help and guidance. First and foremost, I owe the deepest debt to Charles Konzelman for his financial support and tutelage during the largest part of this investigation. Other commitments prevented Charles from being listed as a co-supervisor for this thesis but there is no doubt that he was exactly that. Charles is also the best course instructor that I have had and I have had no shortage of excellent instructors. My undying gratitude and respect goes to David Topham who provided the data and the ideas for this thesis. Without his expertise and guidance this thesis would not have happened. Thanks go to Sadık Dost, who also provided financial support, and to the other committee members for their valuable time and comments. Selçuk Gürsan, my friend, was way too patient with my incessant questions on grammar. Sarah Sinclair was my confidant, editor, image typesetter, and last but not least, my wife.

Dedication

For **Sarah**
and the one
who is soon to be.

Chapter 1

Introduction

1.1 Background

This project had its beginning in 1985 with the Ice Keel Experiment [17, 18] in the Arctic. Originally, the goal of the experiment was to investigate two-layer flow past a land-locked ice keel in the Barrow Strait. The experiment used underwater acoustics to probe the structure of the water column beneath the ice. This was implemented at a number of stations on the upstream and lee sides of the keel.

The density interface between the two differing layers of water was the source of particular interest. As the water flowed in two layers past the keel, a super-critical jet formed which evolved into a turbulent region downstream of the keel. This turbulent region contained microstructure that was visible in the acoustic images. Although much of the microstructure was clear in the images, the biological objects in the water column were dominant outside the turbulent regions.

The majority of the biological reflectors serve as passive tracers of the fluid motion, the acoustic signatures providing information on the local flow characteristics. In particular, the vertical velocities can be estimated from the motion of the passive reflectors. To facilitate a computational analysis of the structure of the water column, it is necessary to identify these

particular signals and extract them in a form suitable for computational analysis. Ideally, this would be done automatically, as part of the processing of the acoustic images.

Removing the biological reflectors from the acoustic images became the challenge of detecting and classifying the biological reflectors. This also had its applications in other areas. For example, the classification of the biological reflectors into separate categories, such as fish or crustaceae, would allow those categories to be enumerated. This has the potential to enhance the management and exploitation of the resource.

There are potential limitations, as well. At the beginning of this study, it was not clear exactly what those limitations might be. Would the classification distinguish between markedly different types of organisms, such as those with exoskeletons like shrimp or those with air bladders like fish? Would the classification even go so far as to distinguish fish of different morphometries? The intention was to develop a tool-set to classify and extract biological reflectors from acoustic images. It was hoped that the development of these techniques would also serve to expose their limitations from the point of classifying biological reflectors.

The order in which the analysis of the acoustic images is approached can be broken down into four questions:

1. Can the biological reflectors be distinguished from the other signals in general acoustic images?
2. Can the biological reflectors be removed effectively from the images to allow further processing to access the fine detail.
3. Can the biological reflectors be further split into sub-classes?
4. Can the process be automated to act in an unsupervised manner?

The answers to these questions depend on the nature of the hardware and software used to investigate the water-column.

Presented in this thesis is a series of computational techniques which attempt to improve the detection of isolated pulse-like signals against a noisy background. The acoustic medium in this study was water, both fresh and saline. However, the techniques presented here are not strongly coupled to the type of medium. Assuming that the medium is acoustically isotropic, compressible, and inviscid, the acoustic signals have only a tenuous connection to the specific medium of propagation after they have been transformed into a digital image as described below. For example, media such as water and air will produce similar digital images. The language used in this investigation will be kept general to reflect this symmetry. The emphasis in this paper is the detection of objects in the acoustic image which correspond to reflectors in the medium. The focus is the digital processing of the acoustic images rather than the physical acoustic processes.

This paper is organised into three main sections. The Introduction gives a background on the project and some of the physics behind the signal processing. The second section, Investigation, develops the computational algorithms for the analysis. It is split into five subsections: Image normalisation, Detection, Classification, and Other techniques. These respectively contain the derivations and descriptions of the corresponding algorithms used to process the acoustic data. The Experimental Results section examines the application of these algorithms to a selected set of synthetic and sampled data.

1.1.1 Underwater acoustics

The use of acoustics for underwater echo-location is well established. Typically, a source is used to generate an acoustic pulse that travels through the medium. This pulse reflects from gradients in the acoustic properties of the medium. These gradients could arise from such things as biological organisms or density changes. The reflections are picked up by a receiver that records or displays the results. The results of repeated pulses are displayed, side by side, to create a fabricated image.

The fabricated image is artifice, an abstract construct. It is not a picture of a scene. It

is the time-representation of the medium in the insonification-field of the source. In reality, the two perpendicular axes of the acoustic image are both based entirely in the time-domain.

Since the acoustic wave has a finite velocity through the medium, the travel time of the reflected acoustic wave gives spatial information. The digital image is made of vectors, short-time recordings, captured over a longer time scale. Each vector presents single-line ‘snapshots’ of the insonified area. If the source or the medium is also moving uniformly in a direction perpendicular to the direction of insonification, then each vector will also have spatial coordinates, producing a spatial image.

Arguably, the most powerful techniques at detection and classification are still those that use the human brain. Presenting the data to facilitate the transfer of information is the main goal of acoustic image visualisation. For example, the preliminary investigation of one of the classification schemes required the detected object parameters to be displayed. The parameters were presented in a two-dimensional graph to determine if there was any clustering of certain classes of objects within those parameters. If more than two parameters are under investigation, the two-dimensional graph is not an effective tool. In these cases, an n -dimensional presentation of the object parameters would be more appropriate. Even a three-dimensional display of the parameters would make the examination much more effective. Unfortunately, three dimensional displays like holograms and VR headsets are expensive.

There are two main types of acoustic sources: narrow-band and broad-band. The narrow-band source transmits a pulse of a specific—or nearly specific—frequency. The broad-band source transmits a pulse containing a number or a range of frequencies. An explosive source or a spark generator are examples of a broad-band acoustic transmitter. The class of narrow-band transmitters would include resonators or piezo-electric crystals oscillating at a particular frequency. Some narrow-band transmitters can be used to produce a broad-band signal. These transmitters are swept through a range of frequencies. This is called ‘chirping.’

The receivers can also be narrow- or broad-band. Commonly, the transmitter is also used as the receiver which means that the receiver has similar frequency characteristics as the transmitter. An important feature that affects receivers and transmitters is the beam pattern. There are three main classes of receivers: single-beam, dual-beam, and array.

Single-beam transducers have only one integral receiver. This receiver can detect the incoming signal along or near one axis. The beam pattern for the single-beam transducer is determined with respect to that axis.

Dual-beam receivers are actually a type of array, but they are often used quite differently from the standard employment of array receivers, so they are treated as a separate class. These transducers have two integral receivers with concentric beam patterns. Dual-beam transducers differ from arrays in that the receiver with the smaller beam pattern determines that the signal source is near the centre of the transducer with the wider beam pattern.

This investigation is intentionally restricted to single beam transducers. Single beam systems are much less expensive and less complicated than arrays. In this way, the algorithms will be developed for the lowest common denominator. The techniques presented below may be directly extended to apply to dual beam systems. An extension to acoustic arrays, while possible, would not likely be trivial.

The modern acoustic system consists of two main components: the analogue system and the digital system. The analogue system includes the pulse generator, the transmitter, the acoustic medium, and the receiver. Physically—in this study—the acoustic medium is the water column and any objects in the water column. The medium is often treated as part of the acoustic system since this engenders a much simpler mathematical description. If a recording system is used, it may be analogue or digital. An analogue recording system may store the amplitude and phase of the received signal or it may store only the intensity.

If the recording system is analogue, an additional component must convert the analogue recording into digital data. A digital recording system would have already done this. The digital components of the acoustic system would then be applied to process the data, in

digital form.

The aim of the digital system is to reduce the acoustic data into useable information. This may be completed before the storage stage to reduce the amount of storage required, or after, using computational methods to investigate the important information from the water column. These two approaches are called pre-processing and post-processing, respectively.

1.1.2 Acoustic domain

The propagation of acoustic information through a medium is a mechanical wave phenomenon. This requires a medium inherently containing mass. Most acoustic media, at some small scale, are constituted of many, discrete, rigid masses connected through various forces. At a larger scale, this type of medium can be approximated as a continuum of mass. This allows the use of equations for waves in continuous media to describe acoustic propagation, at least for the intent of this investigation.

The mechanical wave propagation is characterised by various properties, such as density, pressure, temperature, and the relations between them. Relations such as the temporal-, spatial-, and auto-derivatives of these variables can be used to describe the transmission of energy with respect to time and space.

This thesis is concerned with acoustic waves propagating through inviscid media, specifically fresh and saline water. This focus on underwater acoustics should not invalidate the applicability of results of this work to other compressible, inviscid media such as air. Although the acoustic differences between air and water are significant, fresh and saline water are similar enough to be treated in the same fashion for data processing. The most significant exceptions are the different speed of propagation and attenuation characteristics in each medium.

This investigation attempts to detect and classify the objects distributed within the medium. In general, these objects are best described as disturbances in the acoustical properties of the medium. Acoustic waves propagating across these disturbances will be

partially transmitted and partially reflected at their boundaries. The partial reflection and transmission of the acoustic wave splits the energy between the forward- and back-scattered wave. The net acoustic energy of the propagating wave is conserved at the boundaries. Accounting for the possibility of evanescent waves, there is no net loss or gain at these boundaries.

The digital representation of the back-scattered components of a single pulse-modulated carrier is investigated below. The objects encountered are acoustic reflectors.

It is important to note that the acoustic waves travelling through the medium suffer a reduction of energy with distance from the source. This is absorption. Some of the mechanical energy of the travelling acoustic wave is converted into thermal energy, attenuating the acoustic wave and heating the medium. By contrast, spreading causes a decrease only in the energy density of the wave; total energy is conserved. Both of these processes have the additive effect of attenuating the acoustic wave by the time it reaches the receiver.

1.2 Specifics

1.2.1 Solution of pulsating sphere

A simple mathematical description of the acoustical system is required to explain some of the algorithms presented in this paper [11]. To start with, the transducer, acting as a transmitter, produces the initial acoustic pulse which propagates through the medium (see Figure 1.1). For the moment, assume the transducer acts as a pulsating sphere with a solid surface of radius a . The velocity of the surface, $v_s(t)$, can be written as a superposition of complex sinusoids:

$$v_s(t) = \int_{-\infty}^{\infty} V_s(\omega) e^{-i\omega t} d\omega, \quad (1.1)$$

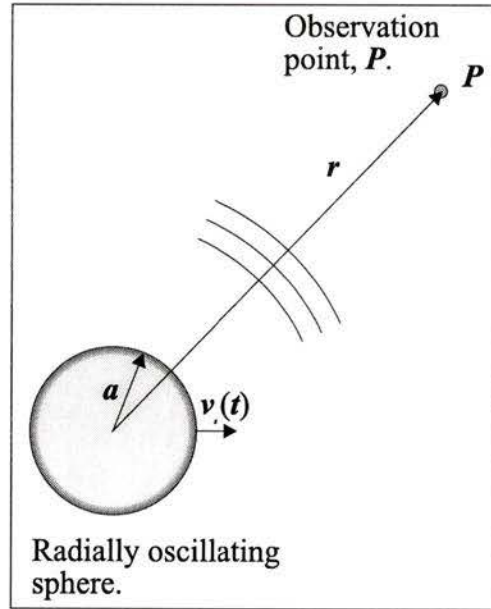


Figure 1.1: Acoustic wave generated by a radially oscillating sphere.

where

$$V_s(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} v_s(t) e^{i\omega t} dt. \quad (1.2)$$

A point about notation: frequency domain operations will be used in a number of places in this paper. Variables in the frequency domain will be referred to with an uppercase letter. The same variables in the time domain will be referred to with the letters in lowercase. The velocity component $V_s(\omega) d\omega$ produces a pressure at the surface of the sphere given by

$$P(r = a, \omega) = Z(\omega) V_s(\omega) d\omega \quad (1.3)$$

where

$$Z(\omega) = \frac{\rho_0 c}{1 + \frac{i}{ka}}, \quad (1.4)$$

is the specific acoustic impedance of the outgoing spherical wave at $r = a$ and $k = \frac{\omega}{c}$. The variable, c , is the speed of the pressure wave in the medium.

Since the complex pressure amplitude for an outward propagating spherical wave is of the form

$$P(r, \omega) = \frac{\hat{A}}{r} \quad (1.5)$$

and satisfies the boundary conditions at the boundary $r = a$, then

$$\hat{A} = ae^{-ika} Z(\omega) V_s(\omega) d\omega. \quad (1.6)$$

The complex pressure is written as

$$P(r, \omega) = \frac{a}{r} e^{ik(r-a)} Z(\omega) V_s(\omega) d\omega. \quad (1.7)$$

The pressure at a particular frequency, ω , is given by

$$P(r, \omega, t) = P(r, \omega) e^{-i\omega t} \quad (1.8)$$

$$= \frac{a}{r} Z(\omega) V_s(\omega) e^{-i(\omega t - k(r-a))} d\omega. \quad (1.9)$$

The total pressure, integrated over all frequencies, is then

$$p(r, t) = \int_{-\infty}^{\infty} \frac{a}{r} Z(\omega) V_s(\omega) e^{-i(\omega t - k(r-a))} d\omega. \quad (1.10)$$

It is expressed fully as

$$p(r, t) = \int_{-\infty}^{\infty} \frac{a}{r} e^{ik(r-a)} \left(\frac{\rho_0 c}{1 + \frac{i}{ka}} \right) V_s(\omega) e^{-i\omega t} d\omega. \quad (1.11)$$

This system is uni-directional; there is no dependence of the pressure field on the dir-

ections perpendicular to the distance, r . In practice, this result serves to provide a general idea of the action of the acoustic waves.

At the observation point, r , the pressure is represented by the complex

$$\frac{a}{r} e^{i\omega\left(\frac{r-a}{c}\right)} \left(\frac{\rho_0 c}{1 + \frac{i}{ka}} \right) V_s(\omega) d\omega \quad (1.12)$$

factor in the integrand of (1.11). Note the change in the exponent of the second factor. At the sphere the pressure is represented by

$$\left(\frac{\rho_0 c}{1 + \frac{i}{ka}} \right) V_s(\omega) d\omega \quad (1.13)$$

Close to the surface of the sphere, the near-field effects become prominent. A closer examination of $Z(\omega)$ reveals that at $ka \ll 1$, $Z(\omega) \rightarrow -i\rho_0 cka$. There is a 90° phase shift between the pressure ($\rho_0 c$) and velocity in (1.13).

Farther from the surface of the sphere, $ka \gg 1$, the near-field effects die away and $Z(\omega) \rightarrow \rho_0 c$. In the far-field an outgoing wave can be approximated as a plane wave. The equation (1.12) becomes

$$\frac{a\rho_0 c}{r} e^{i\omega\left(\frac{r-a}{c}\right)} V_s(\omega) d\omega \quad (1.14)$$

The pressure amplitude decreases with the inverse of r , independent of the frequency. The second factor is frequency dependent. It represents the phase shift of the pressure wave at the point a distance r from the center of the sphere. This phase shift is linear in the frequency domain which converts to a time shift in the time domain [12].

The time shift or delay, according to [12], is $t = \frac{r-a}{c}$. This is the length of time taken for the wave to reach the point at r . It is dependent on the speed, c .

To reduce ambiguity in the following sections, the acoustic speed, c , will hereafter be

referred to as v_w , the speed of acoustic pressure waves in the medium.

1.2.2 Physical medium

The medium contains a number, K , of acoustic reflectors. The k^{th} reflector has depth z^k measured from the surface and horizontal locations x^k and y^k , measured along perpendicular axes. The position of reflector k is referred to in bold vector notation,

$$\mathbf{x}^k = \begin{bmatrix} x^k \\ y^k \\ z^k \end{bmatrix}. \quad (1.15)$$

This is measured with respect to the frame of reference of the medium. The magnitude of \mathbf{x}^k is the equivalent of the variable, r , from the previous section. It is useful to define the position of the k^{th} reflector with respect to the transducer,

$$\hat{\mathbf{x}}^k(t) = \begin{bmatrix} x^k - v_t t \\ y^k \\ z^k \end{bmatrix}, \quad (1.16)$$

where the transducer coincides with the origin of the medium at time, $t = 0$. Both coordinate systems are chosen to have the x -axis parallel to the relative velocity, v_t .

The magnitude of the position vector from the transducer to the reflector is time-dependent. It is given by

$$\|\hat{\mathbf{x}}^k(t)\| \equiv \sqrt{(x^k - v_t t)^2 + (y^k)^2 + (z^k)^2}, \quad (1.17)$$

where $\|\cdot\|$ is the Euclidean distance.

Short time scale

Consider the case where the speed of acoustic propagation in the medium is much greater than the relative speed of the transducer with respect to the medium. The two-way travel time, from the transducer to the reflector and back to the transducer, can be described with the acoustic velocity of the medium, v_w , and the position of the reflector. The expression, with τ as the time variable, is given by

$$\begin{aligned} t_{\Delta}^k(\tau) &\equiv \frac{2}{v_w} \|\hat{\mathbf{x}}^k(\tau)\| \\ &= \frac{2}{v_w} \sqrt{(x^k - v_t \tau)^2 + (y^k)^2 + (z^k)^2}, \end{aligned} \quad (1.18)$$

The variable τ is used to represent the long time scale and to simplify the expression for the distance between the transducer and the reflector. If $v_t \ll v_w$, then (1.18) may be considered constant over the time of propagation to the reflector and back. Therefore, it can be assumed that the position of the reflector does not change significantly over the round-trip time taken by the acoustic pulse between the transducer and reflector.

Examining the system at the short time scale gives an expression for the returned signal from one acoustic pulse. In this system, the reflectors are stationary with respect to the transducer. The ideal continuous signal returned from reflector k is $s^k(t)$. This is a function of the transmitted pulse from the acoustic transducer. The pulse has been convolved with the response of the reflector into $s^k(t)$. Thus,

$$s^k(t) = p^k(t) * q^k(t) \quad (1.19)$$

where $p^k(t)$ is the signal from the transducer. It includes not only the shape of the transmitted pulse, but also the travel time for the pulse to reach the reflector and return to the transducer. It also contains the attenuation component; the reduction in signal energy caused by the distance travelled. The function $q^k(t)$ is the impulse response of the reflector

to a delta function. Since (1.19) represents a physical system, the reflected signal is causal: for $p^k(t) = 0$ and $q^k(t) = 0, \forall t < t_0$, then $s^k(t) = 0, \forall t < t_0$.

This is a linear system [13]. A change in amplitude or time-shift of $p^k(t)$ or $q^k(t)$ will have a corresponding affect on $s^k(t)$. The attenuation of the transmitted pulse due to the distance to the reflector will be represented by the function $\gamma(r)$. With these effects removed from the transmitted pulse, $p^k(t)$, (1.19) becomes

$$s^k(t) = \gamma \left(\left\| \widehat{\mathbf{x}}^k(\tau) \right\| \right) p_{\frac{1}{2}} \left[t - \frac{1}{2} t_{\Delta}^k(\tau) \right] * q^k(t), \quad (1.20)$$

where $\gamma \left(\left\| \widehat{\mathbf{x}}^k(\tau) \right\| \right)$ is the function representing the attenuation with distance at the k^{th} reflector and

$$p_{\frac{1}{2}}(t) = \gamma^{-1} \left(\left\| \widehat{\mathbf{x}}^k(\tau) \right\| \right) p^k \left[t + \frac{1}{2} t_{\Delta}^k(\tau) \right] \quad (1.21)$$

is the un-shifted, un-attenuated transmitter pulse at the reflector. If the transmitter can be expressed as a point-source, i.e. the reflector is in the far-field, then, through linearity, the attenuation of the signal from the transducer to the reflector and back to the transducer can also be expressed,

$$p(t) = \gamma^{-1} \left(\left\| \widehat{\mathbf{x}}^k(\tau) \right\| \right) p_{\frac{1}{2}} \left[t + \frac{1}{2} t_{\Delta}^k(\tau) \right] \quad (1.22)$$

$$= \gamma^{-2} \left(\left\| \widehat{\mathbf{x}}^k(\tau) \right\| \right) p^k \left[t + t_{\Delta}^k(\tau) \right], \quad (1.23)$$

and therefore

$$s^k(t) = \gamma^2 \left(\left\| \widehat{\mathbf{x}}^k(\tau) \right\| \right) p \left[t - t_{\Delta}^k(\tau) \right] * q^k(t) \quad (1.24)$$

$$= \gamma^2 \left(\left\| \widehat{\mathbf{x}}^k(\tau) \right\| \right) p \left[t - \frac{2}{v_w} \left\| \widehat{\mathbf{x}}^k(\tau) \right\| \right] * q^k(t). \quad (1.25)$$

The transducer is directional, so the amplitude of the signal arriving back at the trans-

ducer is further attenuated an amount that depends on the position of the reflector. This effect has not yet been taken into account. Typically, this dependence is a function of the angle subtended by the reflector and the vertical axis of the transducer. The beam pattern of the transducer can be expressed with an attenuation function, $\theta(\mathbf{x})$, with respect to the position, \mathbf{x} , of the reflecting source.

The total attenuation of the acoustic signal due to spreading, dissipation, and the directionality of the transducer can be gathered into one expression as a function of the position of the reflector,

$$\alpha(\mathbf{x}) \equiv \gamma^2 (\|\mathbf{x}\|) \theta(\mathbf{x}) \quad (1.26)$$

In fact, a new expression can be derived after (1.19) that incorporates all of these factors. The amplitude of the signal from the reflector after it has been received by the acoustic system is given by

$$\hat{s}^k(t) \equiv \theta[\hat{\mathbf{x}}^k(\tau)] s^k(t) \quad (1.27)$$

$$= \alpha[\hat{\mathbf{x}}^k(\tau)] p[t - t_{\Delta}^k(\tau)] * q^k(t) \quad (1.28)$$

This is the short time scale expression for a reflector in the medium.

Long time scale

The longer time scale expression follows naturally from the expression for the short time scale. The component of (1.27) which changes over a longer time is the position of the reflector, $\hat{\mathbf{x}}^k(\tau)$, hence the use of τ as the time variable. Really, (1.27) should be expressed as a function of the long time scale variable as well as the short, $\hat{s}^k(t, \tau)$. This conveys the independence of the two time scales.

In the following section, it will be apparent that the dichotomous nature of this representation maps directly to the discrete domain.

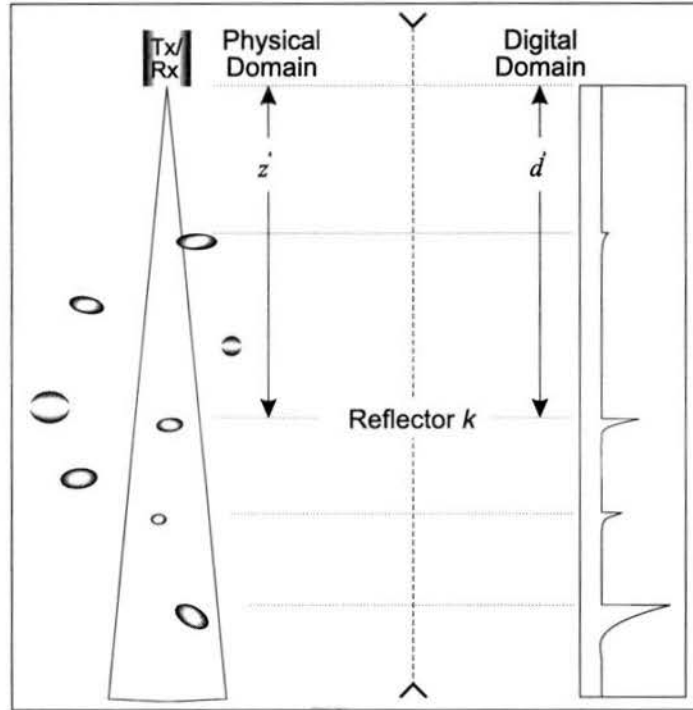


Figure 1.2: A comparison of the physical and digital representation of the reflector.

1.2.3 Digital medium

The digital acoustic image has M columns of N rows each. The columns are also referred to as vectors or echo-returns. The data contained in one vector is the result of recording the reflected information from one acoustic pulse. The reflected pulse is sampled at discrete intervals at a rate r_s , where the i^{th} digital sample is stored in the i^{th} row of the vector. Each vector is acquired at a rate r_v so that the j^{th} vector represents a snapshot of the medium at time, $\tau = \frac{j}{r_v}$. The sample i in vector j is sampled at time $t = \frac{i}{r_s} + \frac{j}{r_v}$, $\forall i, j \geq 0$. Figure 1.2 shows the comparisons between the physical and digital media.

The signal from reflector k will be

$$s_{ij}^k \equiv \mathbf{s}^k \left(\frac{i}{r_s}, \frac{j}{r_v} \right) \quad (1.29)$$

$$= \theta \left[\widehat{\mathbf{x}}^k \left(\frac{j}{r_v} \right) \right] s^k \left(\frac{i}{r_s}, \frac{j}{r_v} \right) \quad (1.30)$$

$$= \alpha \left[\widehat{\mathbf{x}}^k \left(\frac{j}{r_v} \right) \right] p \left[\frac{i}{r_s} - t_{\Delta}^k \left(\frac{j}{r_v} \right) \right] * q^k \left(\frac{i}{r_s} \right) \quad (1.31)$$

$$= \alpha \left[\widehat{\mathbf{x}}^k \left(\frac{j}{r_v} \right) \right] p \left[\frac{i - d^k \left(\frac{j}{r_v} \right)}{r_s} \right] * q^k \left(\frac{i}{r_s} \right), \quad (1.32)$$

where

$$d^k(\tau) \equiv r_s t_{\Delta}^k(\tau) \quad (1.33)$$

is the offset position of the reflector in the acoustic image vector. The time-variable, τ , has been substituted with the vector sample rate, r_v , and the vector index, j . Therefore, the full expansion of (1.32) is

$$s_{ij}^k = \theta \left[\widehat{\mathbf{x}}^k \left(\frac{j}{r_v} \right) \right] \gamma^2 \left[\left\| \widehat{\mathbf{x}}^k \left(\frac{j}{r_v} \right) \right\| \right] p \left[\frac{i}{r_s} - \frac{2}{v_w} \left\| \widehat{\mathbf{x}}^k \left(\frac{j}{r_v} \right) \right\| \right] * q^k \left(\frac{i}{r_s} \right). \quad (1.34)$$

In the acoustic image, the i^{th} element in the j^{th} vector has a value x_{ij} . This is not to be confused with the notation for the physical position of the reflectors, $\widehat{\mathbf{x}}^k(\tau)$. They are separate quantities.

The value of element ij in the digital image will be a function of the signals returned from the reflectors in the water-column and other components such as noise,

$$x_{ij} = f(i, j) = \sum_{k=1}^K s_{ij}^k + \sum_{l=1}^L n_{ij}^l. \quad (1.35)$$

It is important to note that the origin of the contribution of the k^{th} reflector, s_{ij}^k , is the same for all reflectors. This validates the sum in (1.35).

There are L components of noise present in the image. For convenience, the second sum in (1.35) hereafter shall be referred to as N_{ij} ,

$$x_{ij} = \sum_{k=1}^K s_{ij}^k + N_{ij}. \quad (1.36)$$

1.2.4 Equipment

In this investigation, the components of the underwater, acoustic measurement systems had two major sub-systems; analogue and digital. The details of the configuration of these sub-systems are described in the next two sections.

The transmitter and receiver, in this study, are combined in a single-beam transducer. This means that they are located at the same physical point.

The transmitter/receiver pair produces a pulse of mechanical vibration. An approximately square, pulse-envelope modulates a constant frequency carrier of 200 kHz. The pulse has a duration called the pulse length. The pulse insonifies the medium and reflectors, producing a forward and a back-scattered signal, the latter recorded by the receiver. In this investigation, only amplitude information is recorded.

Multiple sets of hydro-acoustic data from two systems were used in this investigation. The first system collected data from the Barrow Strait in the Arctic. It was a modified version of a commercially available Furuno system. The second system, the Water Acoustic Structure Profiler (WASP), recorded data from Kootenay Lake, the Skeena River, and the Mica Reservoir in British Columbia, and the Beaufort Sea in the Arctic. This system was unique, built by the Department of Fisheries and Oceans.

Furuno system

The first acoustic system was a Furuno colour sounder, modified to multiplex eight, single-beam, narrow-band transducers, each in sequence. Each transducer transmitted and received a 200 kHz modulated, acoustic pulse. The system was set up to provide a 350 ms

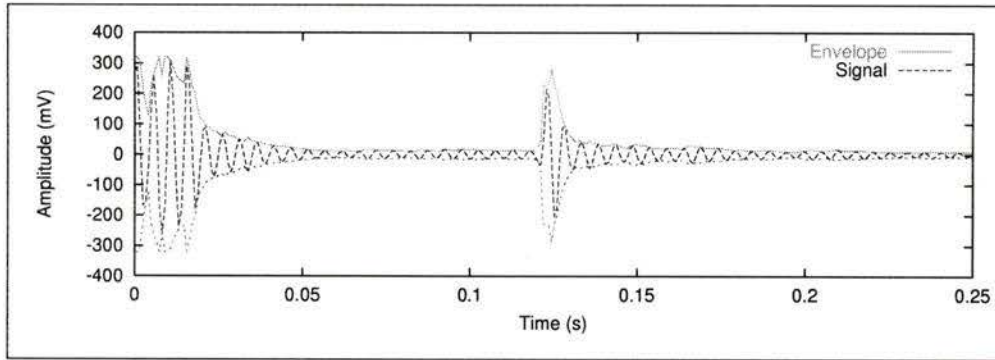


Figure 1.3: 200 kHz modulated carrier with positive and negative envelopes.

interval between the echo-returns from adjacent transducers. This created a total repetition interval of 2.77 s for the entire set of eight transducers. For each acoustic image which was derived from each transducer, this translated into $r_v = 0.357 \text{ vec}\cdot\text{s}^{-1}$. The 200 kHz carrier wave was modulated to produce an approximately square $50 \mu\text{s}$ pulse. The functional characteristics of the transducer included a -3 dB beam-width of 6.5 degrees.

Each transducer produced a square, 200 kHz modulated pulse with a $50 \mu\text{s}$ duration. The returning signal was mixed down to yield a final 8 kHz centre-frequency signal for storage on audio cassette tape. Before the signal was mixed with the 8 kHz carrier, it was low-pass filtered to limit the bandwidth to less than 8 kHz. Figure 1.3 shows an example of the returning signal with the carrier and the envelope of the signal.

The demodulation of the signal removes the carrier from the signal, leaving only the envelope. The analogue 8 kHz signal was demodulated with a RMS-power convertor and the resulting signal digitised. This process also removed the phase information from the signal.

Digitisation converts the amplitude values of the signal at discrete times into discrete digital values. Analogue values have no inherent minimum resolution. The minimum resolution for digital values is well-defined. It is the ratio of one to the maximum number of separate digits representable by the system. In a binary digital system, this is $1 : 2^n$, where

n is the number of bits. The signal was digitised as a twelve-bit binary representation at 15.625×10^3 sample·s⁻¹ and stored in this form. A Tecmar digitising system for PCs was modified to perform the demodulation and the digitisation.

The twelve-bit digital representation created a bottleneck for the Furuno system. However, characteristics of the analogue signal were exploited to increase the apparent resolution of the digital signal. An amplitude-varied gain (AVG) or amplitude compression was employed to reduce the effects of the bottleneck. The signal contained large amplitude spikes of short duration and therefore little inherent information, the larger amplitudes were represented with fewer bits. This conserved binary digits to represent the lower amplitude signals which potentially contained more important information.

The amplitude of the signal was compressed with a logarithmic amplifier before the digitisation stage. This transformed the signal,

$$\hat{s}(t) = A \ln s(t) + B, \quad (1.37)$$

where A and B are coefficients which describe the logarithmic compression.

The amplifier was adjusted to optimise the throughput of the signals under scrutiny. Once past the digitisation bottleneck, the original signal was recovered using the known AVG coefficients. The easiest way to determine the coefficients was to compress and digitise a known signal and calculate the coefficients by comparison with the uncompressed, digital version. In this way, the logarithmic amplifier was calibrated.

In certain cases, leaving the signal in its compressed digital form may provide some advantages in signal detection and classification. For example, an exponential curve, $s^k(t - t_0) = a^k e^{-b^k(t-t_0)}$, $\forall t \geq t_0$, will be transformed into a linear slope,

$$\hat{s}^k(t - t_0) = A \left[\ln a^k - b^k(t - t_0) \right] + B, \quad (1.38)$$

$$\hat{s}^k(t - t_0) = -Ab^k(t - t_0) + A \ln a^k + B. \quad (1.39)$$

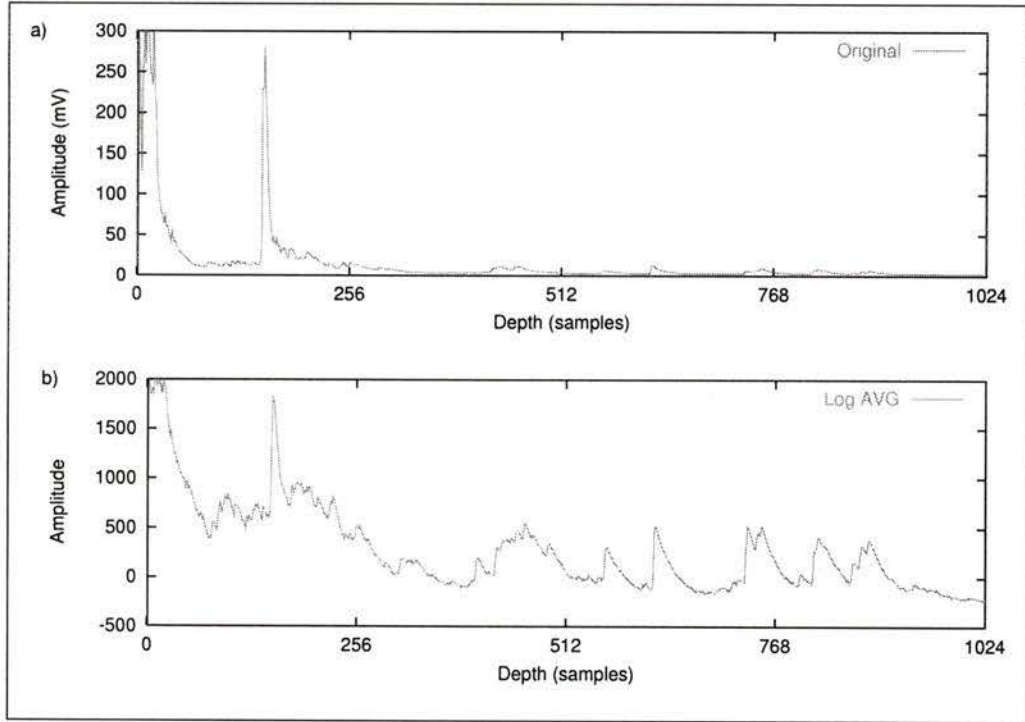


Figure 1.4: Sample image vector, a) without and b) with logarithmic compression.

Respectively, the slope and intercept of the decay of object k are,

$$\left. \frac{d\hat{s}^k(t)}{dt} \right|_{t=t_0} = -Ab^k, \quad (1.40)$$

$$\hat{s}^k(0) = A \ln a^k + B. \quad (1.41)$$

If the decay of an object signal is a sum of exponential curves, then the result of the logarithmic compression could likely be approximated by a piecewise linear function.

A sample vector from an acoustic image is shown in Figure 1.4a. The second chart in Figure 1.4 shows this vector as it would appear with a logarithmic AVG. In this case, the coefficients for the logarithmic compression are $a = 460.8$ and $b = -612$. The increase in the low-level detail is apparent in Figure 1.4b.

There was at least one caveat for this method. Only positive signals could be represented

in the logarithmic domain.

WASP

The other acoustic system which collected the data for this study was the Water Acoustic Structure Profiler (WASP). The WASP is an autonomous, submersible acoustic sampler. The transducer could be deployed on the surface of the water or at the bottom. The onboard hardware consisted of a 200 kHz transceiver, a microcontroller, a flash RAM buffer, and a hard-disk drive. The microcontroller, a 68HC11, collected the digitised data and stored it on the hard drive.

The 200 kHz modulated pulse produced by the WASP had a duration of $50 \mu\text{s}$. Unlike the Furuno system, the backscattered signals were received, demodulated, and digitised directly onboard the WASP. Eight-bit digitisation at a preselected rate was used.

Each image vector was the sum of m successive n -sample echo-returns. The values for m and n varied depending on the application. The KL930 image, examined in section 3.2, had $m = 3$, and $n = 498$. The sample rate was set at $12.1 \times 10^3 \text{ samples}\cdot\text{s}^{-1}$ which gave a maximum depth of 30 m. The summed data was stored in a 16 bit format to accommodate a large number of returns for averaging.

Chapter 2

Investigation

Typically, the process of examining objects in an acoustic image begins with the detection of the objects. The properties of the objects, such as their amplitude or location in the image, are stored or passed on to subsequent stages in the processing.

Techniques for detecting acoustic reflectors are described in section 2.1. The first of these is a normalisation technique that is used to flatten the image. This removes the dependence of the object amplitudes with distance from the transducer. It is a special version of a time-varied gain. The second technique, dPeak, is distinguished from the two remaining detection techniques. It operates in the time domain, directly on the acoustic image with no transformations. The last two approaches transform the image into special domains that allow certain objects to be detected more efficiently.

Acoustic signals can fall into two categories: desirable and undesirable. The desirable signals are objects of interest. These could be caused by biological or physical reflectors in the water-column which reflect some of the energy from the downward-travelling acoustic pulse, or possibly a coherent signal from another source. The undesirable signals tend to fall under the general rubric of noise. However, these might be valid coherent signals caused by biological or physical reflectors. It may be that they are simply objects that are not being classified.

After the chosen objects have been detected, they can be classified. This involves splitting the objects into a number of groups based on the parameters derived from the detection and other information from the image. Two main routes for classification are followed in section 2.2 below. The first approach is based on the transient properties of the object, the second on the persistent properties. The transient properties are derived from one insonification of the reflector in the water-column. This is the information present in one image vector. The persistent properties are derived from multiple insonifications of the reflector. If examining the transients can be thought of as a one-dimensional approach to classification, the persistent properties are a two-dimensional classification.

2.1 Detection

2.1.1 Image normalisation

Image normalisation is treated in this section due to its significant effect on the detection methods in the following sections. In a way, normalisation is an integral part of detection. Giving the acoustic image a uniform consistency will reduce the dependence of the detection performance on the position in the image.

Two types of normalisation were applied in this investigation. The first is an amplitude-varied gain that changes the gain of the signal with respect to amplitude. This was described in section 1.2.4. The second, a time-varied gain (TVG), adjusts the amplitude of the signal with respect to time. The application of a TVG to normalise acoustic data is not new. However, in this case the TVG is based on the reciprocal of a long term mean of the vectors in the acoustic image. This has a number of advantages.

The target strength of objects in an acoustic image usually decay with range. This is caused by the attenuation of the outgoing and returning acoustic waves plus the saturation of the receiver. An example of this can be seen in Figure 1.4a. These factors emphasise the amplitude of the signal from reflectors at early times and attenuate the amplitude at later

times. Precise corrections can be made in the case of point reflectors or volume scatterers, separately. However, combinations of the two can not be corrected simultaneously.

A TVG can be applied to the returning signal in either the hardware or software. A TVG smoothly changes the amplification of the signal with respect to time. The relation between the amplification and time is often an exponential curve as this closely approximates the attenuation of the signal caused by absorption. The other component of attenuation is that of spherical spreading, a $\frac{1}{r}$ relation. Over long times, the attenuation by absorption has a stronger effect than the attenuation caused by spherical spreading. The spherical spreading of the acoustic wave affects the signal more significantly at shorter times. The attenuation due to spreading usually dominates at the frequencies used in this investigation.

Here,

$$\hat{s}(t) = G(t)s(t), \quad (2.1)$$

where $s(t)$ is the original signal, $G(t)$ is the TVG function, and $\hat{s}(t)$ is the amplified signal. Now, if the original signal has a component of noise, $n(t)$, which is stationary, $E\{n(t)\} = \text{constant}$, and uncorrelated to the uncontaminated signal, $\bar{s}(t)$, and $s(t) = \bar{s}(t) + n(t)$, then clearly

$$\hat{n}(t) = G(t)n(t). \quad (2.2)$$

Therefore, the expected value, with respect to time, of the amplified noise component would be

$$G(t)E\{n(t)\} \neq \text{constant}. \quad (2.3)$$

In other words, the noise floor of the signal will be amplified with respect to time. The noise component of the signal becomes non-stationary with the application of a TVG.

A TVG function is usually estimated and applied before the signal is recorded, but it can also be calculated in the post-processing stage. Of the many techniques to calculate and apply a TVG, two were used in this investigation. Both methods involved the calculation of the long-term mean.

The long-term mean (LTM) is the average of a series of vectors in the acoustic image. If the image is denoted with discrete notation as, x_{ij} , where i and j are the indices for the sample and vector, respectively, the LTM is,

$$\bar{x}_i = \frac{1}{N} \sum_{j=1}^N x_{ij}, \quad (2.4)$$

where N is the number of echo-returns in the LTM.

Now, for N very large and $\text{cov}(x_{ij}, x_{ik})$ sufficiently small, $\forall j \neq k$, where

$$\text{cov}(x_{ij}, x_{ik}) = \frac{1}{M-1} \sum_{i=1}^M (x_{ij} - \bar{x}_i)(x_{ik} - \bar{x}_i), \quad (2.5)$$

and M is the number of samples in each echo-return, \bar{x}_i will be an approximation to the non-stationary components in the signal.

To apply this TVG then

$$G_i = \frac{1}{\bar{x}_i} \quad (2.6)$$

is used in a discrete version of (2.1), namely,

$$\hat{x}_{ij} = G_i x_{ij}, \quad (2.7)$$

where \hat{x}_{ij} is the value of the i^{th} sample of the j^{th} vector modified from the original value, x_{ij} , in the acoustic image. This TVG is called the divisor-LTM, or dLTM for short.

When N is not large or $\text{cov}(x_{ij}, x_{ik})$ is not sufficiently small, $\forall j \neq k$, then \bar{x}_i will not

be a good approximation to the non-stationary components in the signal. In this case, it is better to fit a smooth curve to \bar{x}_i and use the curve to calculate a new \tilde{G}_i ,

$$\tilde{G}_i \equiv \left(a e^{b i} \right)^{-1} \quad (2.8)$$

where a and b are coefficients chosen to fit the curve to \bar{x}_i . This is the exponential dLTM (edLTM). The Least Mean Square algorithm was used to find the coefficients in this investigation. It should be noted that this algorithm may not be appropriate for this application due to the potentially significant influence from objects with large peak or near zero values.

The first form of the TVG, using (2.6), has a number of interesting properties, including:

1. If $x_{ij} > 0, \forall j$ and some i , then $\hat{x}_{ij} > 0$.
2. If $\bar{x}_i < 0$, then $\text{sign}(\hat{x}_{ij}) = -\text{sign}(x_{ij}), \forall j$.

The first property states that all-positive values at some sample index i will remain all-positive after the normalisation. The second property states that a negative LTM will reflect the signal amplitude about the zero value. The latter property holds for both types of divisor-LTMs. For this reason, divisor-LTM normalisation will be used only for positive signals.

The advantage of the dLTM appears with the presence of prominent stationary reflectors in the acoustic image. This reflectors may be caused by boundary or stationary density structures of the acoustic domain. It may also be caused by other instruments, such as a current meter in the acoustic medium. These reflectors will cause the appearance of many horizontally-correlated objects with nearly the same amplitudes. These objects can mask smaller amplitude information from the image. The dLTM normalisation will effectively suppress the non-varying objects, allowing the varying objects to be detected.

An example of the dLTM normalisation is shown in Figure 2.1. The original signal is overlaid with its calculated LTM in the first part of the figure. The integration of

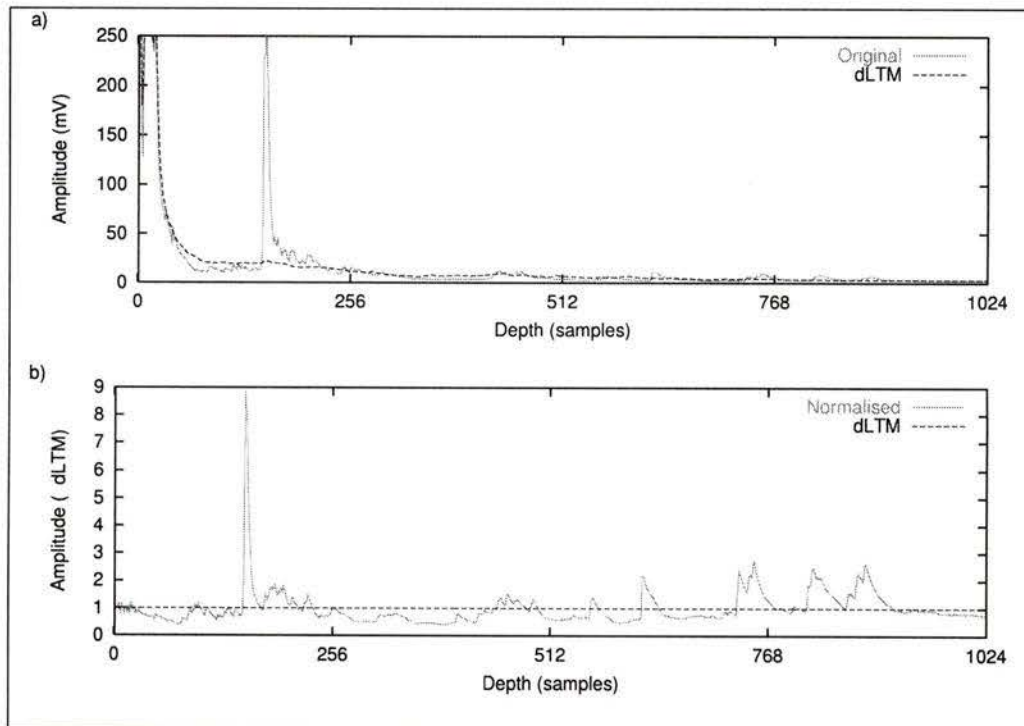


Figure 2.1: dLTM normalisation example. The original image vector, a), is normalised with dLTM to yield b).

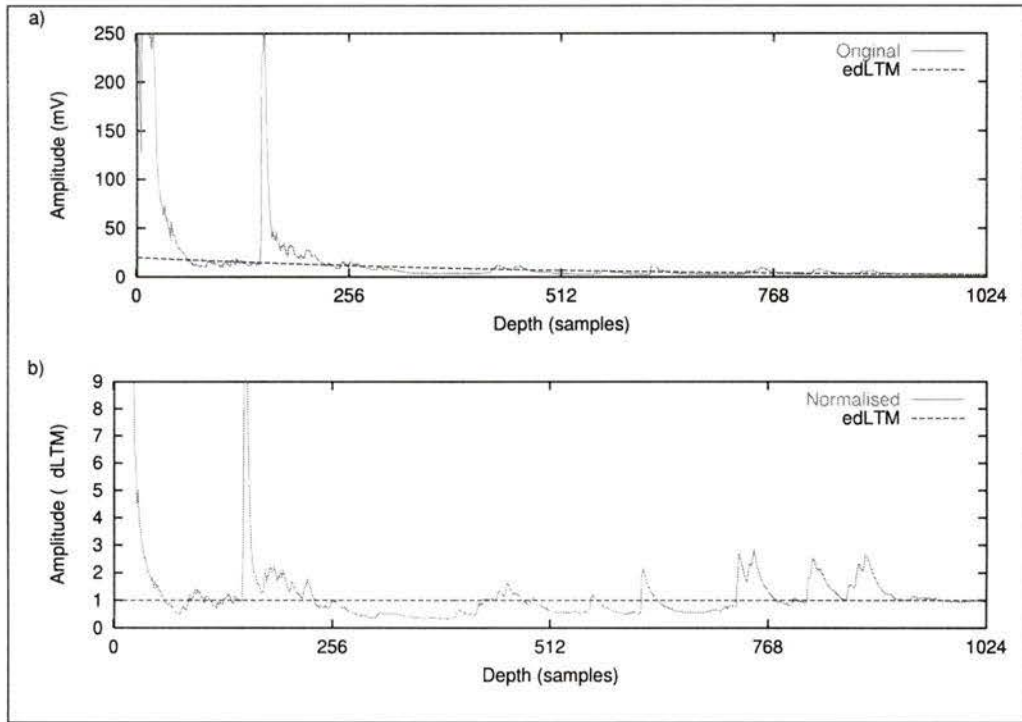


Figure 2.2: edLTM normalisation example. The original image vector, a), is normalised with edLTM to yield b).

correlated reflectors in the original image is apparent in the LTM. The second part of the figure shows the result of the normalisation. The LTM is also presented in this graph. It is a constant, 1.0. Figure 2.2 shows a corresponding example of the edLTM. An important difference here is the exponential curve (labelled edLTM) was fitted to the LTM from the hundredth sample downward. This was done to avoid having the $\frac{1}{r^2}$ -dominant portion of the signal skew the exponential-curve fit. This is the reason that the edLTM curve does not match the signal very closely for samples where $i < 100$.

In both normalised examples can be seen the result from (2.3). The noise floor becomes dominant at greater depths due to the normalisation. This is an expected consequence of this operation.

If the signal from each object changes over a much smaller time scale than the LTM and $\text{cov}(x_{ij}, x_{ik})$ is sufficiently small, the application of the LTM normalisation will not

have a significant effect on the shape or characteristics of any object signal. In other words, $\bar{x}_i \approx \text{constant}$ for small changes in i . This also states that individual object signals are transients and do not contribute to the LTM. This is not the case for Figure 2.1.

Another advantage of this normalisation is that it removes at least some of the non-stationary aspects of the signal and allows each object to be treated on par with every other object, regardless of position.

2.1.2 Peak detection

The most straightforward and most oft used method for detecting objects in the acoustic images is a search for peaks. A peak is defined in the discrete domain as a location at which there is a discontinuity in the first derivative of the signal such that the second derivative is negative. For the discrete signal, x_{ij} , this can be established with

$$\frac{\partial^- x_{ij}}{\partial i} \equiv \lim_{h \rightarrow i^-} \frac{x_{ij} - x_{hj}}{i - h}, \quad (2.9)$$

$$\frac{\partial^+ x_{ij}}{\partial i} \equiv \lim_{h \rightarrow i^+} \frac{x_{ij} - x_{hj}}{h - i}, \quad (2.10)$$

and

$$\text{sign} \frac{\partial^- x_{ij}}{\partial i} > 0, \quad (2.11)$$

$$\text{sign} \frac{\partial^+ x_{ij}}{\partial i} < 0. \quad (2.12)$$

It should be noted that the conditions (2.11) and (2.12) are strict inequalities.

This method is named dPeak. Algorithmically, it may be implemented efficiently by index-sorting the x_{ij} values in the discrete domain. An index-sort operator, $S_i \equiv f[x_{ij}]$, is defined, such that, $x_{S_i j} \geq x_{S_{i+1} j}$, $\forall i, j$. The index, i , is iterated and at each step the first and second derivatives of the signal are evaluated as in (2.11) and (2.12) above. The difference with this implementation is that the equations (2.9) have the index-sort operator,

S_i , substituted for the sample index.

The dPeak method has also been extended to detect more complex object features, as in section 2.2.1. In this study, the limitations of the peak detection required that it be augmented by other methods.

2.1.3 Deconvolution

Each object in an acoustic image has a characteristic shape or energy signature that depends on the properties of the associated reflector in the medium. Deconvolution is used to detect objects in an acoustic image by re-modelling this energy signature into a large-amplitude transient. In this case, a simple threshold might then suffice to find the objects.

Ideally, the amplitude of the signal from reflector k can be represented by a continuous function, $s^k(t)$. In discrete notation, this can be converted to s_{ij}^k as in (1.34), where i is the sample index of vector j in the acoustic image. Two prominent items appear in (1.34). The first is the function $q_k(t)$ that represents the response of the reflector to an impulse with unit area. This function contains all the information about the shape or energy signature from the object in the image. No other terms in the equation contain any information about the reflector, other than its position. The second item is the convolution operator. The acoustic pulse from the transducer is convolved with the response from the reflector to produce the signal from the reflector.

If the function $q_k(t)$ was known or estimated, the signal could be deconvolved to reveal $p(t)$. If $p(t)$ is known, the time-shift and the attenuation component, $\alpha(\mathbf{x})$, could also be recovered [13]. From (1.18), the time shift, t_{Δ}^k , gives the distance of the reflector from the transducer. To examine the deconvolution of the signal, the convolution must exist.

Given two continuous signals $g(t)$ and $h(t)$, their convolution is defined [14, 12] as

$$g(t) * h(t) \equiv \int_{-\infty}^{\infty} g(\tau)h(t - \tau)d\tau \quad (2.13)$$

If

$$\mathcal{F}[g(t)] \equiv \int_{-\infty}^{\infty} g(t)e^{-i\omega t} dt \quad (2.14)$$

$$= G(\omega) \quad (2.15)$$

and, similarly,

$$\mathcal{F}[h(t)] = H(\omega) \quad (2.16)$$

where $\mathcal{F}[\cdot]$ is the Fourier transform and ω is the angular frequency, then

$$\mathcal{F}[g(t) * h(t)] = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} g(t)h(t-\tau)d\tau \right] e^{-i\omega t} dt \quad (2.17)$$

Changing the order of integration and using the shift property [12] of the Fourier transform,

$$\mathcal{F}[h(t-\tau)] = \int_{-\infty}^{\infty} h(t-\tau)e^{-i\omega t} dt \quad (2.18)$$

$$= e^{-i\omega\tau}H(\omega), \quad (2.19)$$

gives,

$$\mathcal{F}[g(t) * h(t)] = \int_{-\infty}^{\infty} g(\tau)e^{-i\omega\tau}d\tau H(\omega) \quad (2.20)$$

$$= G(\omega)H(\omega) \quad (2.21)$$

In other words, convolution can be performed as a complex multiplication in the frequency domain. It follows then that deconvolution is a complex division in the frequency domain. With

$$s(t) = g(t) * h(t), \quad (2.22)$$

and knowing $s(t)$ and $h(t)$, the deconvolution to obtain $g(t)$ from $s(t)$ knowing $h(t)$ is

$$g(t) \equiv \mathcal{F}^{-1} \left[\frac{S(\omega) H^*(\omega)}{|H(\omega)|^2} \right], \quad (2.23)$$

where the $*$ denotes complex conjugation. In theory, this is an elegant result. What would normally have been the solution to a large number of linear equations in the time domain now becomes quite trivial in the frequency domain. In practice, the continuous Fourier transform is computationally intensive to compute. The fast Fourier transform (FFT), a discrete version of the Fourier transform, reduces the computational time from $\mathcal{O}[N^2]$ to $\mathcal{O}[N \log_2(N)]$. There are two caveats to note, however. The first, transformed functions are assumed periodic. This is particularly important for convolution and deconvolution operations performed using the FFT. End-effects of the components can contaminate the result of the deconvolution. The second consideration is that the discrete interval in the sampled data must be constant. That property is satisfied by most acoustic systems, including the application here.

From [14], the discrete Fourier transform of a discrete function, h_k , is

$$X_m = \sum_{k=1}^{M-1} x_k e^{-\frac{ikm}{M}} \quad (2.24)$$

In (1.36), the acoustic image, x_{ij} , contains the sum of the contributions from all the reflectors in the water-column. If these objects had the same impulse response, $q^k(t) \equiv q(t)$, then x_{ij} would be composed of a linear combination of $q_i \equiv q\left(\frac{i}{r_s}\right)$. The Fourier transform of (1.36) would be,

$$X_{mj} = \sum_{k=1}^K \alpha_j^k e^{-im d_j^k} P_m Q_m + N_{mj} \quad (2.25)$$

where α_j^k and d_j^k are versions of (1.26) and (1.33), respectively, redefined similarly to (1.29), converting τ to the subscript j . P_m and Q_m are the discrete Fourier transforms of p_i and

q_i , respectively. Note that (2.25) calls on the linearity and shift properties of the Fourier transform [12]. Rearranging (2.25) will give,

$$X_{mj} = P_m Q_m \sum_{k=1}^K \alpha_j^k e^{-imd_j^k} + N_{mj} \quad (2.26)$$

After the complex division by Q_m , (2.26) becomes

$$\hat{X}_{mj} = \frac{X_{mj} Q_m^*}{\|Q_m\|^2} \quad (2.27)$$

$$= P_m \sum_{k=1}^K \alpha_j^k e^{-imd_j^k} + N_{mj} \frac{N_{mj} Q_m^*}{\|Q_m\|^2} \quad (2.28)$$

The inverse Fourier transform gives,

$$\hat{x}_{ij} = \sum_{k=1}^K \alpha_j^k p_{i-d_j^k} + \hat{n}_{ij} \quad (2.29)$$

This exists for every value of the subscript j . In other words, the deconvolution is carried out vector-by-vector. If any components of Q_m are zero, the corresponding component of X_{mj} will be undefined. This is one of the drawbacks with this method.

It is interesting that the noise term has been transformed into a modified noise term that maintains its statistical properties. If n_{ij} is uncorrelated with q_i , its expectation value and other statistical properties will be relatively unchanged after the deconvolution.

Now, if p_i is known, it too can be removed by division from (1.36), leaving

$$\tilde{x}_{ij} = \sum_{k=1}^K \alpha_j^k \delta_{i-d_j^k} + \tilde{n}_{ij} \quad (2.30)$$

where δ_i is the Dirac delta-function at location i . This gives a clear picture of the location of the object in the image. It occurs at an offset d_j^k . This relies on a couple of assumptions. The most obvious is that the impulse response functions of the objects match the response

function for the deconvolution, $q_k(t) = q(t)$, or $q_i^k = q_i$.

If q_i matches only a subset of the reflectors, say $\forall k = 1..K_q$, the discrete Fourier transform of the result will look like

$$\tilde{X}_{mj} = \sum_{k=1}^{K_q} \alpha_j^k e^{-imd_j^k} + \sum_{k=K_q+1}^K \alpha_j^k e^{-imd_j^k} \frac{Q_m^k Q_m^*}{|Q_m|^2} + \frac{N_{mj} (P_m Q_m)^*}{|P_m Q_m|^2}. \quad (2.31)$$

The first term in (2.31) is similar to (2.30). The objects, q_i , will be transformed into approximations of a δ -function. The second term in (2.31) is the most crucial component. The noise term, if it is random before the deconvolution, will remain random after the deconvolution. The characteristics of the second term are not obvious. They depend on the similarity of Q_m^k and Q_m , $\forall m$.

If the objects, q_i^k , are very similar to q_i , with the exception of a dilation in time, the results from the deconvolution might be more predictable. The deconvolution models the dilation of q_i , namely $q_{(\eta i)}$, where $\eta, \forall \eta > 0; \eta \neq 1$, is the dilation factor, with some function that is dependent on the response function, q_i .

If $q_i = ae^{-bi}$, $\forall i \geq 0$ an exponentially decaying response function, $q_{(\eta i)} = ae^{-b\eta i}$ is modeled with a δ -function and an exponentially decaying response with a much reduce amplitude. More complicated q_i can produce varying effects.

Seldom is $q(t)$ known in practice. Typically, the response function is constructed from the objects of interest, identified in a sample acoustic image. This often requires the use of an alternate or non-acoustic method to determine the identity of the reflectors. The objects in the image, which correspond to those reflectors, are gathered from the image and used as the response function for the deconvolution. This response function is actually the result of the convolution of $p(t)$ and $q(t)$. The deconvolution of the image then produces a result similar to (2.30) or (2.31).

A number of ways exist to obtain a response function from the image. The difficulty is to find a response which represents at least a subset of desirable objects. The least

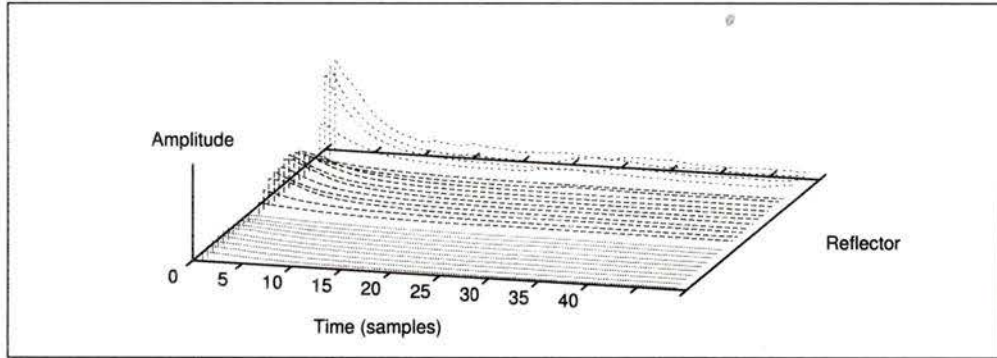


Figure 2.3: Three distinct sets of response functions from the 3B6 acoustic image. The set of objects in the foreground were the deepest, those in the background, the shallowest.

complicated method is to take the values, x_{ij} , for some range of i and some fixed value of j , corresponding to some object of interest. Figure 2.3 shows three sets of sample response functions taken from the 3B6 acoustic image. The set in the foreground has the lowest amplitudes. This set is from a reflector that was farther away from the transducer (i.e. deeper) than for the other two sets. The set in the background was taken from a shallow reflector, the nearest of the three.

An acceptable way to choose the response function would be to take one representative curve from each set, or one from all of the sets combined. An important criterion: The response function should be chosen to be relatively free from noise. It is also important to examine the frequency components of the response function to determine if any component is near zero. The best way to do this is to graph the function $\frac{Y_m^*}{\|Y_m\|^2}$, call it $\frac{1}{Y_m}$, where Y_m is the discrete complex response function in the frequency domain. Any large values above a preset limit can be set to the limit or zero. While this is not analytically correct, it is practical in most cases. The function y_i , where

$$y_i = \mathcal{F}^{-1}[Y_m]$$

should be compared to the original range of values removed from x_{ij} . This will help to

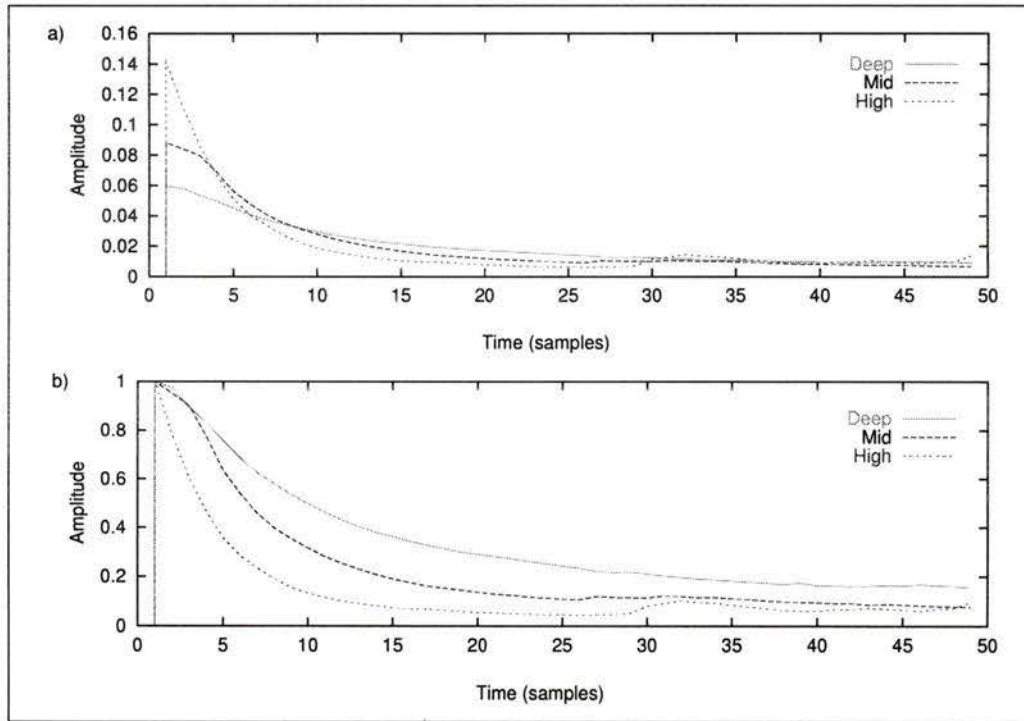


Figure 2.4: Each response function is the average of a set of extracted objects from 3B6 a) normalised to have an integral of 1.0 and b) normalised to have a peak value of 1.0.

ensure that the function will behave as expected.

Another approach averages each set of normalised objects that come from the same—or same type of—reflector. The set of objects must be aligned in the i -direction. This alignment has already been performed in both Figure 2.3 and Figure 2.4. Before the average is performed, each of the objects from the image must be normalised to have the same peak value, or the same integration value. The choice of normalisation is dependent upon the reflector objects. Figure 2.4 gives examples for the two methods. Figure 2.4a was normalised with the value of its integral. Each curve in Figure 2.4b was divided by the peak value. Mathematically, a set of objects, y_i^k , shifted an amount h^k , amplified by g^k , and averaged

are

$$\bar{y}_i = \frac{1}{K} \sum_{k=1}^K g^k y_{i-h^k} \quad (2.32)$$

The value h^k may be the index of the peak value of y_i^k , or the index of the peak cross-correlation coefficient between the objects. Two possibilities for g^k are

$$g_1^k = \max_i (y_i^k) \quad (2.33)$$

$$g_2^k = \sum_{i=1}^M y_i^k \quad (2.34)$$

The Fourier domain deconvolution is very sensitive to noise in the response function. Relatively small changes in the near-zero, frequency components of the response function will create large changes in the result. To avoid this, a response function must be chosen which has no near-zero components in the frequency domain. Certain types of discontinuities in the response function will also wreak havoc with the deconvolution. For example, an object might look similar to Figure 2.5a, an exponentially decaying spike with a linear attack, exaggerated in this case. The Fourier domain representation of the response function, Y_m , and its reciprocal are given in Figure 2.5b. There are clearly prominent poles in the complex Fourier transform of the response function which make it useless for deconvolution.

Contrast this with Figure 2.6. The only difference between Figure 2.5 and Figure 2.6 is the attack of the response function. In the latter figure, the attack, similar to the decay, is an exponential curve.

The exponential decay, by itself, may also be a useful response function. The response functions in this investigation, shown below in the experimental results (Chapter 3), used the exponential decay component only (Figure 2.7). Except for a few illustrative cases, the attack component was not required.

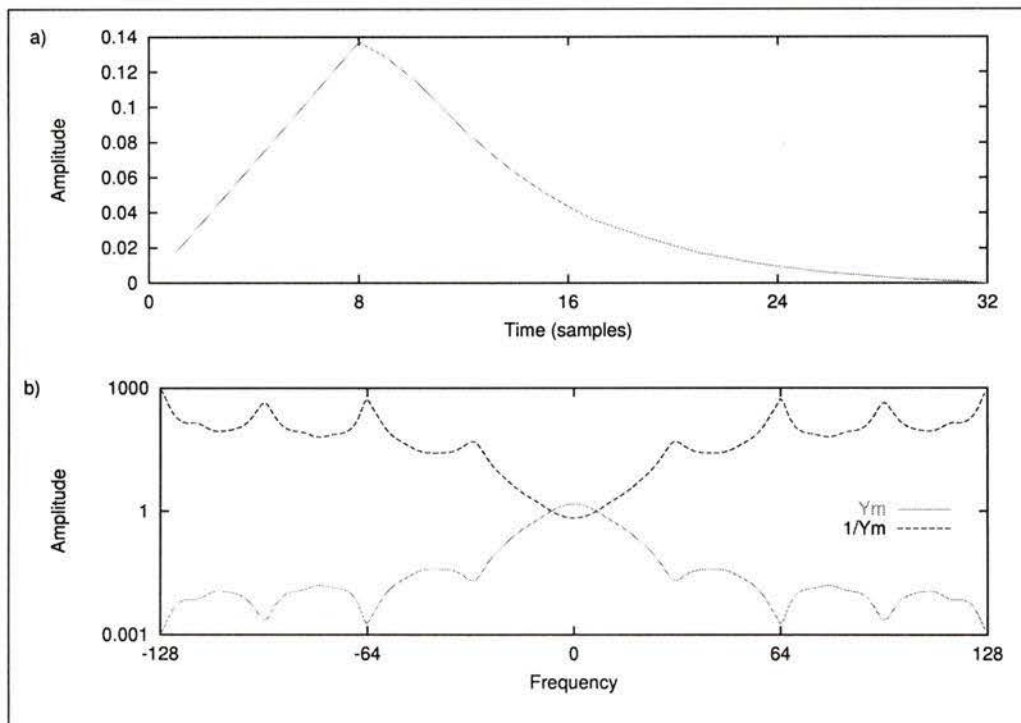


Figure 2.5: a) This response function is constructed from a linear attack and an exponential decay. b) A number of singular poles are evident in the Fourier domain.

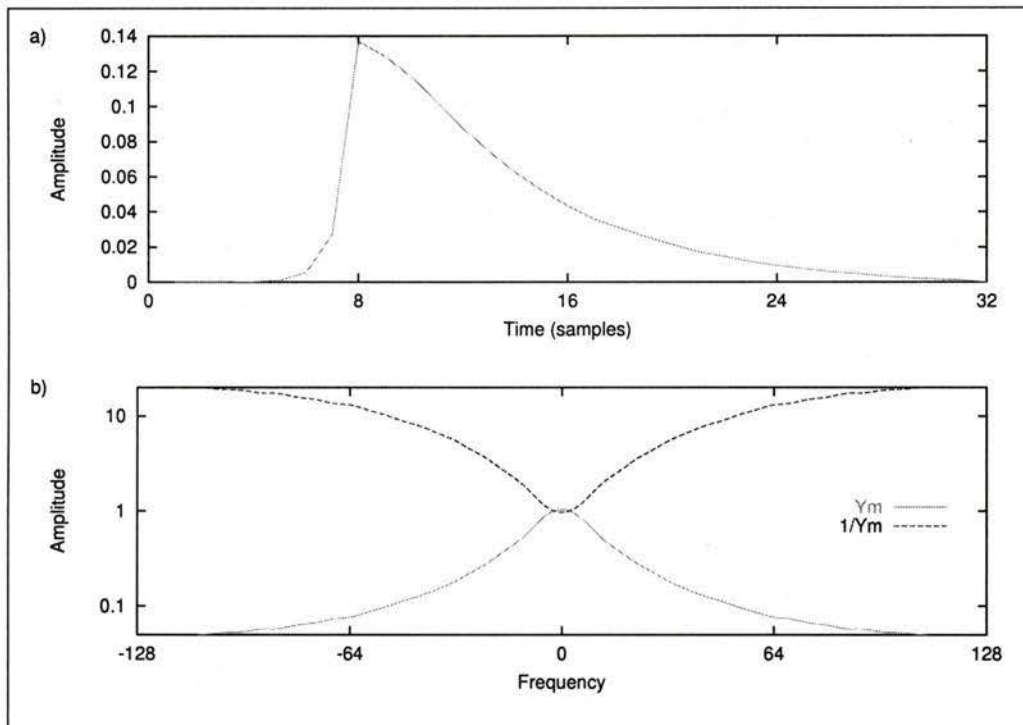


Figure 2.6: a) The response function with an exponential attack. b) The representation of a) in the Fourier domain.

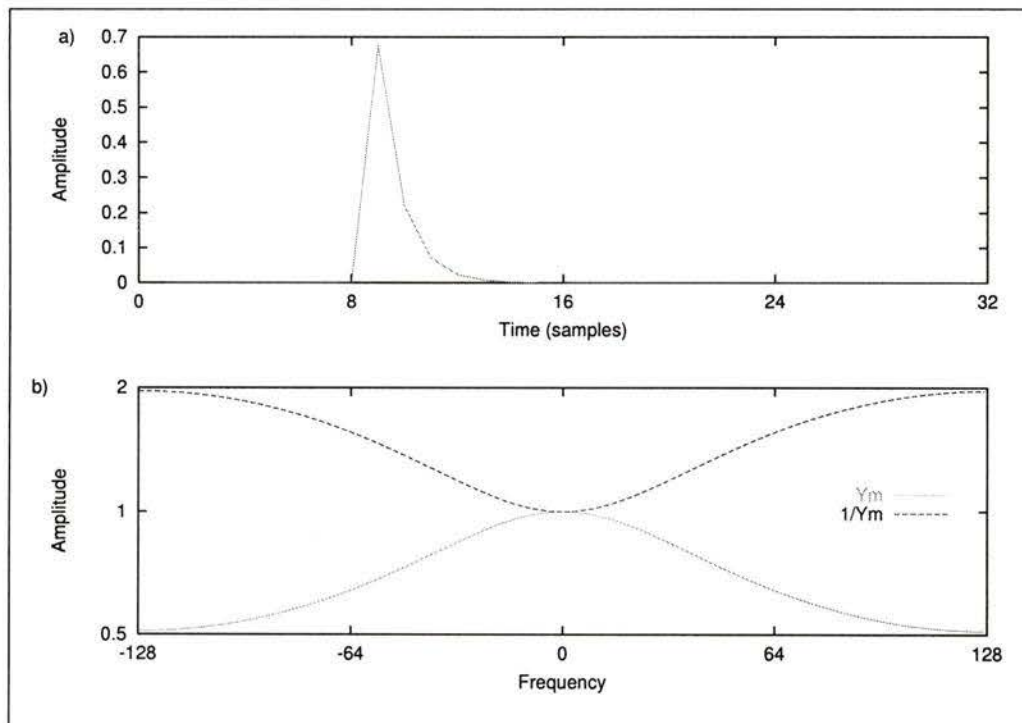


Figure 2.7: a) This response function is constructed from an exponential decay only. b) The Fourier domain of the response.

2.1.4 Wavelet detection

Deconvolution does not perform well when the objects in the image are dilated with respect to the response function. The dilation of a function is obtained with the use of a multiplicative factor on the time component. For example, $f(\eta t)$ is a dilation of $f(t)$, $\forall \eta > 0; \eta \neq 1$. Even a small difference in the dilation of the object or response function can introduce a wildly oscillating result in the deconvolution. Separate classes of objects are likely to have different characteristic durations or dilations. Even objects within the same class but of different size may have variable durations.

Wavelet transforms offer to bridge the detection gap for objects with a number of marked size classes. The basic premise of the wavelet transform is this: Model the signal with a series of shifted and dilated functions. By contrast, deconvolution describes the image with a series of shifted functions only.

The wavelet transform is implemented in a similar fashion to the Fourier transform [16, 14]. Both transforms look at the data through a time-frequency window. Both use basis functions to model the data. The difference is in the nature of the basis functions. The Fourier transform uses infinitely long sinusoids to represent the signal. The wavelet transform uses finite wavelets that are adjusted for each frequency band in the signal [3].

The Fourier transform uses the same window to examine all frequencies and times. The Fourier transform is very good at representing long-duration signals in the frequency domain, but poor at resolving the time-domain position of the frequency components. The location of the frequency components of short signals can be more accurately determined, but the frequency resolution is degraded. Even tailored Fourier methods such as the short-time Fourier transform still suffer from this type of inflexibility.

The wavelet transform follows a slightly different approach. It uses different windows, each optimised for the particular components of the signal under scrutiny. This extra degree of freedom offers the potential to improve detection of transient signals significantly. The basis functions used by the wavelet transform are localised in time and frequency. These

functions set the wavelet transform apart from the Fourier transform.

Basis functions

The choice of basis functions is likely the most significant step in using the wavelet transform. There is formally no such thing as the plain “wavelet transform.” The wavelet transform is rigorously described as the “wavelet transform with the ‘x’ wavelets.”

Wavelets are often referred to in the plural sense. The wavelet transform requires a set of wavelets to operate. Usually, these wavelets are generated inside the transform from one or two functions [14, 15]. This generating function can be a mother wavelet or a scaling function. The transform constructs the series of wavelets from the mother wavelet by dilation and shifting as the transform operates. The scaling function constructs the wavelets iteratively.

The general criteria for an orthogonal wavelet are quite restrictive. A wavelet must have a beginning and an ending. Obviously, infinite sinusoids do not fall into this category. The wavelet must also be orthogonal to dilated or shifted versions of itself.

Daubechies derived a series of mother wavelets [4, 15] which satisfy these criteria. The scaling functions she developed are the most compact representations of wavelet bases for the given number of coefficients. Daubechies’ four-coefficient mother wavelet (DAUB4) is efficient at modelling piecewise-linear signals. The DAUB6 wavelets are derived from six coefficients and provide a similar ability for the first derivative of the signal, and so on. The DAUB4 and DAUB12 wavelets are presented in Figure 2.8. They garnered most of the attention in this investigation. They were easy to implement and fast. The linear-modelling property of the DAUB4 wavelet had some other advantages, as well. These are described in section 2.3.1.

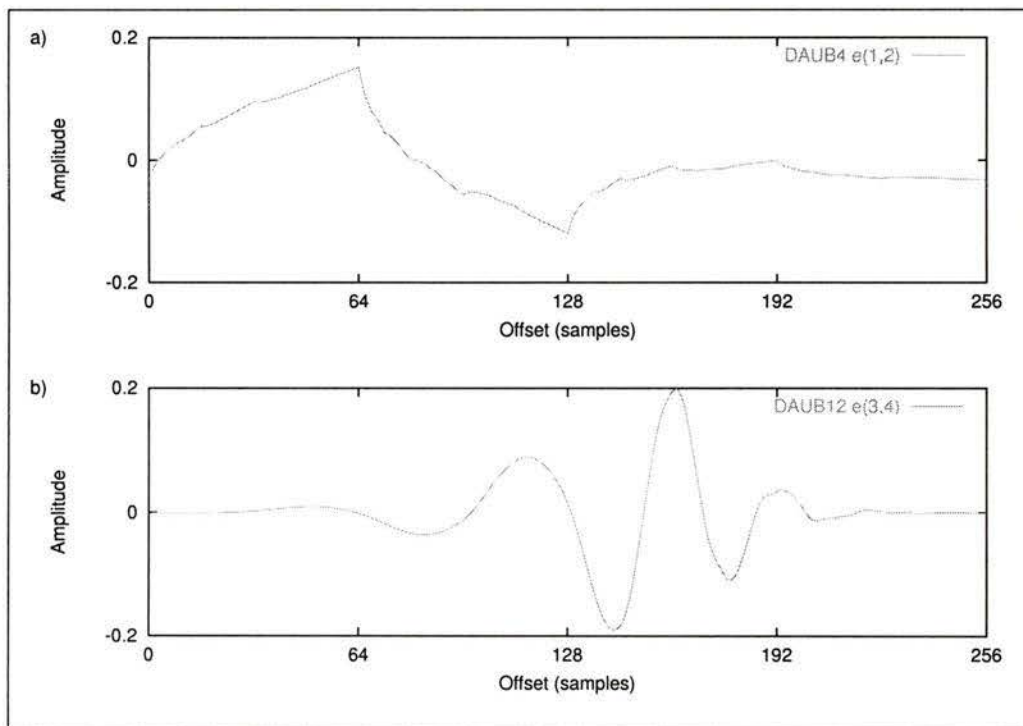


Figure 2.8: Daubechies wavelets: a) DAUB4 and b) DAUB12.

The process is then iterated for the data vector, \mathbf{s}_1 ,

$$\hat{\mathbf{s}}_{i-1} = \mathbf{W}_{i-1} \mathbf{s}_{i-1} \quad (2.39)$$

$$\mathbf{s}_i = \downarrow_{\text{odd}} \hat{\mathbf{s}}_{i-1}, \quad (2.40)$$

$$\mathbf{d}_i = \downarrow_{\text{even}} \hat{\mathbf{s}}_{i-1}, \quad (2.41)$$

At each iteration, the matrix is reduced to one quarter the size of the previous multiplication so that, if

$$\mathbf{W}_{i-1} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \quad (2.42)$$

where \mathbf{A} , \mathbf{B} , \mathbf{C} , and \mathbf{D} have the same dimensions, then

$$\mathbf{W}_i = [\mathbf{A}] \quad (2.43)$$

The process stops at (2.39) when $i = n$. At that point, there are $n + 1$ vectors that contain a number of hierarchies of information: \mathbf{d}_i where $i = 1 \dots n$, and \mathbf{s}_n . The vector \mathbf{d}_i contains the detail information at hierarchy $h = (n + 1 - i)$. The vector \mathbf{s}_n contains the scaling information for the original data vector. This is often referred to as the smooth or mother-wavelet coefficients. It can also be called the zeroth hierarchy.

The sum of the lengths of the n vectors, \mathbf{d}_i , where $i = 1 \dots n$, and \mathbf{s}_n , is 2^{n+1} . This is the same length as the original data vector. This means that the results from the DWT can be concatenated in the original vector, saving computer memory space and easing the

visual representation of the wavelet transform. Call the result of the DWT, $\hat{\mathbf{s}}$, where

$$\mathbf{s} = \begin{bmatrix} \mathbf{s}_n \\ \mathbf{d}_n \\ \vdots \\ \mathbf{d}_1 \end{bmatrix} \quad (2.44)$$

Each detail subvector in (2.44) has half the length of the preceding, lower subvector. The last two subvectors, \mathbf{s}_n and \mathbf{d}_n , contain two elements each. The subvector \mathbf{d}_1 is half the length of the original vector and contains the highest level of detail.

For the Daubechies wavelets with more coefficients, the only difference to this procedure is in the matrix \mathbf{W}_0 . The DAUB12 coefficients are contained in the matrix in a similar fashion.

Detection

If certain objects can be efficiently modelled for a particular set of wavelets, they will show up in specific hierarchies of the detail vectors, \mathbf{d}_i . The information for the objects may be spread between any number of hierarchies, or it may appear in only one. The best result, for the purpose of this study, is one that has the information for the object at one location in one hierarchy. This means that the signal from the object matched one, scaled and shifted version of the wavelets exactly. For an imperfect match, the information from the object may be smeared across many hierarchies.

Figure 2.9 shows a sample vector and its DWT. The vector contains a response function from section 2.1.3. The difference between this and the Fourier representation of the same function (Figure 2.7b) is that most of the wavelet coefficients are zero. There are only a few coefficients with non-zero values and most of these are quite conspicuous. In many cases, edge effects are also prominent at the hierarchy boundaries. These boundaries occur at $2^h + 1$ and 2^{h+1} in hierarchy h . The edge effects are often caused by mismatch between

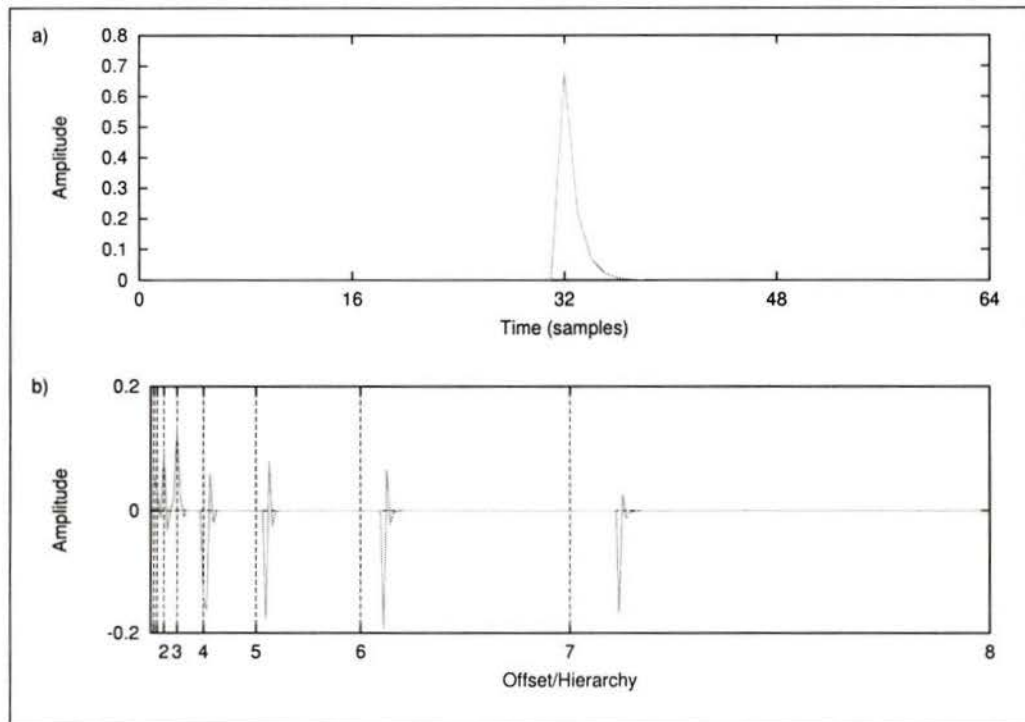


Figure 2.9: Example of the discrete wavelet transform using the DAUB4 wavelets. The graph in b) is the DWT of a).

the beginning and ending amplitudes of the original vector. Like the Fourier transform, the wavelet transform is periodic. In this example, the sample vector contains only a localised transient and therefore produces no detectable effects at the boundaries.

Two approaches make the wavelet transform useful for detection. The first approach depends on the similarity of the desirable objects to the wavelets. The similar desirable objects will model as well-defined, large-amplitude values in the wavelet domain. These spikes may be located with a threshold or peak-detection algorithm. The position in the wavelet domain can be converted to a position in the time domain. This is called the Wavelet-dPeak hybrid detection scheme.

The second approach requires only that the signals from desirable and undesirable objects are separate in the wavelet domain. If this is the case, the portions of the detail vectors that contain the undesired objects can be set to zero. The inverse wavelet transform will create an image *sans* the undesirable objects. This is called the Wavelet detection method. This new image which has only desirable objects can then be put through a regular threshold or peak-detection algorithm to find the location of the desirable objects.

The Wavelet-dPeak detection algorithm used in section 3 is a hybrid of the wavelet transform and the dPeak detection algorithm (see section 2.1.2).

A location in the time domain has a corresponding location in each hierarchy of the wavelet domain. For a compact representation of the object k in vector j of the wavelet domain, the location, ι_h , is approximated by

$$\iota_h \approx 2^h \left\lfloor \frac{d_j^k}{N} \right\rfloor + 2^{h-1} + 1, \quad (2.45)$$

where h is the hierarchy and N is the number of samples in the wavelet transform.

The dPeak algorithm searches the time domain for peaks as described in section 2.1.2. With the Wavelet detection method, dPeak is given the values ι_h as approximations to the location of the peaks. This has the potential to improve the detection. The peaks chosen

by the Wavelet-dPeak algorithm are more likely to have compact representations in the wavelet domain. This reduces the likelihood of the detection of undesirable objects such as noise.

The drawback to this method is the granularity of the detection for lower hierarchies, since (2.45) is an approximation with a precision that decreases with h . The distinction should be made at this point between the accuracy and the precision of the approximation. It is the precision of the approximation that decreases due to the reduced resolution of the lower hierarchies. This is represented by the floor operator in the first term of (2.45).

2.2 Classification

Once the objects in the image have been detected, the classification process separates them into categories. If the results of the detection also contain undesirable objects, this stage may serve to distinguish them from the desirable objects.

The objects are categorised based on the information available from the image. In the two-dimensional acoustic images described in this paper there are actually three dimensions of information: amplitude with respect to the sample and vector indices. Once the detection is complete, another type of information exists: *the location of the object*. This is also measured with respect to the sample and vector indices. Along the sample axis, the amplitude information conveys more information than just the location of the object. The amplitude *versus* sample index, is called transient information and contains such information as the acoustic response of the reflectors. The location of the object *versus* the sample and vector indices is called persistent information. It describes reflector behaviour.

2.2.1 Transient properties

The envelope of the reflector response to the acoustic pulse has been discussed in section 2.1, above. Some of the aforementioned detection techniques may be put to the further task

of classifying the objects in the image. This transient classification can be carried out in each of the three domains discussed so far: time, Fourier, and wavelet. It is possible that the detection and the classification may be carried out concurrently.

The time domain methods are easy to implement, but they all suffer from the noise problem. Noise and the objects of interest are often difficult to distinguish. They are mixed together in the time domain. The energy signatures of the desirable objects are further clouded by other objects and their own variability.

Fourier domain methods offer the potential to see through the clouding effect by other objects and noise. By separating the components of the signal into different regions in the frequency domain, the objects may be examined with reduced interference from other objects or noise. This depends on the differences in the energy signatures of the different objects in the frequency domain.

The wavelet domain offers the same advantages as the time and Fourier domain techniques with the added benefit of localisation. Objects which have a finite location in the time domain will also have a finite location in the wavelet domain. The objects are more likely to be separate from interference in the wavelet domain.

Time domain

The time domain classification requires the location of the objects from the detection stage. With this location, the objects can be classified by the shape of the spike caused by the reflector. The simplest method is to examine the peak amplitude of the spike.

It is important to note that the time domain discretisation can influence the peak amplitude. For now, assume the shape of the spike caused by the reflector has the form $y = e^{-\alpha t}$, $t \geq 0$; $y = 0$, otherwise. If the digitisation interval is ΔT , the apparent peak can vary from 1.0 to $e^{-\alpha\Delta T}$, the latter being the case where a digitally sampled point comes just before the spike front.

This indicates that measuring the spike width at some level may have some advantages

over measuring the peak value. A slightly more complex method fits a curve to the decay of the spike.

In the absence of noise, the peak amplitude can be used to obtain an indication of the size of the physical reflector. Alone, this value does not contain much other information about the type of reflector. After the objects in the image have been separated into classes using a different method, the peak value may provide an easy method for further resolving the size classes. It suffers from two main disadvantages. First, the peak value must be corrected for the position of the reflector. Off-axis reflectors or an incorrect normalisation will have a more significant effect on the peak amplitude than the size classes of the reflectors. Second, the peak value is sensitive to noise and signal aliasing in the image.

The width of the spike must be measured at a consistent amplitude level. That level may be the half-height or the base of the spike. Determining these levels for each object can be quite difficult in light of the variability of the signals from the same reflector. The height of the spike may be measured from the base to the peak of the spike or from the zero-amplitude level to the peak. The base of the spike represents the background signal level. For this to be obtained easily, the spike must be separate from all other objects in the image.

The zero-amplitude method has a disadvantage. It requires a normalised image with the non-stationary components removed. This can be difficult to achieve in practise. However, this method has the advantage of being resistant to the noise at the base of the spike.

The most straightforward way to find the base of the spike is to start from the peak and examine the slope of the amplitude on both sides. The peak detection algorithm from section 2.1.2 can be extended to provide this functionality. The base of the spike occurs where the amplitude, on one side of the peak or the other, stops decreasing. The base amplitude can be the lowest point or average of the lowest point on both sides. This is an arbitrary definition and is chosen for convenience. The disadvantage here is the variability of this method from noise or other objects near the spike. This can be estimated in a

statistical sense across many spikes by the discrepancy between the two sides of the spike base.

The spike width is the difference between the indices of the two locations. There are a number of factors which influence this value, such as

1. the physical size or geometry of the insonified reflector,
2. the frequency of the acoustic pulse,
3. more than one object at similar offsets, $t_{\Delta}^{k_1}(\tau) \approx t_{\Delta}^{k_2}(\tau)$, and
4. noise in the image.

More robust is the method of fitting a curve to the decay of the spike. The coefficients from this fit, including the quality of the fit, can be used to classify the objects into separate categories. Ideally, the decay curves would be approximated with the particle response model (PRM), described in section 2.4.1. Due to the complexity of the PRM, a simple exponential function, $f(t) = be^{-at}$, might serve to provide a rough estimate of the decay of the spike objects. However, in the preliminary stages of this investigation, the simple exponential function was found to be inappropriate for modelling the decay curves produced by the PRM.

The logarithmic compression described in section 1.2.4, has the potential to improve the transient classification of the object. If the spike decay is a sum of exponential curves as in the PRM, it will be at least a piecewise linear curve with the logarithmic amplifier. This would ease the task of fitting a curve to the decay of the object since only linear fits would be required. The caveat here is that the AVG can compress only a strictly positive signal.

If the shapes or classes of objects are unknown before the classification, an n-dimensional unsupervised clustering method can be used to determine the categories. For this method, curves are first fit to the detected objects in the image. Next, a clustering algorithm such as the generalised minimum-volume ellipsoid (GMVE) method [1, 10], can be used to determine

a number of classes. The objects can be clustered by their measured parameters including, but not limited to, peak value, spike width, position, and decay coefficients. The clustering categorises objects into separate classes. At the least, this might separate the spikes from desirable objects and those from noise into different categories.

Fourier domain

Here, classification in the Fourier domain follows the same lines of reasoning as detection in the Fourier domain. The signal is deconvolved with different response functions to convert the objects into close approximations of the δ -function. Those objects that are closest in signature to the response function will achieve a certain threshold value where they may be detected. Different response functions can be used to separate the objects into different classes.

This requires one application of the deconvolution for each category of reflectors. The response function for each category must be known *a priori* or determined from the image. This is not exactly an unsupervised approach. If the response functions for each class do not change across separate images, then an automatic procedure can be implemented, but this has not been found to be the case.

The sensitivity of this procedure to noise has inhibited its usefulness for classification. In some cases, the variation of δ -function amplitudes across classes is on the same order as the variation within the class under scrutiny.

Wavelet domain

A localised object in the acoustic image will model into localised regions in the hierarchies of the wavelet domain [3, 16]. The question is whether or not the hierarchical distribution will be unique for each class of object.

The aim here is to get the wavelet representation of the objects to become spatially distinct. If this is the case, then the objects can be resolved in the wavelet domain, possibly

more easily than the time or Fourier domains. Figure 2.9 from section 2.1.4 above, shows the transformation of a spike into the wavelet domain. As expected, the object is localised in the separate hierarchies of the wavelet domain (where hierarchy h is bounded inclusively by 2^h and $2^{h+1} - 1$). Vertical lines mark these boundaries in Figure 2.9.

In this figure, it is clear that the components of the object occur in all hierarchies. What is desired is to have the object appear in a single hierarchy only. To realise this, the choice of wavelets must efficiently represent the object at some particular dilation. This will cause the object in the image to model into fewer hierarchies in the wavelet domain.

For a relatively close match between object and wavelet, the object, though not localised to a single hierarchy, may be more prominent in one hierarchy. This would indicate that more of the energy from that object appears in one hierarchy as opposed to other hierarchies.

In the ideal case, where only the objects in question model into a single hierarchy, the classification is straightforward. All other hierarchies are set to zero and the result is inverse-transformed. All the energy from the desired class of objects will appear in the image intact. Only a small portion of the energy from the other objects or noise that occurred in the chosen hierarchy would appear in the result.

If, instead, the separate object classes did not each occur in a single hierarchy but had distinct distributions throughout a number of hierarchies, the classification might be performed by sorting the coefficients in the wavelet domain with respect to amplitude and location.

In some cases, the transient information is not enough to categorise the objects in the image. For example, in this investigation, one data set was dominated by the response function of the equipment (section 3.2.1). The object responses occurred on a much smaller time-scale than the response of one component of the acoustic system. This made the desirable objects relatively indistinguishable using the transient properties. Another means was required to classify the objects in the image.

2.2.2 Persistent properties

Object properties which exist over longer time scales than the transient information can give details about the behaviour of the reflector. For example, changes in the depth of the reflector between adjacent image vectors will describe the motion of the reflector through the medium. Another possibility is the change in the transient properties of an object over time might signal a change in the orientation or geometry of the reflector.

It is difficult to collect the persistent information for unique reflectors in the acoustic image. Gathering accurate persistent data requires the reflector identity of each object to be determined uniquely across adjacent image vectors. However, without establishing some assumptions there are no direct connections between two objects in adjacent vectors. There is no absolute certainty that any two objects were caused by the same reflector. The best that can be achieved is a probability that two objects have the same source.

Particle tracking

The position of the object in the acoustic image is related to the position of the reflector by the Euclidean norm, from (1.18) and (1.33). If the relative velocity between the transducer and the reflector is only the horizontal velocity v_t , the object in the image will follow an hyperbolic path across the image vectors (see Figure 2.10). The presence of the reflector within the beam of the transducer is detected with each insonification of the water column. Since the relative path of the reflector is constrained by physical limits, there is a non-zero probability that the reflector will be detected in consecutive echo-returns. This probability can be improved by increasing the vector sample rate, r_v , or reducing the relative velocity of the transducer.

If the reflector is represented in multiple echo-returns, its trajectory will produce a “particle track” in the acoustic image. With a sufficiently high vector sample rate or a conversely low reflector velocity, this track will be contiguous and may have a characteristic shape that can be used to classify the objects of which it is constituted.

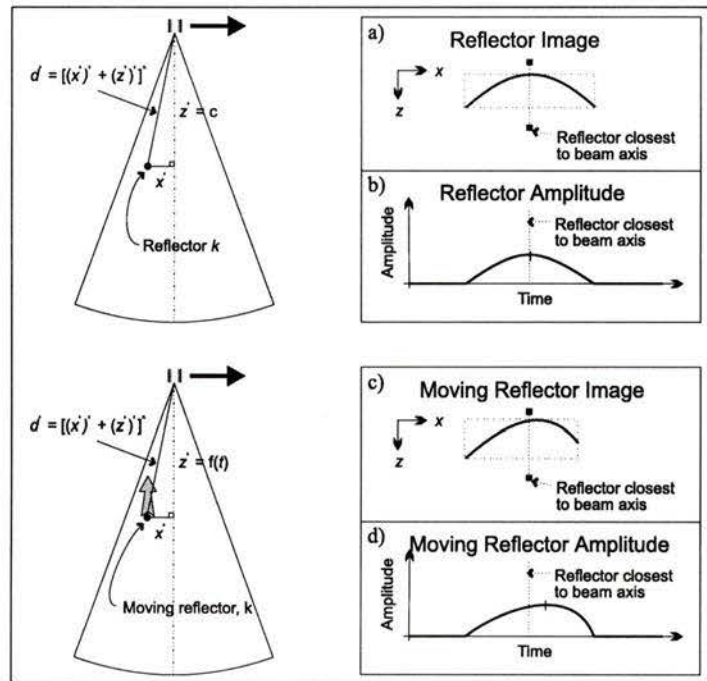


Figure 2.10: Particle properties for non-moving and moving reflectors. Reflector images appear in a) $z^k = \text{constant}$ and c) $z^k = f(t)$. Amplitude profiles are in b) and d), respectively.

The most immediate and straightforward classification is that of motility. Is the particle moving under its own power or passive within the medium? Motile reflectors will have a particle track which is different from the track of the passive reflectors nearby.

This classification has some useful properties. Passive reflectors act as indicators for movement of the medium. Vertical velocity and turbulence in the medium may be resolved using these reflectors. On the other hand, the movement of motile reflectors often gives some clue to the reflector identities.

Algorithmically, motility is difficult to establish in practice as passive reflectors from one location in the image may have particle tracks different from passive reflectors at another location. At what point is a reflector determined to be motile? At which point is it passive? The answer is often obscured by the general motion of the medium. In most cases involving water, it is not often that the medium will remain motionless over the time of the acoustic image. It is desirable to establish the background motion of the medium with respect to location so that the motile reflectors may be resolved.

The motile reflectors are not necessarily constrained by the same physical limitations as the medium. This includes local acceleration of the medium. Often, anomalous vertical accelerations or discontinuous reflector trajectories can be used to establish the motility of the reflector. Ideally, this approach may be combined with other information to identify the class of reflectors. For example, fish and shrimp are both motile reflectors but they have different maximum accelerations. Large vertical discontinuities in the particle tracks might correspond with reflectors having a large maximum acceleration. However, the opposite argument, that small vertical discontinuities are caused by reflectors with small maximum accelerations, is not true. The most that can be said for small discontinuities with respect to the general motion of the medium is that the reflectors are motile. The maximum potential acceleration of any reflector can never be established from the acoustic image.

Assume a particular reflector passes through the beam of the acoustic receiver smoothly. Also assume the received signal strength from that reflector varies with its position, $\hat{\mathbf{x}}^k(\tau)$,

only. The amplitude of the signal from the reflector will vary proportionally to $\alpha [\widehat{\mathbf{x}}^k(\tau)]$, from (1.26).

If the reflector passes through the centre of the beam along a line of constant depth, $z^k \equiv z$, the expression for the position of the object becomes

$$d_j^k = \frac{2r_s}{v_w} \left\| \widehat{\mathbf{x}}^k \left(\frac{j}{r_v} \right) \right\| \quad (2.46)$$

$$= \frac{2r_s}{v_w} \sqrt{\left(x^k - v_t \frac{j}{r_v} \right)^2 + z^2}. \quad (2.47)$$

This is the same as the expression for an hyperbolic curve,

$$y^2 = a^2 x^2 + b^2. \quad (2.48)$$

where, $b = \frac{4r_s z}{v_w}$, describes the point of closest approach to the transducer for the particle.

The attenuation of the object in the image is

$$\alpha_j^k = \theta \left[\widehat{\mathbf{x}}^k \left(\frac{j}{r_v} \right) \right] \gamma^2 \left[\frac{r_s}{v_w} \sqrt{\left(x^k - v_t \frac{j}{r_v} \right)^2 + z^2} \right]. \quad (2.49)$$

If the only non-zero velocity—excluding the acoustic velocity—in the system is v_t then nothing would move within the medium. Only the transducer would have a velocity, v_t , with respect to the medium and reflectors. The corresponding objects in the image would all trace out hyperbolic paths and shown in Figure 2.10a. This is the zero-velocity norm.

Reflectors with non-zero velocities in the medium can be compared to the zero-velocity norm. A non-zero horizontal velocity for a reflector would be manifested as a shorter particle track in the image. The particle spends less time within the beam resulting in this shorter particle track. This result is similar to a reflector path which is offset from the centre of the beam. In the offset case, the particle experiences a shorter path through the cross-section of the beam. This shorter path will also produce a shorter particle track

in the image. Potentially the best method for distinguishing between these two cases is to compare the amplitude profile of each particle track. The unique shape of the beam pattern may be used to establish the true particle track. This requires a calibrated transducer with a known beam pattern. In the absence of a calibrated transducer, statistical methods may provide a mean horizontal velocity

In the initial stages of this investigation an attempt was made to estimate the mean horizontal velocity of a large number of particles. A statistical approach examined large groups of reflectors that passed through the beam. However, the variability of the horizontal velocities and the limited number of reflectors did not produce useful results.

Vertical velocities, $z^k \equiv z(\tau)$, occur near the axis of the beam and are therefore much more apparent and more easily measured. Roughly, the change in depth of an object in the image related to the time difference between image vectors gives the vertical velocity.

A reflector possessing a vertical velocity causes the normal curvature of the particle track in Figure 2.10a, to be warped into a shape resembling that in Figure 2.10c. The two important differences between the two figures are the offset of the shallowest depth of the particle track and the difference in depths of the ends. The ends of the particle track are usually defined as the point at which the target strength of the reflector drops below the threshold of detection.

The amplitude profile of an object in the particle track is affected similarly to the depth profile. This is shown in Figure 2.10b and d. A deep reflector or a wide beam angle will cause both of these effects to be more pronounced.

The Streaks algorithm organises the detected objects into particle tracks. This algorithm uses a “brute force” method: The detected objects from the image are stepped through, one at a time. The particle tracks at each iteration are examined to determine which detected object would best extend that track. If no objects fit the criteria, the track is complete. Any objects or tracks which do not meet certain criteria are discarded.

The particle tracks were evaluated with four criteria: minimum and maximum length,

maximum gap, and minimum significance. Completed tracks shorter than the minimum length were discarded. Tracks under construction that exceeded the maximum length or that happened upon a gap which exceeded the maximum allowable gap were declared complete. A minimum significance criterion was also established. This involved the reciprocal of the χ^2 value acquired by fitting the amplitude profile of the particle track to a known function. The function used for this investigation was a second-order polynomial. The minimum significance was set at 1.0. Tracks with a lower minimum significance were deleted.

The following is the algorithm used for the Streaks method.

1. Start at the first vector in the image.
2. Initiate the active particle tracks using every reflector object in the vector.
3. Iterate the following for every vector in the image:
 - (a) Iterate for every active particle track:
 - i. If there are more than three reflectors in the current particle track, estimate the position of the next reflector in the particle track using a second-order polynomial fitted to the current track. Else, use the depth and amplitude of the last point in the track as the estimate.
 - ii. If the second-order polynomial gives a negative curvature, use a linear fit on the most recent four points to estimate the next object in the particle track.
 - iii. Calculate the distances from the objects in the current vector to the estimated location of the next object in the current track. The distances between reflector objects are based on the position only.
 - iv. If the closest object is within the maximum gap distance, add it to the particle track.
 - v. Advance to the next active particle track.

- (b) Examine all the active streaks to determine if they meet certain criteria. These criteria include minimum track-length and an amplitude profile “quality” factor. Set those that fail, inactive.
- (c) Advance to the next vector.

The distances between the objects are calculated using the depth and vector indices. If p_k is the position of the object k , then the distance between object k and object l is given by

$$d_{kl} \equiv |p_k - p_l| \tag{2.50}$$

$$= \sqrt{(i_k - i_l)^2 + (j_k - j_l)^2} \tag{2.51}$$

where object k has a depth i_k , vector j_k , and amplitude a_k

GMVE Clustering

Unsupervised clustering has been developed to the stage where it may prove useful in determining the number and segregation of size classes in the acoustic image. In the initial stages of this investigation, an unsupervised clustering method was used to classify acoustic objects.

Properties, including the coefficients of fitted decay curves (section 2.2.1), from the detected objects were clustered to provide a classification relying on the transient and persistent properties of each object. Unfortunately, the chosen curve that was fitted to the decay of the objects was not appropriate. This blurred the transient properties of the objects enough to inhibit a proper classification using this method. Generalised minimum-volume ellipsoid (GMVE) clustering is presented in this investigation without results as it is felt that it is a potentially powerful approach for unsupervised classification.

GMVE is a method for splitting the objects into separate classes, based on the object properties. This n -dimensional object clustering was developed by the author [1], based on

the work by Jolion *et al* [10]. GMVE performs best if the optimal clustering of the object parameters is ellipsoidal in shape. However, strict adherence to this rule is not required for adequate performance.

The GMVE clustering algorithm was originally developed by the author to group vanishing point intersections in photogrammetry. It has features which make it useful for the classification of reflector properties in an acoustic image. First, the GMVE clustering is an n -dimensional weighted clustering. It treats any number of parameters as a weighted object. In contrast to k -means clustering, GMVE does not require that the number of clusters be known *a priori*.

The GMVE clustering is iterative and divisive. The clustering commences by extracting the most significant cluster on the first iteration. Subsequent iterations produce successive, less significant clusters. The algorithm can be described as a series of steps:

1. Put all objects in one cluster.
2. Get a random set of $(n + 1)$ -tuples of objects ($\in \mathfrak{R}^n$) from the first cluster.
3. Using the random set, find the set of objects from the first cluster which has a distribution most similar to the smallest gaussian distribution.
4. Extract the resulting cluster of objects from the first cluster.
5. Add the cluster to the list of clusters.
6. Repeat steps 2–5 until the significance of the extracted cluster is less than 0.1 or there are no more objects.

The most complex part of the GMVE clustering is contained in the third step. The algorithm for this step is presented separately from the main algorithm:

1. For each value of h from the maximum to the minimum size:

- (a) For each random sample:
 - i. Sort the Mahalanobis distances of the objects in each random sample.
 - ii. Calculate the volume of the ellipsoid containing the fraction, h , of the summed weights.
 - (b) Find the centroid of the minimum-volume ellipsoid.
 - (c) Create a temporary cluster containing the objects which carry less than the fraction, h , of the summed weights of all objects.
2. Compare all temporary clusters to find the cluster with a distribution which is closest (in a Kolomogorov-Smirnov sense) to a gaussian distribution.
 3. Remove all other temporary clusters.

The most important concept in the GMVE clustering is the distance function for the objects, called the Mahalanobis distance function. The squared Mahalanobis distance, \hat{d}^2 between the n -dimensional location of an object, \mathbf{x} , is given by,

$$\hat{d}^2 \equiv (\mathbf{x} - \bar{\mathbf{x}})^T \mathbf{C}^{-1} (\mathbf{x} - \bar{\mathbf{x}}) \quad (2.52)$$

where $\bar{\mathbf{x}}$ is the origin—possibly the centroid of the objects—and \mathbf{C}^{-1} is the inverse of the weighted covariance matrix of the cluster,

$$\mathbf{C} = \frac{\sum_k w_k (\mathbf{x}_k - \bar{\mathbf{x}}) (\mathbf{x}_k - \bar{\mathbf{x}})^T}{\sum_k w_k - 1} \quad (2.53)$$

The values \mathbf{x}_k and w_k are the respective properties and weight of the k^{th} object of the cluster.

Rephrased, the GMVE clustering is an hierarchical method, used at different resolutions, h , to determine the optimum partitioning for that resolution. The best partitioning is

determined using the weighted covariance matrix. It is the partitioning that lies within the n -dimensional ellipsoid with the smallest volume.

This implementation uses Singular Value Decomposition (SVD) to calculate the pseudo-inverse [19] of the weighted covariance matrix, \mathbf{C} , for the cluster. This has an advantage in the case where objects are co-planar. Co-planarity occurs when the locations of the objects lie along a plane of dimension $n - p$, where $0 < p < n - 1$. The volume of the n -dimensional ellipsoid is zero in this case and (2.53) is singular. SVD removes this degeneracy through the use of a threshold related to the condition of the matrix. This automatically reduces the problem to the $(n - p)$ -dimensional clustering.

2.3 Other techniques

2.3.1 Image compression

One of the other advantages of the wavelet transform is that it offers the potential for image compression [2]. With the correct choice of wavelets, the transient information in the acoustic images can be described efficiently with this transform.

The priority for good compression is to describe the image with as few values as possible. The information in the acoustic image consists of transient information, localised in time. The wavelet transform has the potential to model this kind of information efficiently. Frisch *et al* examine this as well [9], from a different aspect.

There are two main types of compression; lossless and lossy. Lossless compression allows the image to be decompressed intact. There should be no difference between the original and the decompressed image. An exception to this rule is usually made for round-off error caused by the process of compressing and decompressing the image.

Lossy compression is very similar to lossless except that information is discarded to give higher compression ratios. The lossy compression technique is evaluated for its ability to throw away unimportant information only. The distinction between important and unim-

portant makes lossy compression subjective. It is highly dependent on the requirements for the image following decompression.

Wavelet compression begins with the transformation of the object into the wavelet domain. If the choice of wavelets is effective at modelling the information sparsely, there will be many elements in the wavelet image that have a small or zero value. Exploiting this redundancy is the key to compressing the image.

The version of lossless, wavelet image compression discussed below applies standard lossless compression techniques to the wavelet image. The improved representation of the acoustic information in the wavelet domain makes the image inherently more compressible than with standard techniques alone. This is likely not the most effective use of the wavelet transform to compress images, but it serves as a preliminary examination. The image is decompressed by reversing the process.

With the assumption that the important information will have higher amplitudes in the wavelet domain, a level can be chosen that represents a threshold for importance. Any coefficients below this value can be set to zero. Again, the wavelet image can be compressed using the standard lossless compression methods. This is lossy compression.

This assumption relies on a useful property of the wavelet transform; information is sorted by wavelet contribution with respect to amplitude. For reflected acoustic pulses, importance comes in the form of oscillatory phenomena, read wavelets. This is the reason the wavelet transform is well-suited to the compression of acoustic images.

Both lossless and lossy compressed images are recovered by the reverse of the processes above. First, the image is decompressed with the standard techniques. Next, the result is inverse-transformed back into the time domain.

A brief examination of the acoustic image compression with the wavelet transform can be found in [2].

2.4 Synthetic data

It is difficult to evaluate the performance of the detection or classification without a direct knowledge of the constituents of the acoustic images. Synthetic images were created to help evaluate the algorithms used in this investigation. The results of the analysis of the synthetic data is presented in section 3.1, below.

The construction of the synthetic images followed a progression of steps.

1. Create transient reflector information.
2. Attenuate reflector based on location.
3. Add reflectors to image vectors.
4. Apply TVG to image.
5. Add noise to image.
6. Scale and store image with appropriate data representation.

2.4.1 Particle Response Model

The reflectors that were added to the image vectors in the synthetic image were created with a particle response model (PRM). The PRM had an empirical foundation. It was originally derived to simulate the reflectors in the acoustic images produced by the Furuno acoustic system (section 1.2.4). The PRM was modified and used in this investigation to evaluate the detection and classification algorithms.

The response of reflector k , $q^k(t)$, was approximated with a linear attack followed by an exponential decay. The decay function was the sum of three exponential curves. Thus,

the response of the simulated reflector signal is

$$\tilde{q}^k(t) = \begin{cases} 0, & \text{for } t < t_0 \\ \frac{a}{t_0}(t_0 - t), & \text{for } t_0 \leq t < 0 \\ \frac{a}{3}(e^{-b_1 t} + e^{-b_2 t} + e^{-b_3 t}), & \text{for } 0 \leq t \end{cases} \quad (2.54)$$

If $t_0 = 0$, then there is no linear attack, only a discontinuity with respect to time.

From (1.29), the contribution of reflector k to the image is

$$\tilde{s}_{ij}^k = \alpha \left[\hat{\mathbf{x}}^k \left(\frac{j}{r_v} \right) \right] p \left[\frac{i - d^k \left(\frac{j}{r_v} \right)}{r_s} \right] * \tilde{q}^k \left(\frac{i}{r_s} \right), \quad (2.55)$$

where the reflector has coordinates $\hat{\mathbf{x}}^k \left(\frac{j}{r_v} \right)$.

The attenuation component $\alpha \left[\hat{\mathbf{x}}^k \left(\frac{j}{r_v} \right) \right]$ was approximated in two steps. First, the factor comprising the attenuation with respect to offset in the beam $\theta \left[\hat{\mathbf{x}}^k \left(\frac{j}{r_v} \right) \right]$ was approximated with a parabolic equation. Second, the factor for the attenuation with respect to distance was applied.

The beam pattern can be used to determine the amount of energy returned to the transducer from an off-axis reflector. The main parameter is the angle subtended by the reflector and the transducer axis, call it β , measured in radians.

The parabolic approximation was implemented thus,

$$\tilde{\theta}(\beta) \equiv 10 \left[\frac{c\beta^2}{20} \right], \quad (2.56)$$

where c is the negative parabolic coefficient determined by the beam pattern. The beam angle of the transducer is specified in radians. It is the angle at which the received and transmitted energy is attenuated by 3 dB. In this study, the beam angle, η , was used to calculate the beam pattern coefficient, $c = -3/\eta^2$. The coefficient c has units of decibels.

The angle subtended by the reflector was then

$$\beta^k(\tau) = \tan^{-1} \left(\frac{x^k - v_t \tau}{z^k} \right) \quad (2.57)$$

or

$$\beta_j^k = \tan^{-1} \left(\frac{x^k - v_t \left(\frac{j}{r_v} \right)}{z^k} \right) \quad (2.58)$$

The attenuation component due to distance was simulated with an exponential curve. The equation

$$\tilde{\gamma} \left[\hat{\mathbf{x}}^k \left(\frac{j}{r_v} \right) \right] = A e^{-B d_j^k}, \quad (2.59)$$

where A and B are the chosen coefficients, provided the factor by which the signal was next attenuated.

In the interest of computational efficiency, (2.59) was not applied as a k^{th} reflector component of \tilde{s}_{ij}^k , it was applied to the j^{th} image vector of \tilde{x}_{ij}^k as in a TVG,

$$\hat{\gamma}_i = A e^{-B i}. \quad (2.60)$$

This is not accurate as it implies that the signal from each reflector decays with distance. Attenuation components notwithstanding, (2.54) shows the time-decay of the signal from a reflector is not a function of distance. It is a function of time. The assumption here is that the signal from each the reflector, while theoretically infinite in extent, is, in a practical sense, considerably shorter than $\hat{\gamma}_i$. Another way to describe this: B from (2.60) $\ll \min_{l=1..3} \frac{b_l}{r_s}$, where b_l is the set of coefficients from (2.54).

Noise was added to the image as well,

$$N_{ij} = Cr_{ij} + D \quad (2.61)$$

where $-1 < r_{ij} < 1$ and r_{ij} varies randomly with $E\{r_{ij}\} = 0$, C is the amplitude of the noise, and D is the bias on the image. The noise had a normal distribution about a zero mean.

The synthetic acoustic image becomes

$$\tilde{x}_{ij} = \sum_{k=1}^K \tilde{s}_{ij}^k + N_{ij} \quad (2.62)$$

2.4.2 Model Construction

The model was constructed by performing the operation described by (2.62). The parameters of the model described in Table 2.1 were chosen. The reflectors were given positions and velocities clustered about a set of values with a normal distribution. To describe this, the “ \pm ” notation is used, as in the table in section 3.1. The one exception to the normal distributions of the parameters was the x -axis component of the position of the reflector which had an uniform distribution.

For the purposes of increasing the realism of the model, jitter was added to the positions and velocities of the reflectors and the coefficients of the PRM. The jitter added to the model was different from the variability of the chosen parameters, described above. It was used to cause variability within each reflector rather than between separate reflectors. The use of jitter is also described with “ \pm ” notation where appropriate. In all cases the jitter had a normal distribution about the original values.

Table 2.1: List of parameters for the synthetic image model.

Var	Description
a	PRM linear attack coefficients
$b_1..b_3$	PRM exponential decay coefficients
c	2nd-order beam pattern coefficient
η	Beam angle
A	TVG maximum amplitude
B	TVG exponential decay coefficient
C	Noise amplitude
D	Image bias
$\hat{\mathbf{x}}^k$	Positions of reflectors
$\hat{\mathbf{v}}^k$	Velocities of reflectors
r_s	Sample rate
r_v	Vector rate
v_t	Velocity of transducer
v_w	Acoustic velocity of water

Chapter 3

Experimental Results

The detection, classification, and extraction algorithms were evaluated using acoustic images. Two types of acoustic images were used in this study. The first set of images was synthetically derived. The remaining images were physically sampled during controlled investigations of marine and fresh-water environments. The following sections describe the images, any parameters involved in their construction, and the results of the processing.

Only certain parts of the analysis were valid for particular digital images; not all the images had the same algorithms applied. Figure 3.1 is a data-flow diagramme which shows the processes that were followed in this examination. The first level names the images that were used. In the second level, the normalisation is represented. In this investigation, only the sampled images required normalisation. The normalisation was the edLTM, described in section 2.1.1. This was required to suppress the effects of horizontally correlated amplitude information such as the transducer saturation at early times. The detection level in the diagramme shows the four different operations that were used for detecting the objects in the images.

The Deconv operation actually involved two operations. The first operation was the deconvolution from section 2.1.3. This was followed by a threshold (Thresh) or the peak-detection algorithm (dPeak) to provide the detection. The result of dPeak was passed to the

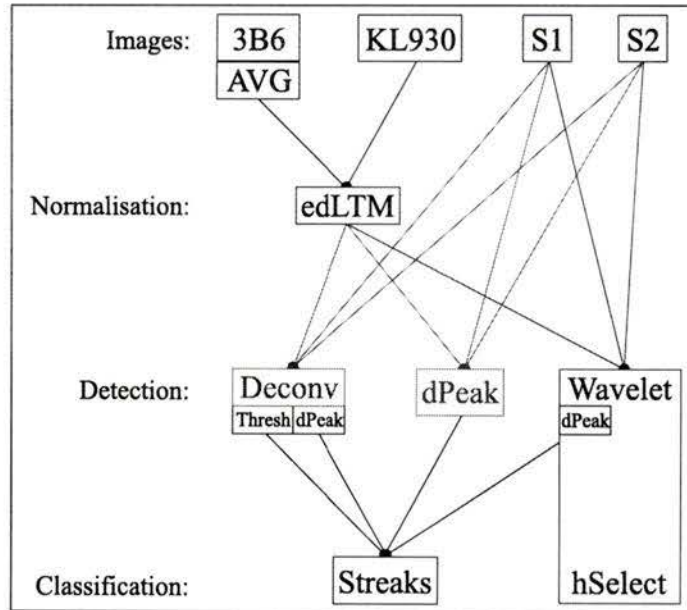


Figure 3.1: The data-flow diagramme for the evaluation of the detection, classification, and extraction algorithms. The separate levels of the diagramme represent the separate stages of the analysis.

Streaks operation in the classification level. The WASP image, KL930, was not deconvolved as the objects in the image were of short duration. The dPeak detection was defined in section 2.1.2 as the time-domain peak detection. This operation was used on both the synthetic and sampled data sets. The Wavelet detection examined the hierarchies of the wavelet domain for the largest coefficients. These were taken to represent the detected objects. The most effective implementation of the wavelet detection was a hybrid of the Wavelet and dPeak detection operations. It is shown in the flow diagramme as a dPeak sub-component of the Wavelet operation. Both wavelet detection methods are described in 2.1.4.

The Streaks classification grouped the objects found by the previous stages into particle tracks. The wavelet classification, as described in section 2.2.1, used the ability of the wavelet transform to model different classes of reflectors into different regions in the wavelet domain. The wavelet classification was integrated with the wavelet detection. The classi-

Table 3.1: Basic synthetic image, S1, with four reflectors.

Target Size (dB)	x (m)	z (m)	v_z (m·s ⁻¹)	a (m)	b_1	b_2	b_3
-40	10.0	6	0	0.05	-4.5	-3.0	-1.0
-40	20.0	7	0.01	0.05	-4.5	-3.0	-1.0
-40	30.0	20	0	0.05	-4.5	-3.0	-1.0
-40	22.0	30	0	0.05	-4.5	-3.0	-1.0

fication, hSelect, preserved a selection of hierarchies which contained the desired objects; the other hierarchies were removed. In this study, at most one hierarchy was used.

3.1 Synthetic data

The synthetic data exists to evaluate the different methods in the toolset. In the following section, the various detection algorithms are put to the test with synthetically generated acoustic images. For these images, the chosen medium was water and the chosen reflector parameters matched those of fish. Section 2.4, above, describes the meaning of the parameters used in the construction of the synthetic image.

Two synthetic images were constructed, each consisting of a total of 128 image vectors containing 1024 points. The sample frequency of 15.625×10^3 samples·s⁻¹, combined with an acoustic velocity of 1450 m·s⁻¹ in water, translated to a simulated depth of 47.5 m. The vector frequency, v_w , was set at 2.0 vectors·s⁻¹ and the relative horizontal velocity, v_t , was 1.0 m·s⁻¹.

The first image was the most basic. It contained only four reflectors and no noise. The parameters for this model appear in Table 3.1 and the image itself, in Figure 3.2a. A TVG was applied to the image to simulate the attenuation of the signal caused by spherical spreading and absorption. The attenuation coefficient was -0.05 dB·m⁻¹.

The parameters for the second image, S2, are shown in Table 3.2 (Figure 3.2b). The image was biased positively from zero at -63 dB of the maximum signal value. This image

Table 3.2: Synthetic image S2.

Target Size (dB)	\perp Dist (width)	z (m)	v_z (m·s ⁻¹)	a (m)	b_1	b_2	b_3
-37 ± 4	0.0 ± 10	25 ± 12	0.1 ± 0.02	0.1 ± 0.01	-4.5 ± 0.1	-3.0 ± 0.5	-1.0 ± 0.1
-45 ± 5	0.0 ± 4	12 ± 6	0.05 ± 0.01	0.05 ± 0.01	-5.0 ± 0.1	-3.5 ± 0.5	-1.0 ± 0.1
-60 ± 10	0.0 ± 4	6 ± 3	0	0.02 ± 0.005	-5.0 ± 0.1	-3.0 ± 0.5	-1.0 ± 0.1

was constructed with background noise. The TVG coefficient was $-0.05 \text{ dB}\cdot\text{m}^{-1}$.

The image was constructed using three basic size classes of reflectors. These were meant to simulate large, medium, and small fish in the water column. A total of 99 reflectors were placed in the image. Approximately 9 % of the total were large fish, 11 % were medium and the remainder were small fish.

The added background noise was random with an amplitude of -72 dB and a normal distribution about a zero mean.

3.1.1 Detection

Normalisation

The LTM normalisation was not required for the synthetic images. The construction of the synthetic images controlled the introduction of many of the artifacts and horizontally correlated reflectors that appeared in the sampled data. The dLTM, the normalisation without the fitted curve, was inappropriate due to this low number of reflectors in the image. The individual reflectors contributed strongly to the LTM. This combined with the lack of horizontally correlated reflectors invalidated the use of the dLTM. As well, the effect of the edLTM was negligible due to the controlled environment under which the images were created. In spite of this, it was used in place of a standard TVG which would have been applied, in any case, to remove the simulated effects of spreading and attenuation.

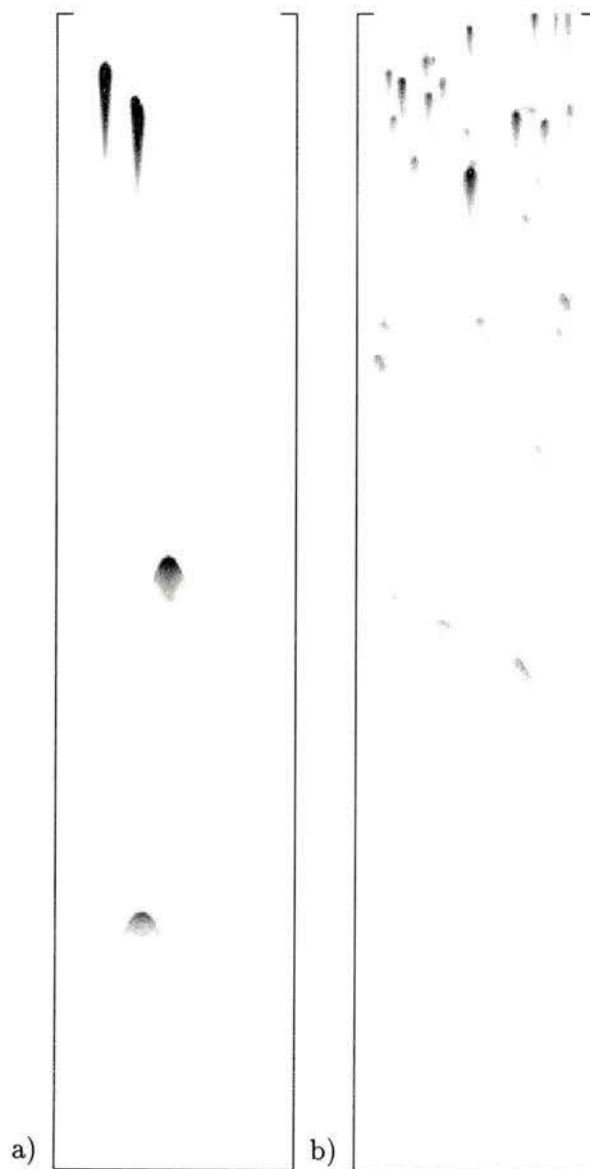


Figure 3.2: The synthetic images a) S1 and b) S2.

Time domain

Of the different techniques for detecting the objects in the time domain, the most robust was the simple peak detection described in section 2.1.2. This algorithm is given the dPeak moniker in the data flow diagramme in Figure 3.1. It is used in three cases in the detection stage: Deconv-dPeak, dPeak, and Wavelet-dPeak. It should be noted that the first 100 samples (4.6 m) of the image were omitted from the dPeak detection in all cases. The result of the peak detection on S1 and S2 is shown in Figure 3.3.

The time domain detection appears to be excellent at finding the objects in S1. The objects are visible well below the visible threshold of the original image. This was the expected result.

The complications occur in processing S2 which had added noise. For example, there is a noticeable effect caused by the tails of the reflector objects combined with the noise. The peaks caused by the noise are biased up into the same amplitude region as the peaks of the objects. The peak detection in the time domain does not handle this well. While it seems adequate at finding the peaks of the objects, it also picks up the peaks of the noise.

Fourier domain

It was relatively easy to extract the simulated response functions from S1. The positions of the reflectors were known *a priori*. They were extracted by taking x_{ij} for a range of i and a value of j which corresponded to the position of the objects. This is the same procedure that was described in section 2.1.3. The response was captured as the simulated reflector crossed the axis of the beam. The shallowest response function is shown in Figure 3.4. Since the four reflectors in the image were designed to have the same PRM coefficients, the differences between the response functions were slight. The reflectors at different depths had slightly different pulse shapes, caused by the application and imperfect removal of the TVG which was applied to the image vectors during the image construction.

S1 was deconvolved with the response from the shallowest reflector object. This object

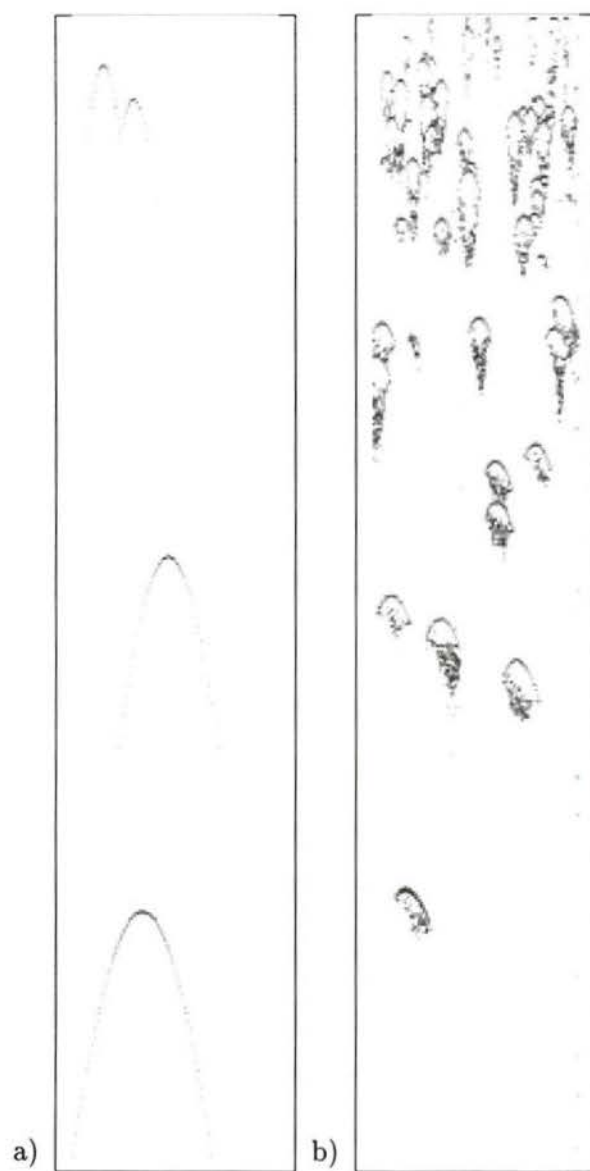


Figure 3.3: The result of the time-domain peak detection algorithm, dPeak, on images a) S1 and b) S2, respectively.

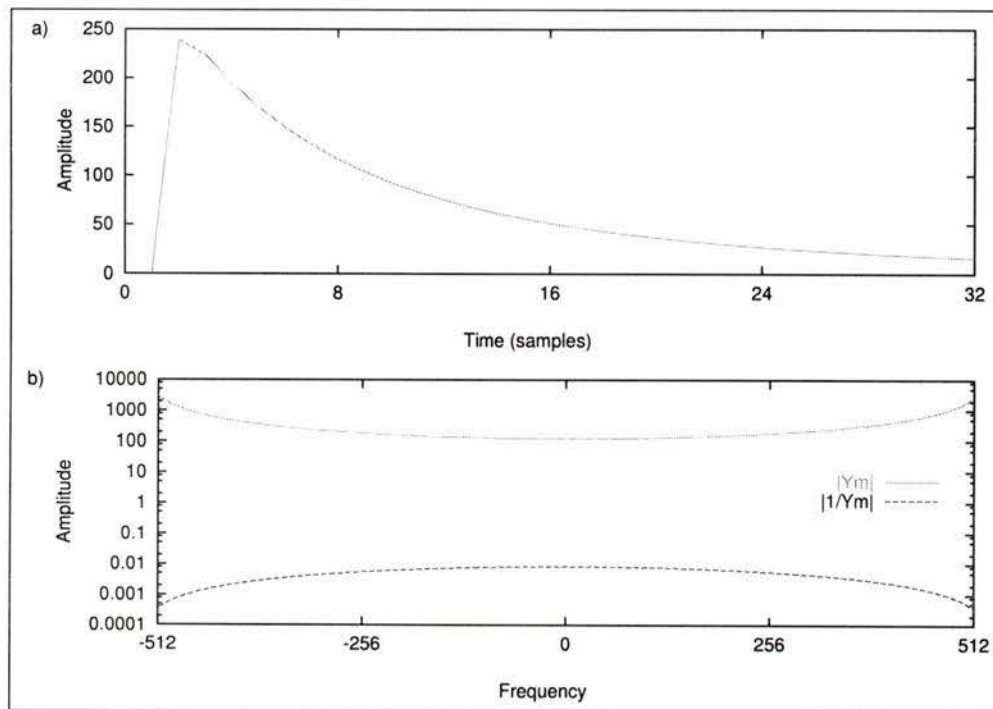


Figure 3.4: This a) response function and, its associated b) Fourier domain representation is extracted from the synthetic image S1.

was the offspring of the PRM with parameters: $a = 0.05$, $b_1 = -4.5$, $b_2 = -3.0$, $b_3 = -1.0$. This image appears in Figure 3.5a.

The change in the visible reflector properties is quite evident. Below the visible level, there are noticeable artifacts which occur as a repeated reflector with a much reduced amplitude. This echoing is caused by the the discontinuity of the response function. It is clearly visible in the result of the classification in Figure 3.10, below. The beginning of the reflector has a discontinuous jump from zero to the maximum amplitude. The infinite response of the reflector, as described by the PRM, is truncated before the deconvolution. This discontinuity picks up on the longer decay of the reflector response in the original image. In fact, the artifacts are repeated at intervals equal in length to the response function. They appear in the images (Figure 3.5) as reflector objects which are echoes of the originals.

The artifacts in the images can be attenuated with a longer response function for the deconvolution. This does not pose a problem in the synthetic images as the entire response function is known or can be calculated, but it is not practical for real data. In many cases, the response function is known for only a few data points. This sensitivity to the response is a clear and well-documented weakness of the deconvolution [14].

Image S2 was also deconvolved with the same response function as image S1, above. This image is shown in Figure 3.5b as well.

An amplitude threshold was applied to the deconvolved images from S1 and S2 to provide a simplistic detection scheme. The result of this is shown in Figure 3.6. The objects which appear in the image have been detected as desirable objects using the Thresh detection algorithm. Thresh was set to take values > 0.05 % of the largest amplitude value.

Wavelet domain

The wavelet domain representation of S1 and S2 is shown in Figure 3.7a and b, respectively. The separate hierarchies are visible in the last image, Figure 3.7c; the lowest hierarchy at the top and the highest at the bottom. This is a special image enhancement of Figure

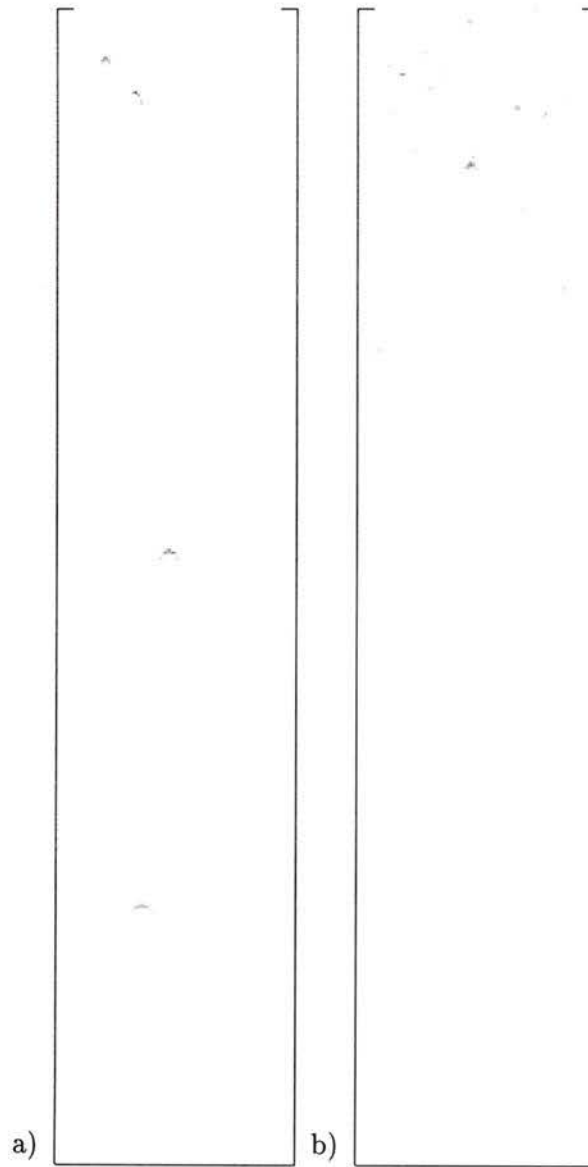


Figure 3.5: The image a) S1 and b) S2, deconvolved with the PRM response function: $a = 0.05$, $b_1 = -4.5$, $b_2 = -3.0$, $b_3 = -1.0$.

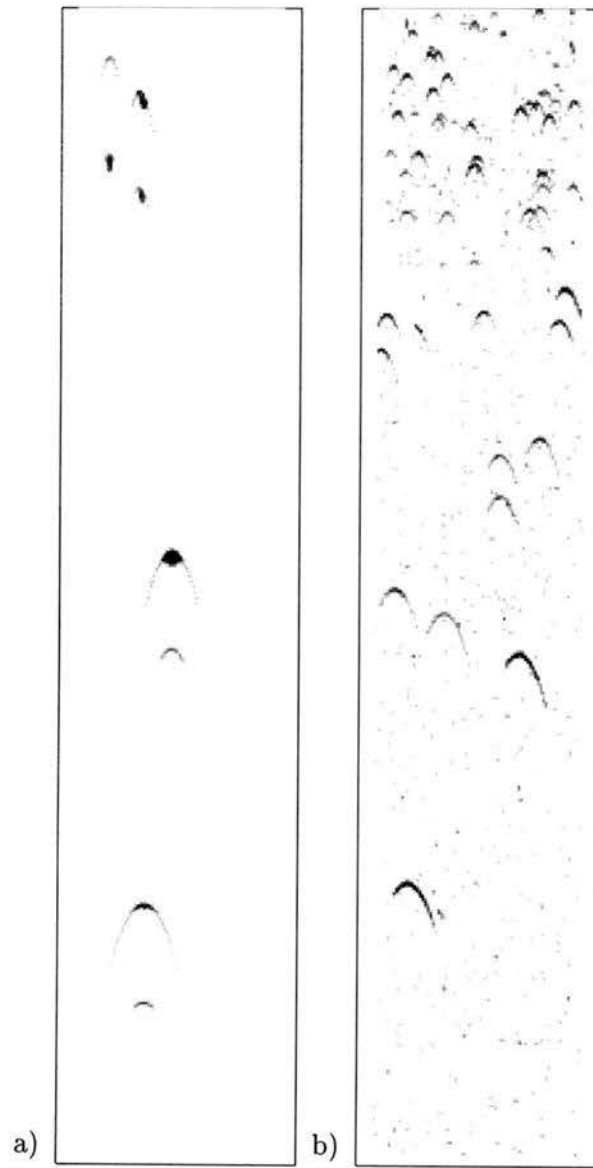


Figure 3.6: The result of the application of a threshold to the deconvolved images a) S1 and b) S2.

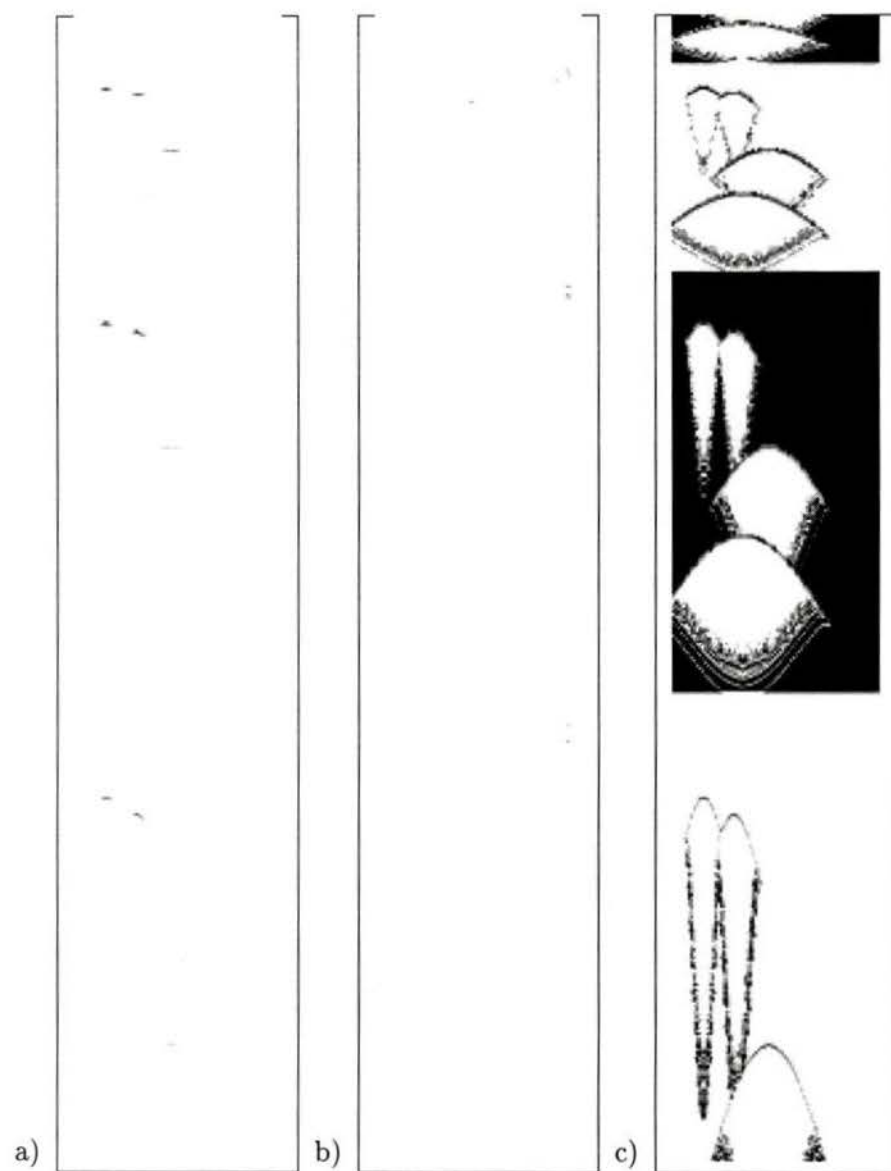


Figure 3.7: The wavelet domain representation of a) S1, b) S2, and c) an enhanced S1. In the last image the separate hierarchies are clearly visible.

3.7a, intended to demonstrate the separate hierarchies of the wavelet domain. It has no significance other than for visualisation. The outlines of the areas of influence from each reflector are clearly visible in each hierarchy. The background colour is alternated between black and white in consecutive hierarchies. In the second image, Figure 3.7b, the objects are most prominent in the lower hierarchies and the noise is less prominent when compared to the time domain of Figure 3.2b. In many cases, this hierarchical localisation can be exploited to operate on the objects apart from the noise. In this analysis, it was found that the last hierarchy actually gave the best results. While this does not appear to agree with the wavelet domain images, the discrepancy can be attributed to the manner in which the Wavelet-dPeak detection proceeds.

The reflector detection in the wavelet domain used the 9th hierarchy exclusively. The Wavelet-dPeak algorithm examines hierarchy h of the wavelet domain for the largest peaks. The algorithm uses that information to find the peaks in the time domain image. The algorithm is limited in the maximum number of objects that are detected in any vector of the acoustic image. This prevents the algorithm from interpreting low level artifacts as valid objects in the image. The strength of the wavelet algorithm lies in the ability to distinguish wavelet-like objects at any scale and amplitude in the image. The objects are sorted by importance. The limit on the number of objects can be thought of as a cut-off or threshold for the importance of the reflector objects. The two results are shown together, respectively, in Figure 3.8. The horizontal line in Figure 3.8a indicates that the wavelet algorithm has not reach the limit of the number of objects in the image. The edge effect from the blanked area in the image is a discontinuity that is detected as a valid reflector.

The horizontal lines also occur in Figure 3.7a and b, although they are not visible in the figures. They are also the result of edge effects from the wavelet transform. The hierarchy boundaries in Figure 3.7c reveal the locations of the edge effects of the boundaries. As noted above, the dPeak algorithm omitted the first 100 samples (4.6 m) from the detection. Since the sampled data images have blanked regions to improve efficiency of detection and

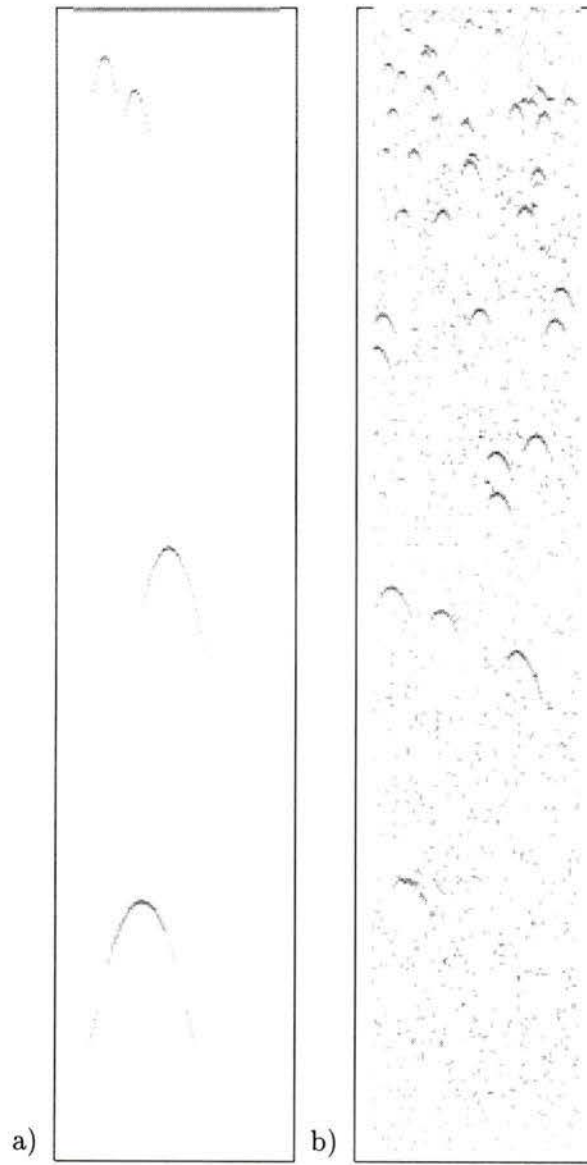


Figure 3.8: The result of the wavelet detection of images a) S1 and b) S2.

classification, the synthetic images were also blanked to examine the effects.

The use of lower hierarchies had the benefit of being more resistant to the noise in the images. The trade-off with using the lower hierarchies was the granularity of the final image. The detected peaks occurred at discrete intervals in i . The granularity is caused by the downsampling involved in performing the wavelet transform. After the detection the locations of the peaks are converted back into the time domain. This necessarily involves upsampling. An example of this granularity is given in Figure 3.9. It is manifested in the particle tracks in the image. The particle tracks appear to be constructed from a series of discrete discontinuous jumps. One way to minimise this granularity is to use the wavelet domain to detect the approximate location of the objects. This information is then used to search for the peaks in the time domain as in Wavelet-dPeak.

3.1.2 Classification

The artificially constructed images in Figures 3.10 and 3.11 represent the result of the Streaks algorithm on images S1 and S2. In each figure there are three sub-figures, labelled **a**, **b**, and **c**. The **a** sub-figure was constructed by feeding the dPeak results into the Streaks algorithm. The **b** and **c** sub-figures used the detected objects from the Deconv-dPeak and Wavelet-dPeak algorithms, respectively, in Streaks. While it is difficult to show the particle tracks in an image, the objects in the images belong to the particle tracks found by the Streaks algorithm. There are no reflector objects in the images that were not part of a particle track.

Changing the criteria such as minimum and maximum track length as well as maximum gap can have an effect on the result. If the minimum track length is set too large, valid tracks will be eliminated from the result. If it is set too small, unwanted and invalid particle tracks will start to appear. Adjusting the parameters involves a trade-off between limiting the amount of noise in the result while providing a useful detection and classification. A low maximum track length will cause constructed particle tracks to be stopped before they

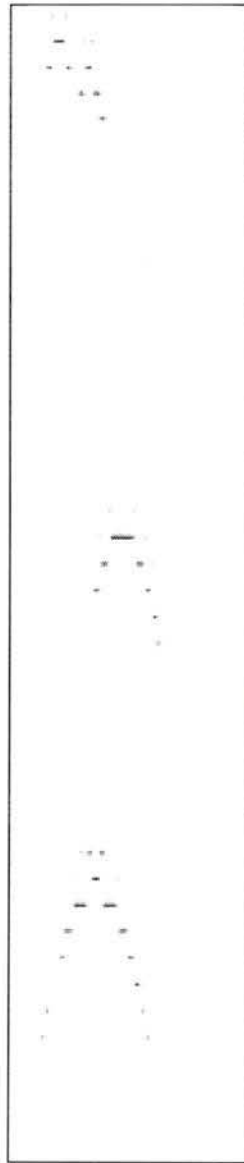


Figure 3.9: An example of the granularity of the wavelet detection.

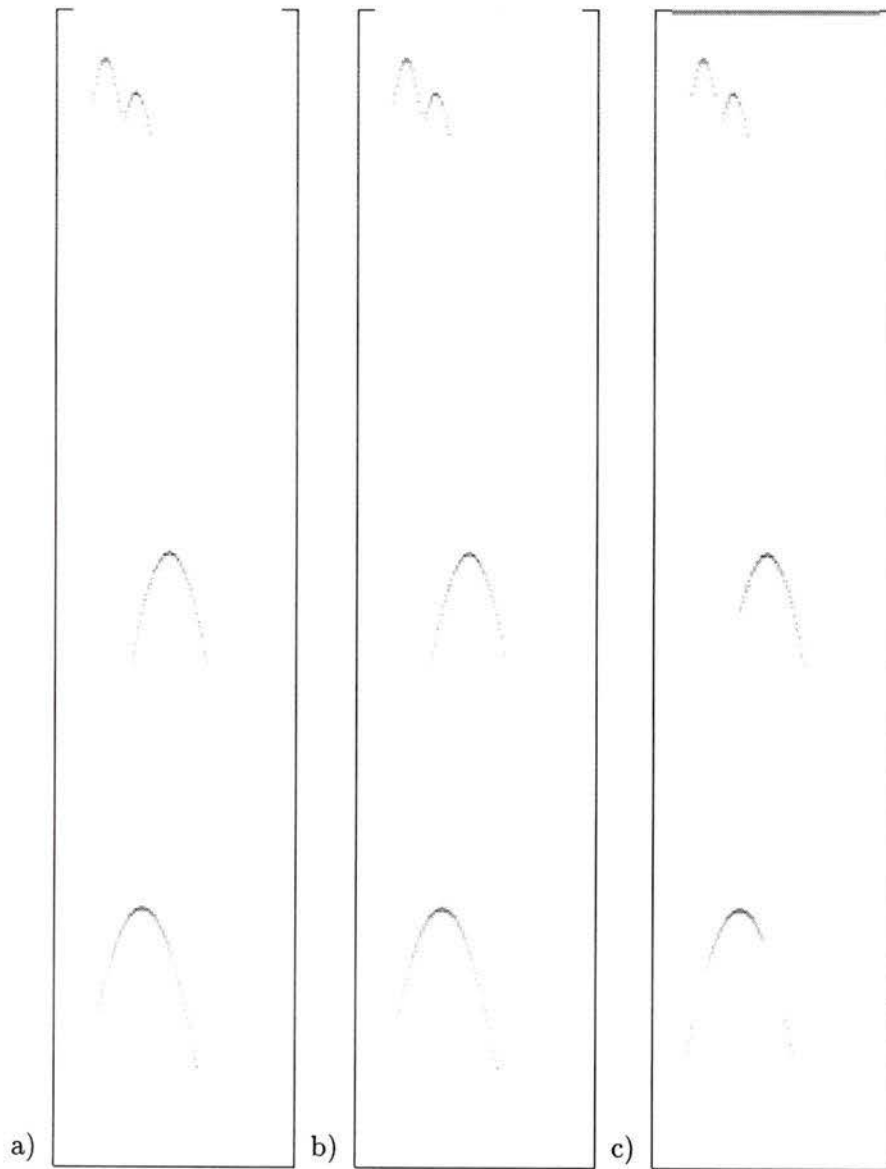


Figure 3.10: The particle tracks found in image S1. The Streaks classification used the results of the a) dPeak, b) Deconv-dPeak, and c) Wavelet-dPeak algorithms.



Figure 3.11: The particle tracks found in image S2. The Streaks classification used the results of the a) dPeak, b) Deconv-dPeak, and c) Wavelet-dPeak algorithms.

have gathered all of the objects from the same reflector. This will have the effect of splitting a single particle track into many contiguous particle tracks for the same reflector trajectory. Setting this value too high may increase the computational time required and may join two tracks which are from separate reflectors. The maximum gap is also used to determine the separation between reflectors.

In Figures 3.10 and 3.11, the minimum track length was set to 2, the maximum track length was 128, and the maximum gap in any track was allowed to be as large as 6 for Figure 3.10, but was limited to 4 for Figure 3.11. The maximum number of objects that could be detected in any one image vector was limited to 5 in Figure 3.10 and 7 in Figure 3.11.

The choice of detection also makes a difference to the classification. The Wavelet-dPeak-Streaks classification appears to be the optimal choice for the synthetic images. The dPeak-Streaks classification performs well in the absence of noise, but is confused by the peaks which occur in the tails of the spikes. The Deconv-dPeak-Streaks algorithm provides a mediocre performance, influenced dramatically by the edge effects of the deconvolution.

The hSelect classification does not seem adequate for classifying the reflectors in the synthetic images. The success of this classification hinges on the ability of the wavelets to model the desirable reflector objects in regions separate from the undesirable objects, preferably in a few regions only. In this case, the choice of wavelets did not prove adequate. The DAUB4 wavelet transform of the synthetic images S1 and S2 (shown in Figure 3.7) shows this result. The desirable reflector objects appear accross many hierarchies instead being prominent in only a few.

Figure 3.12 shows the result of keeping only hierarchies 6 (Figure 3.12a) and 9 (Figure 3.12b) of image S1 and inverse transforming the result. The result is a good indication that the reflectors are not sparse in this wavelet domain. The shallowest reflector object is the most prominent. It appears in both the 6th and 9th hierarchies.

Since the choice of wavelets is almost entirely to blame, it is possible that a different

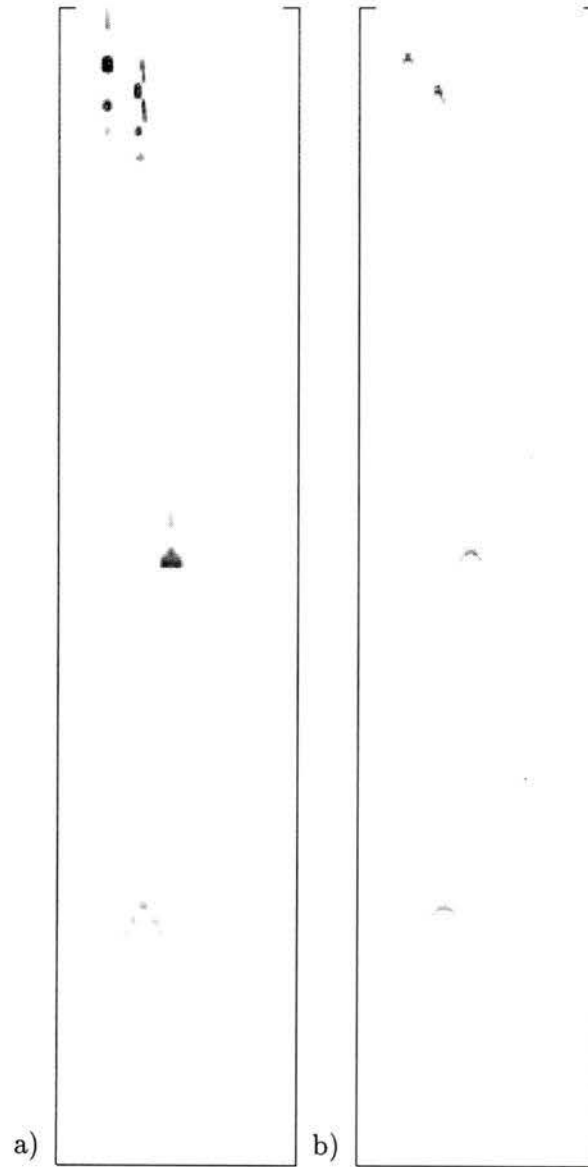


Figure 3.12: The inverse wavelet transform of hierarchies a) 6 and b) 9 from S1.

wavelet would be proved better at modelling the reflector objects. One of the most significant limitations of the DAUB4 wavelets is the fact that they are invertible. That is, the wavelet transform followed by its inverse will return the original image. No information is lost in the process.

For the purposes of the detection and classification, an invertible wavelet transform is not a strict necessity. As equation (2.45) points out, the locations of points in the wavelet domain may be approximated in the time domain with varying degrees of accuracy. Since the location of the objects in the time domain is the most important property for the detection, it is not necessary to limit the wavelets to those that are invertible only.

The classification can also benefit from no invertibility restriction. The wavelets must satisfy a number of criteria in order that they may provide a reversible transform. Some of these criteria interfere with the representation of the reflector objects in the acoustic images. They are not required for wavelets in general.

Relaxing these criteria and finding a set of wavelets that modelled the objects sparsely could significantly improve the detection and classification.

3.2 Sampled data

Two images are presented in this section. The images were chosen to provide a well-rounded view of the results of the analysis. In fact, there were a total of twelve images that were used to evaluate the algorithms in this study. Of those two were chosen to provide the most valuable information about the performance of the algorithms.

The first image, called 3B6, is a sample from the Ice Keel Experiment (see section 1.2.4). It consists of $M = 640$ image vectors each containing $N = 1024$ sampled points. The points were stored in a 16-bit AVG-compressed format. The operations in this section were performed after the removal of the AVG-compression. The image originally contained $N = 1792$ but was reduced to $N = 1024$ to allow the other images to be compared to it at

Table 3.3: Parameters for the two sampled images 3B6 and KL930.

Image	N	M	r_s (ksam·s ⁻¹)	r_v (vec·s ⁻¹)	v_w (m·s ⁻¹)	Pulse (μ s)	Beam (deg.)	Freq. (kHz)
3B6	1792	1024	15.6	0.357	1450	50	6.5	200
KL930	499	1024	12.1	0.333	1450	50	11.0	200

various stages in the processing. The other parameters appear in Table 3.3.

The other image, KL930, was collected in a freshwater environment with the WASP (section 1.2.4). The image dimensions are $N = 499$ samples and $M = 1024$ vectors. The number of samples in the image was extended to $N = 1024$ for consistency with the other images.

The Furuno system, used for the first image, stored every image echo-return that it received as one image vector. The WASP used successive sums of a number of echo-returns for each image vector. In the KL930, three successive echo-returns were summed into each image vector. The digitisation used 8-bits, but the image vectors were stored in a 16-bit format.

The images are shown in Figures 3.13 and 3.14, respectively. In 3B6, the biological reflectors are clearly visible near the centre of the image. The apparent noise in the upper portion of the image is actually microstructure caused by turbulence behind an ice keel. The keel is upstream of the transducer. The other object of interest in this image is the current metre which appears near the lower portion of the image at one location. It is indicated in the lowest detail box of the figure. The current metre was deployed during the collection of the data near the transducer and was detected by one of the side lobes of the transducer. These side lobes are visible in the beam pattern.

As mentioned before, the WASP is a bottom mounted device. KL930, Figure 3.14, was captured in the opposite direction from 3B6. In KL930, multiple reflections occur at the top of the image. They are separated from the rest of the image by a dark horizontal band. The multiple reflections were caused by the high acoustic reflectivity of the surface of the

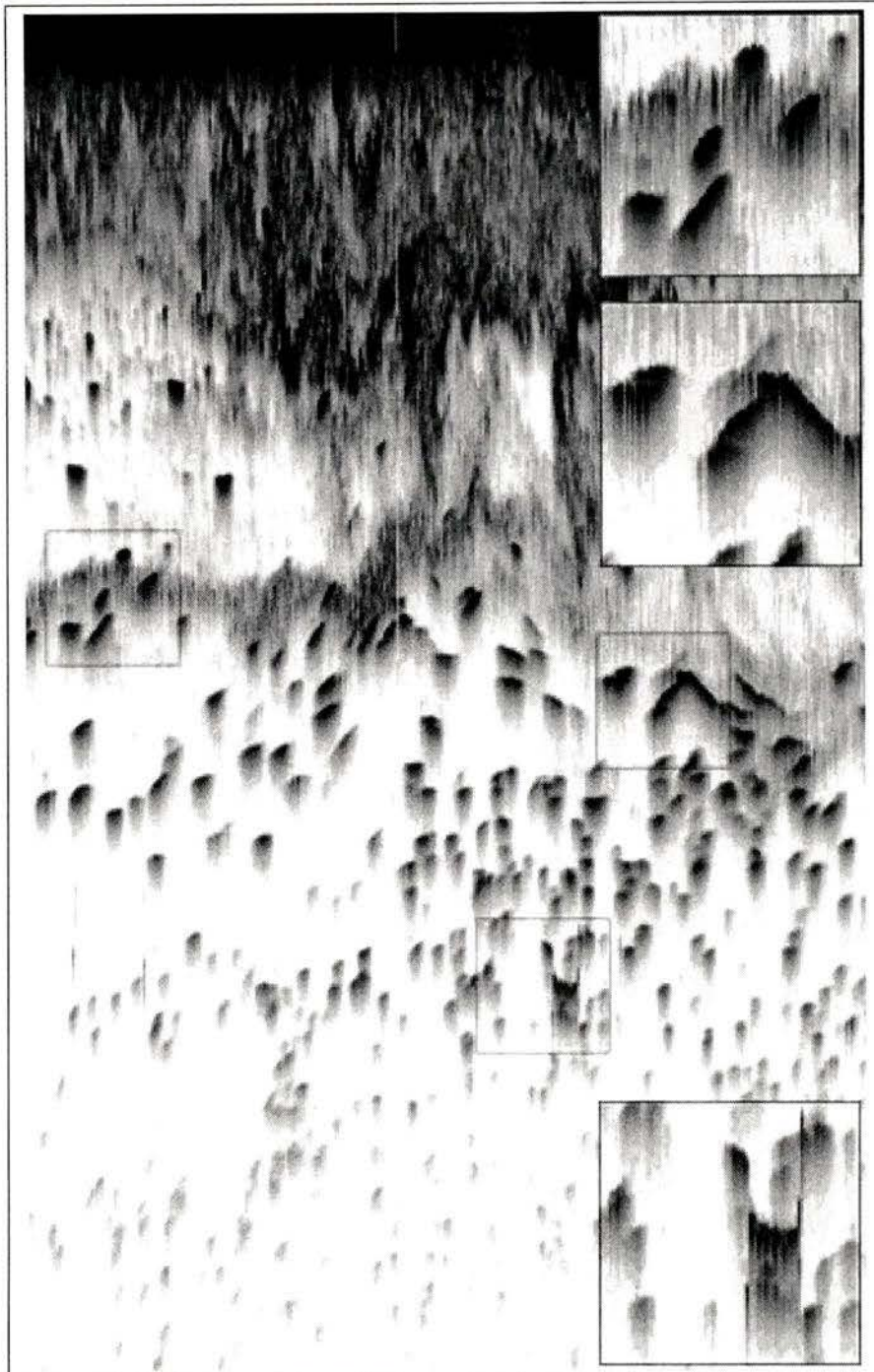


Figure 3.13: The image 3B6.

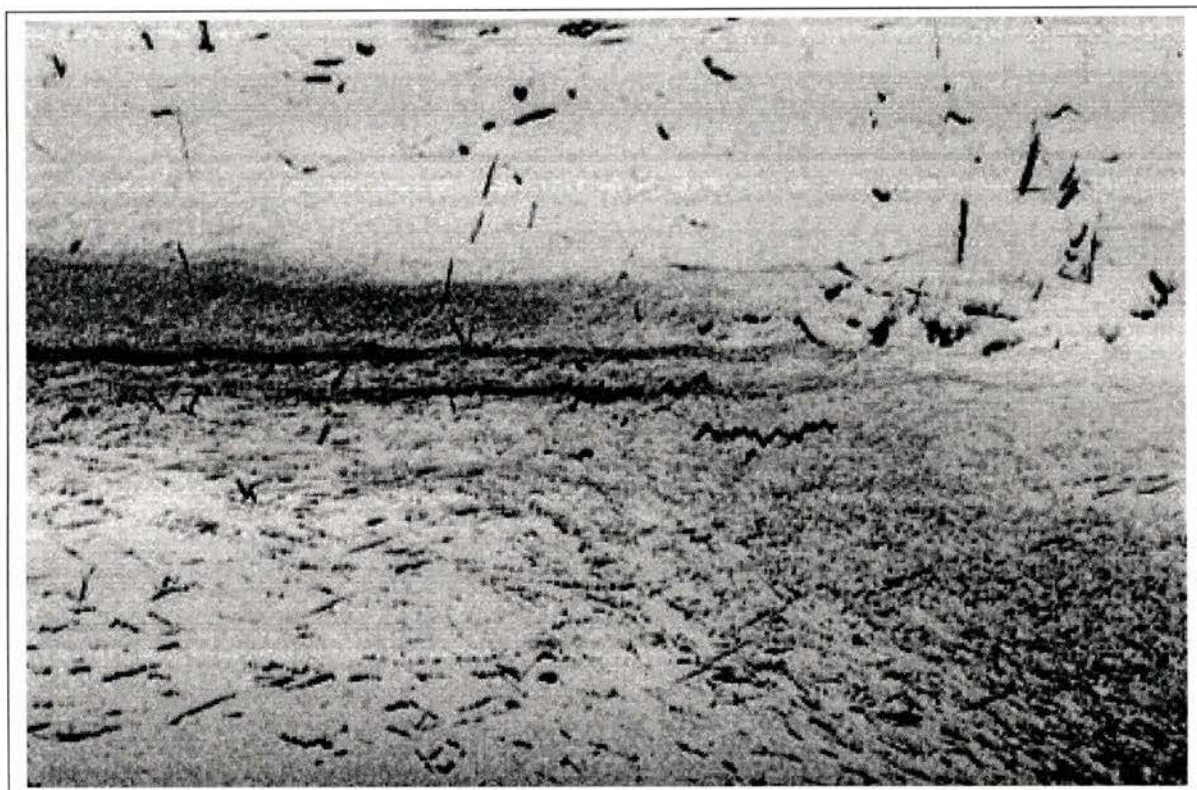


Figure 3.14: The image KL930.

water. The multiples are omitted from most of the analysis here. However, the multiple reflections often contain useful information that can aid in the detection and classification of the reflectors.

3.2.1 Detection

Normalisation

The LTM was useful for the sampled images. The number of reflectors was sufficiently high to provide a smooth LTM curve. This was due mostly to turbulence and numerous small reflectors in the images. The dLTM normalisation of image 3B6 appears in Figure 3.15. The artifacts caused by the coherence of the large reflectors are apparent as faint lines which stretch across the image at constant z .

The edLTM did not produce the same artifacts that were caused by the dLTM. Figure 3.16 shows the 3B6 image with an edLTM normalisation. The absence of artifacts makes this normalisation the most visually appealing, but the edLTM does not remove the effects of the saturation of the transducer at early sample times. If the saturation removal is important, the dLTM would be more appropriate for further processing of this image.

The other sampled image, KL930, was also normalised with the edLTM (Figure 3.17). The normalised version of this images will also be used throughout the rest of this investigation.

In some images, the LTM normalisation is inappropriate. Typically, these images will have had a hardware TVG applied during sampling. Another factor which might influence this would be a sparsity of microstructure and small background reflectors. This would cause the calculated dLTM to have zero values. The division by zeros would cause the normalised image to have locations with undefined amplitudes. These zero values would also biased the edLTM dramatically such that a fit to an exponentially decaying curve was inappropriate, as well.

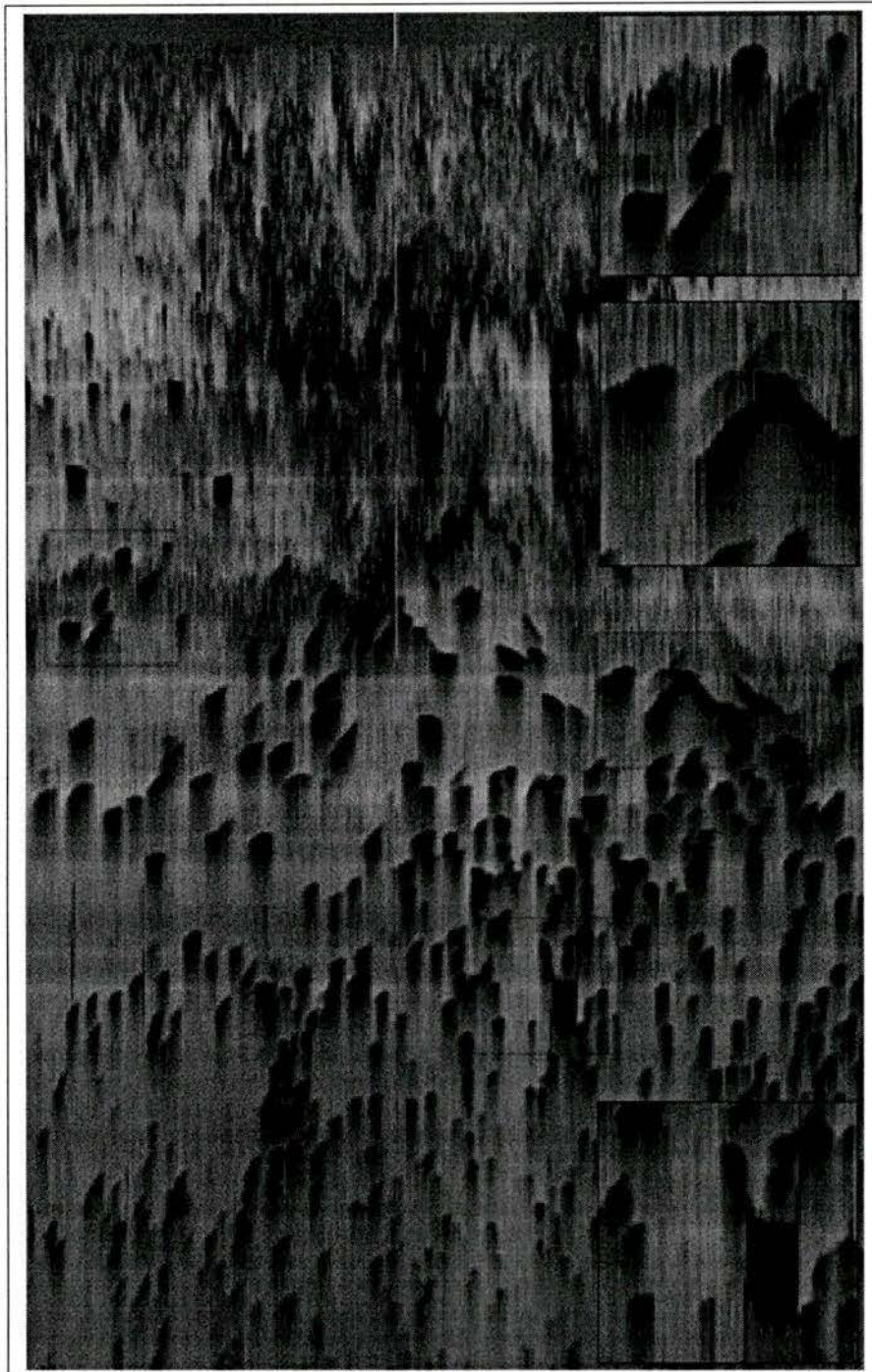


Figure 3.15: The image 3B6, normalised with the calculated LTM.

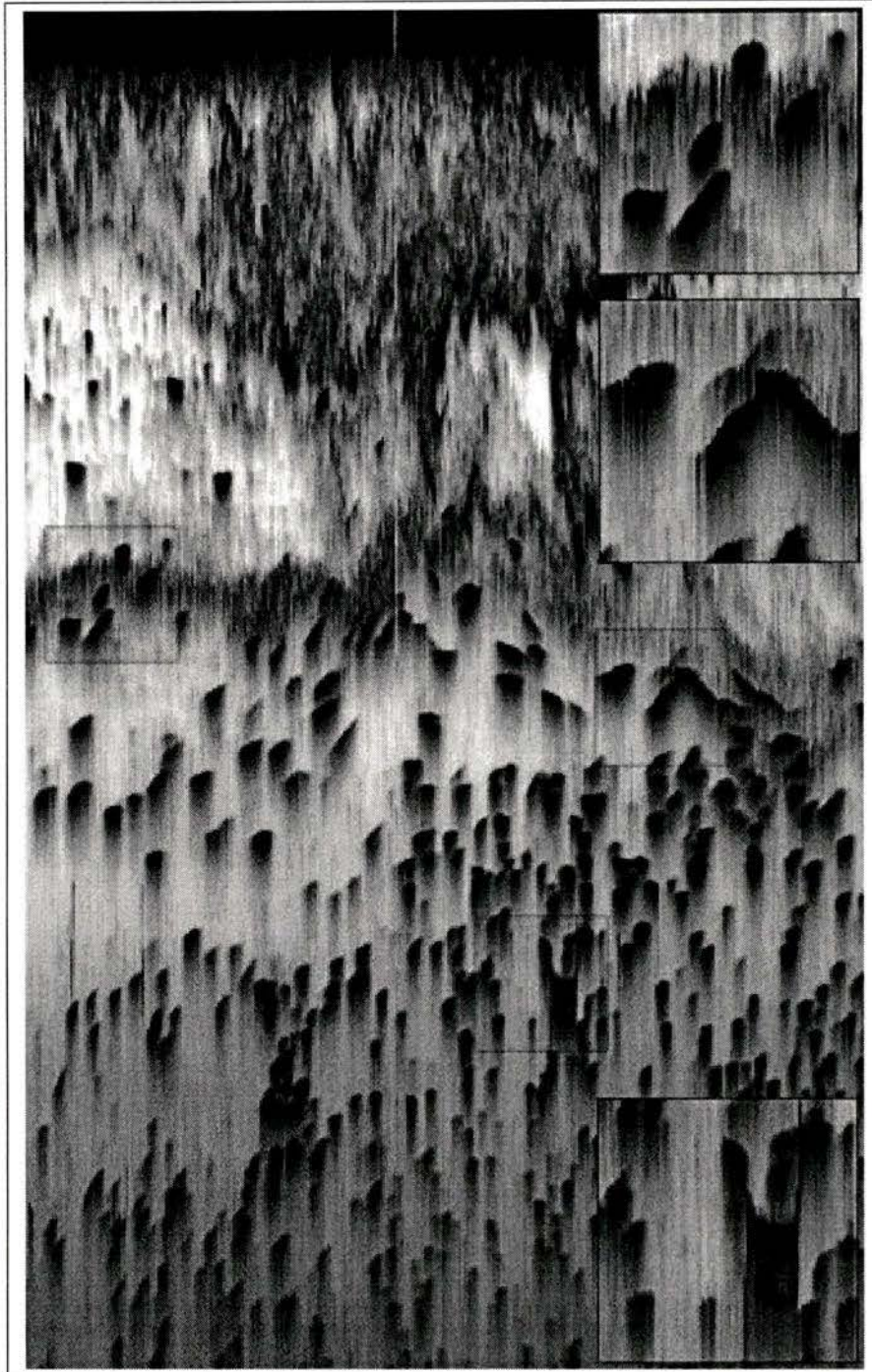


Figure 3.16: The image 3B6, normalised with the exponential fitted LTM.



Figure 3.17: The image KL930, normalised with the edLTM.

Time domain

The examples of the time domain detection on both images are presented in Figure 3.18 and Figure 3.19. The reflectors are readily identified in both images. As with the synthetic images, the time domain peak detection algorithm seems to have difficulty distinguishing between reflectors and noise.

It should be noted that the KL930 image is sampled with the transducer at the bottom of the water-column. The dPeak algorithm omits the first 40 samples from the image to improve the detection and classification in the absence of the saturation of the acoustic system. This produces the white band across the bottom of Figure 3.19.

Fourier domain

The Deconv operation was applied only to 3B6 (see the data flow diagramme in Figure 3.1).

As expected, the response functions from the sampled image were more difficult to extract than with the synthetic images. A typical response function, shown in Figure 3.20, was extracted from 3B6 using the same method as the synthetic images. It was extracted by taking x_{ij} for $i = 152..178$ and $j = 183$ which corresponded to the position of the objects. This is the same procedure that was described in section 2.1.3. In this case, only a single value for j was used. In other words, a single reflector object was extracted from the image and used for the deconvolution. The response was captured at the point where the reflector was closest to the transducer. In the case of a reflector with no vertical velocity, this corresponds with the nearest approach by the reflector to the center of the beam.

The results of the deconvolution of 3B6 with the sample response function are shown in Figure 3.21. . The application of the Thresh algorithm with a threshold of 5 % of the largest amplitude value in the image gives Figure 3.22. While it performs quite well, the problem with this detection method is that the threshold level must be tailored to suit the image. Each image is different. This prevents an unsupervised analysis.

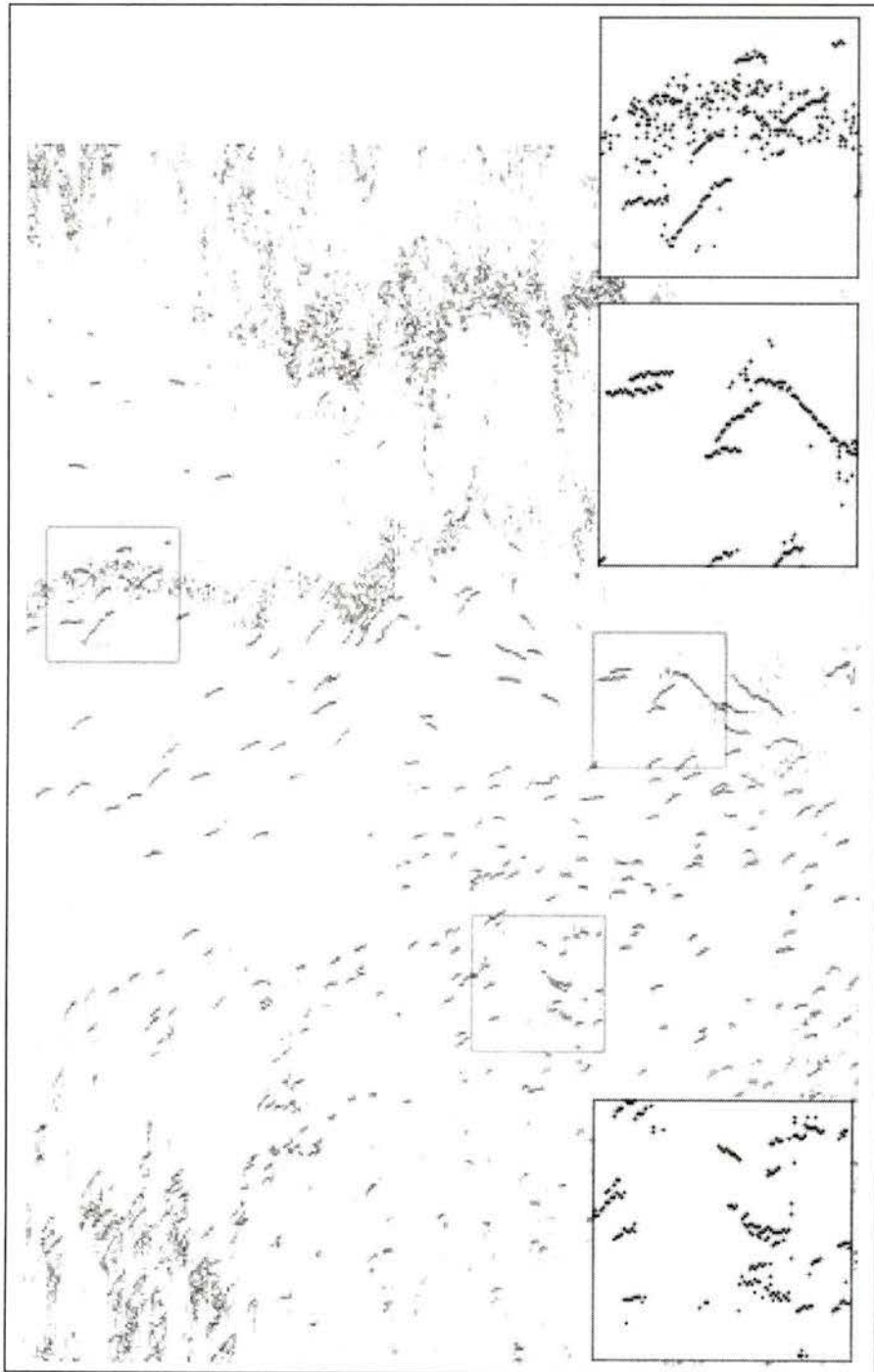


Figure 3.18: The dPeak detection of image 3B6 showing the peak detection in the time domain.

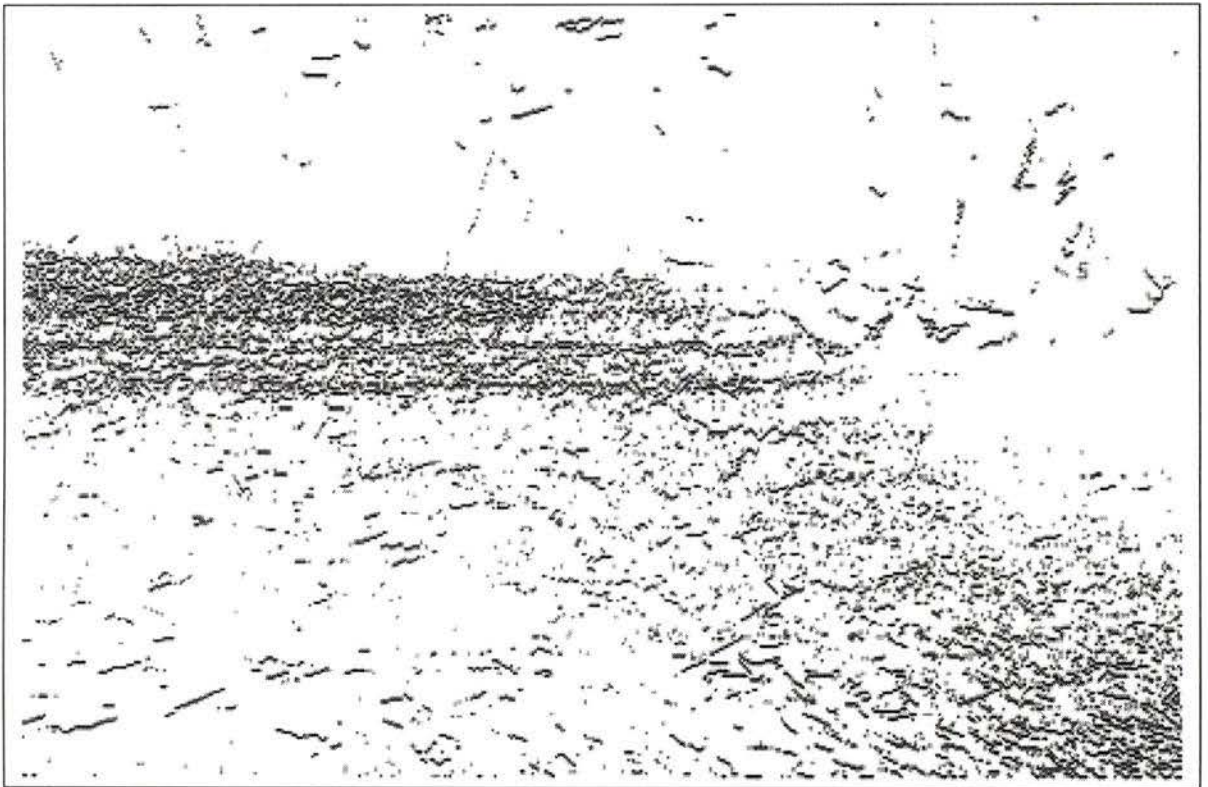


Figure 3.19: The dPeak detection of image KL930 showing the peak detection in the time domain.

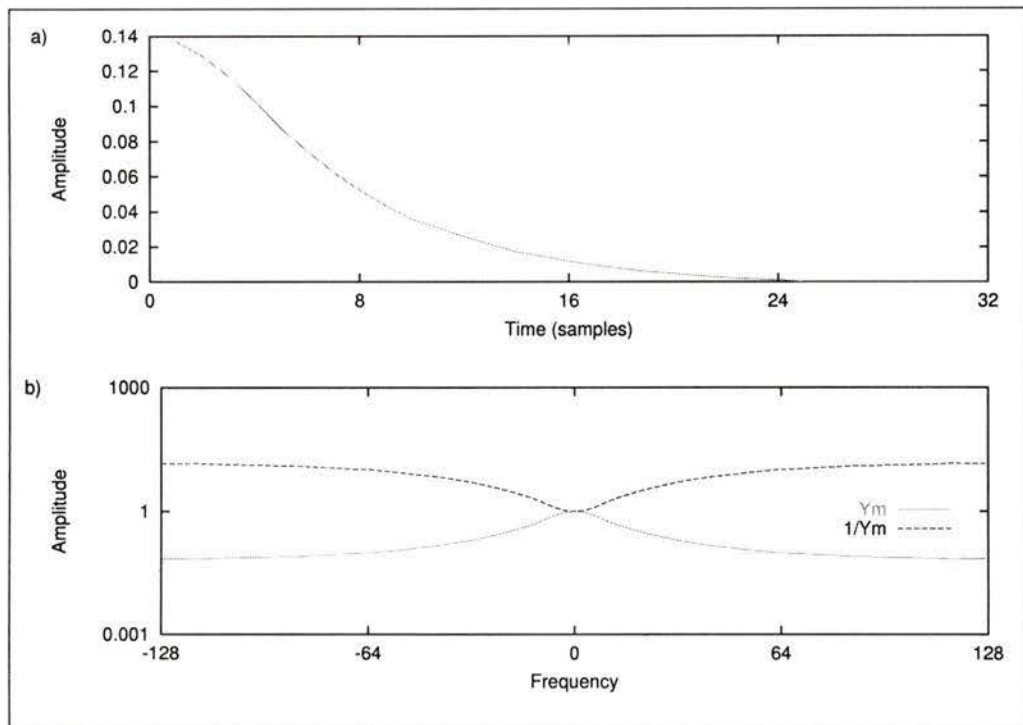


Figure 3.20: a) The response function extracted from image 3B6. b) The magnitude and its reciprocal of the sample response function in the Fourier domain.

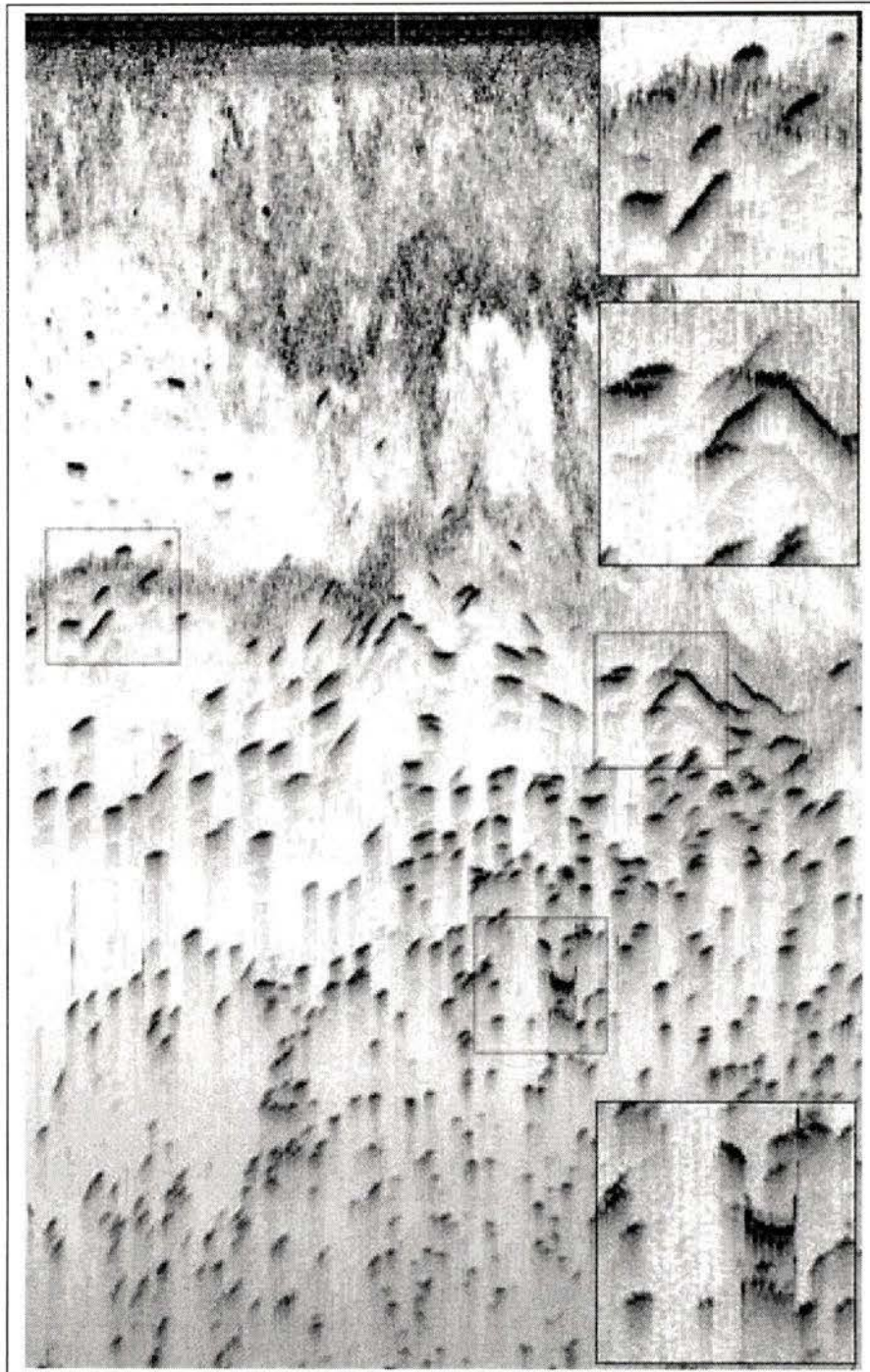


Figure 3.21: The result of the deconvolution applied to 3B6.

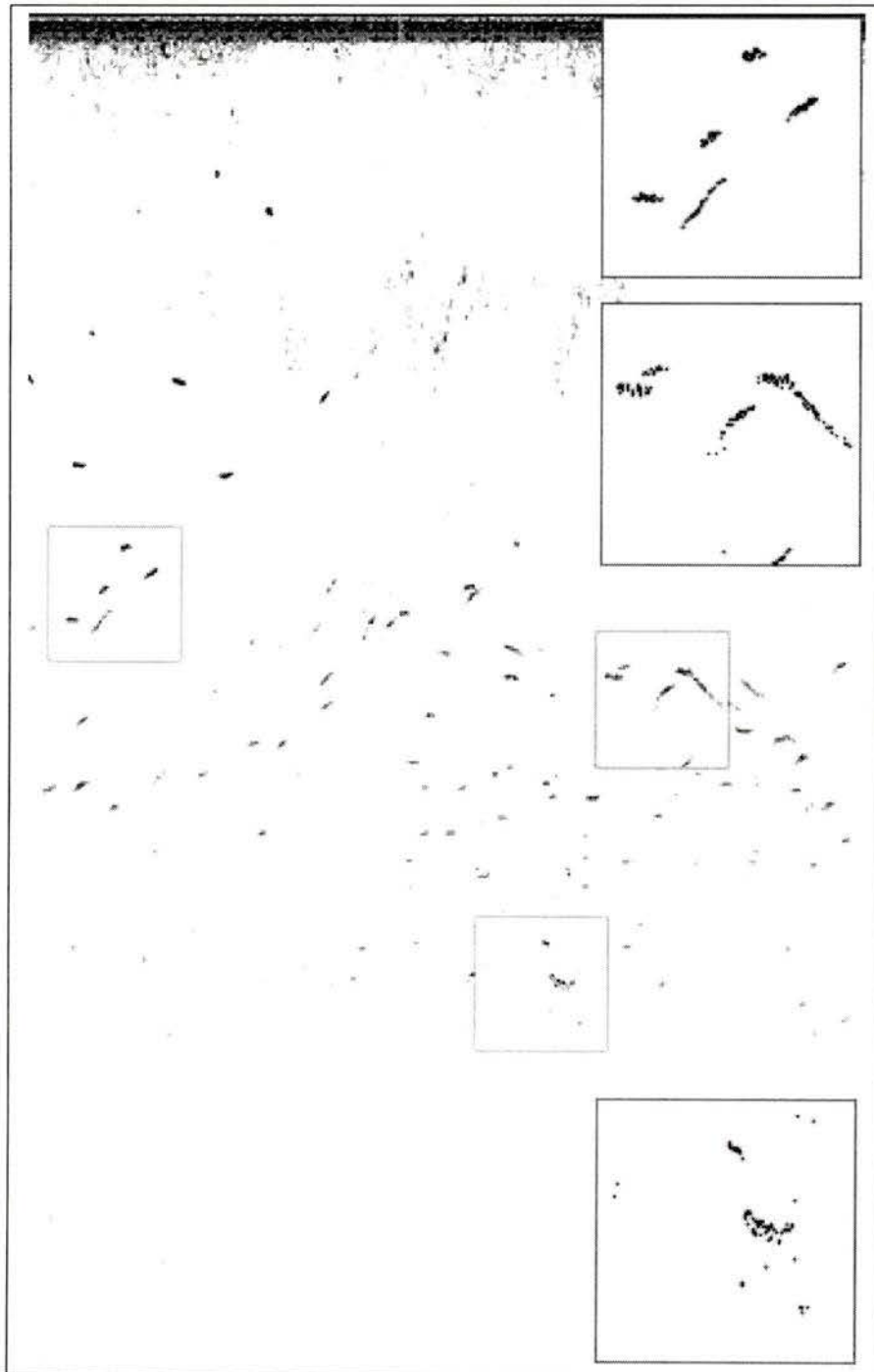


Figure 3.22: The result of Deconv-Thresh applied to 3B6.

Wavelet domain

The wavelet domain detection offered an improvement over the time and Fourier domain detection in the previous two sections. The wavelet transforms of the two sampled images are given in Figures 3.23 and 3.24. With this detection scheme, a single hierarchy from the wavelet domain representation of each image was searched for reflector objects. Due to the granularity of the wavelet domain, caused by downsampling, the reflectors that were found were used to aid the dPeak method in finding the objects. The wavelet domain put the dPeak algorithm into the ball-park. This is described in section 2.1.4.

The detection, performed with the Wavelet-dPeak hybrid, as in section 3.1, is shown in Figure 3.25 and 3.26 for images 3B6 and KL930, respectively.

3.2.2 Classification

Transient information

The transient classification of the sampled images consisted of the wavelet hierarchy selection (hSelect).

In the second method, hSelect, the classification seemed to fare better than with the synthetic images. The wavelet domain representations of the sampled images are available in Figures 3.23 and 3.24. A close examination reveals that the reflector objects are more prominent in certain hierarchies. This bias toward certain hierarchies is exploited with a selection of hierarchies 8–9, 7–8, and 6–7, respectively, for each of the images. These images are shown in Figures 3.27 and 3.28.

The hSelect version of 3B6 shows some definite improvements over the original. The reflector objects of interest are clearly more prominent. As described in section 2.2.1, the wavelet classification is not as black-and-white as the time domain classification. The granularity is visible in the images. A “ringing” effect caused by the incomplete modelling of the reflectors is also apparent. The best approach is to model the reflector objects as

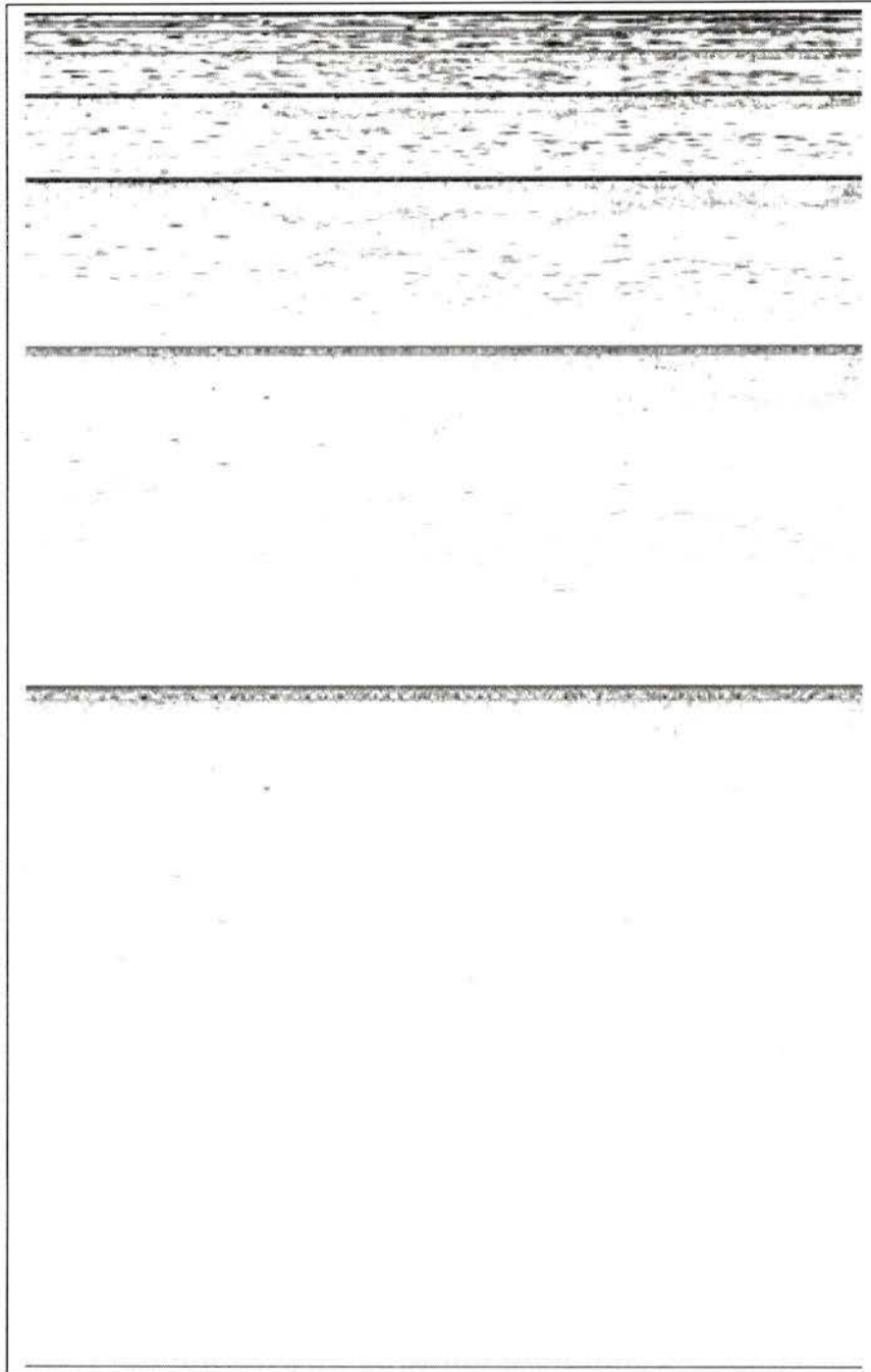


Figure 3.23: The wavelet transform of image 3B6.

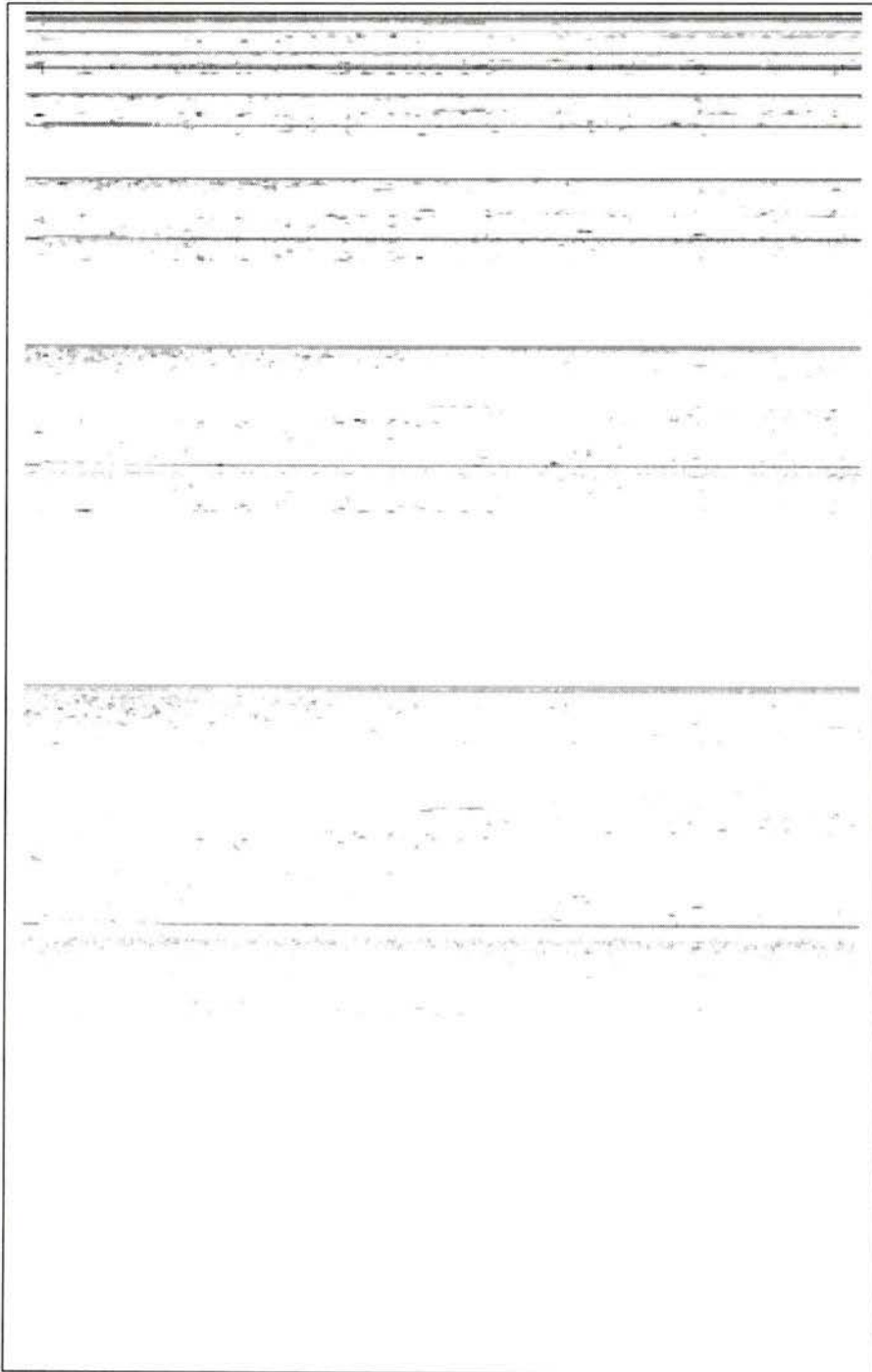


Figure 3.24: The wavelet transform of image KL930.

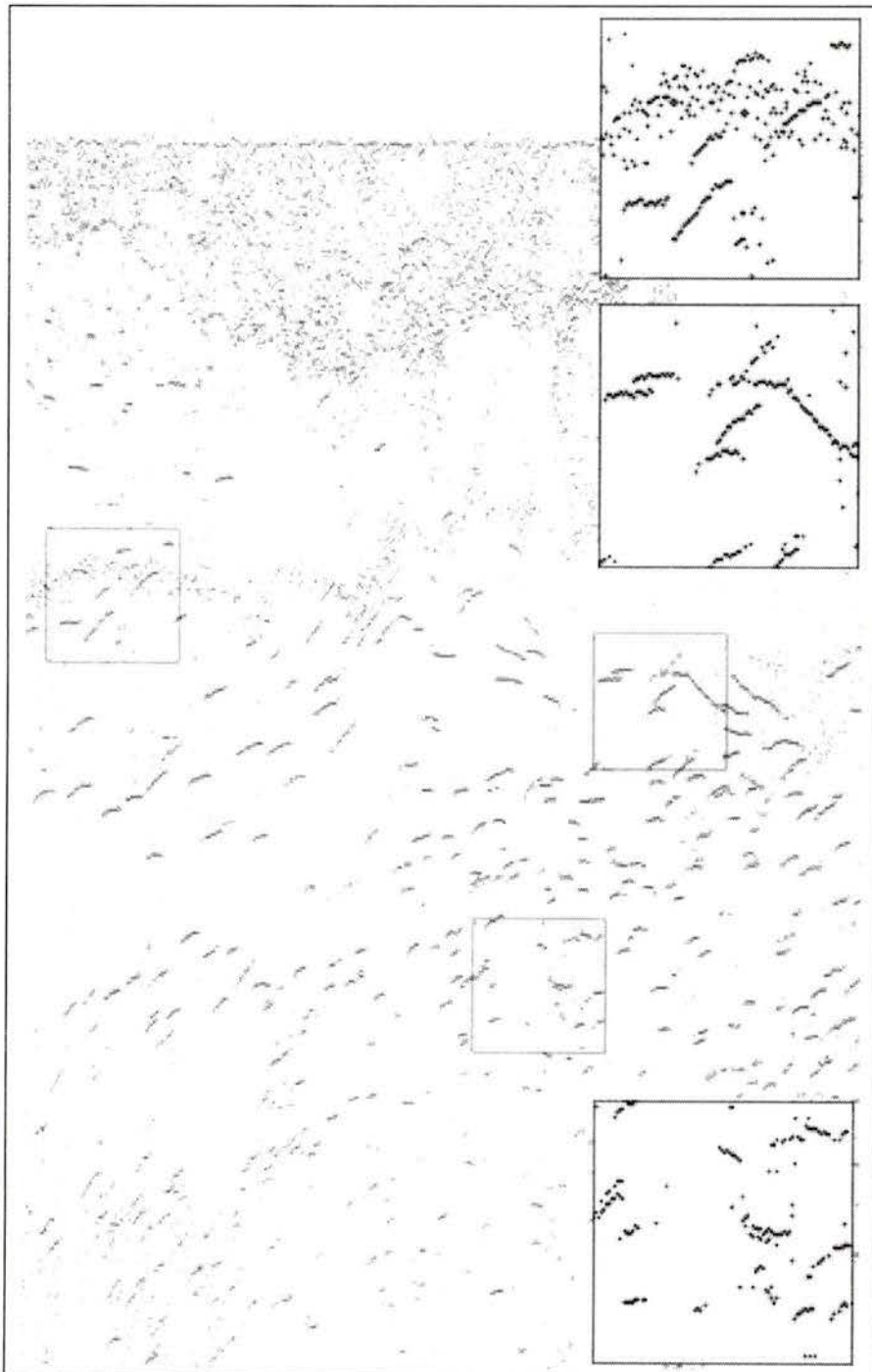


Figure 3.25: The Wavelet-dPeak detection of image 3B6 showing only detected peaks.

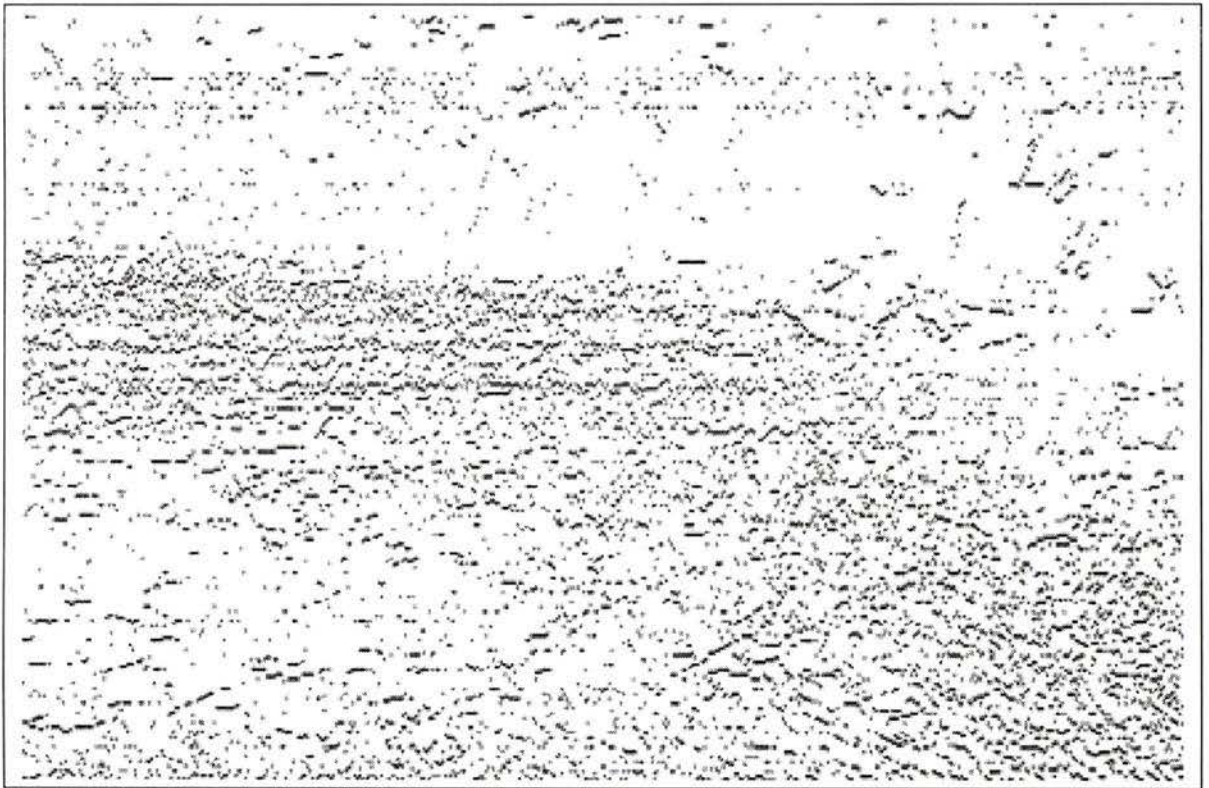


Figure 3.26: The Wavelet-dPeak detection of image KL930 showing only detected peaks.

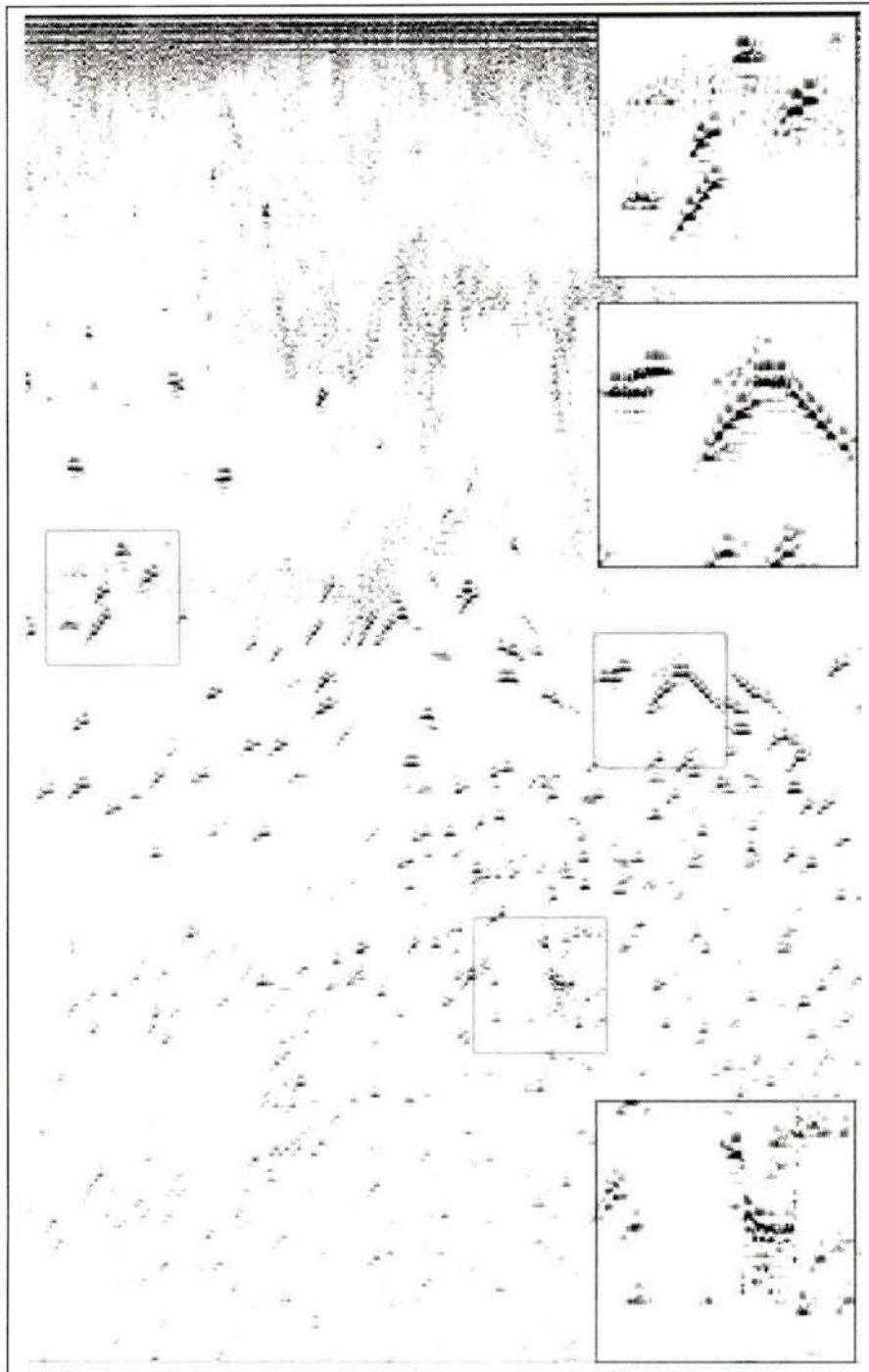


Figure 3.27: The hSelect classification of image 3B6 using hierarchies 8–9.



Figure 3.28: The hSelect classification of image KL930 using hierarchies 7-8.

efficiently as possible. This means that some of the energy from the reflectors may not be removed. Further some of the energy from other sources such as noise and undesired objects may be removed during the classification. This is apparent in the hSelect classification in this section. The chosen wavelets did not model the desired reflector objects efficiently. This also occurred with the synthetic images in section 3.1.2.

Persistent information

In the Streaks classification, objects which did not form particle tracks were assumed to represent objects which were undesirable. This part of the classification could almost be considered to be part of the detection; as if the detection from the previous stage was extended to further remove the undesirable objects. In fact, the undesirable objects may have come from legitimate reflectors such as turbulent microstructure in the water column. These reflectors would not likely form coherent particle tracks. The Streaks classification was found to be effective at omitting these objects.

The Figures 3.29, 3.30, and 3.31 represent the result of the different classifications of image 3B6. The dPeak-Streaks, the Deconv-dPeak-Streaks, and the Wavelet-dPeak-Streaks classifications are presented, respectively, in each of the three figures. As with the synthetic image classification, there are no reflector objects in the images that were not constituents of at least one particle track.

Examination of Figure 3.29 and 3.31 reveals some significant differences. The dPeak algorithm seems to detect more undesirable objects as it did in with the synthetic images. The differences in the images can be seen in the turbulent microstructure at the top of both images. Noticeable differences are also apparent near some of the deeper reflectors in the bottom left corners of the images.

The other image, Figure 3.32, shows the Wavelet-dPeak-Streaks particle tracking result.

As with the synthetic images, the choice of detection makes a difference to the classification of the sampled data. The Wavelet-dPeak-Streaks classification appears to be the

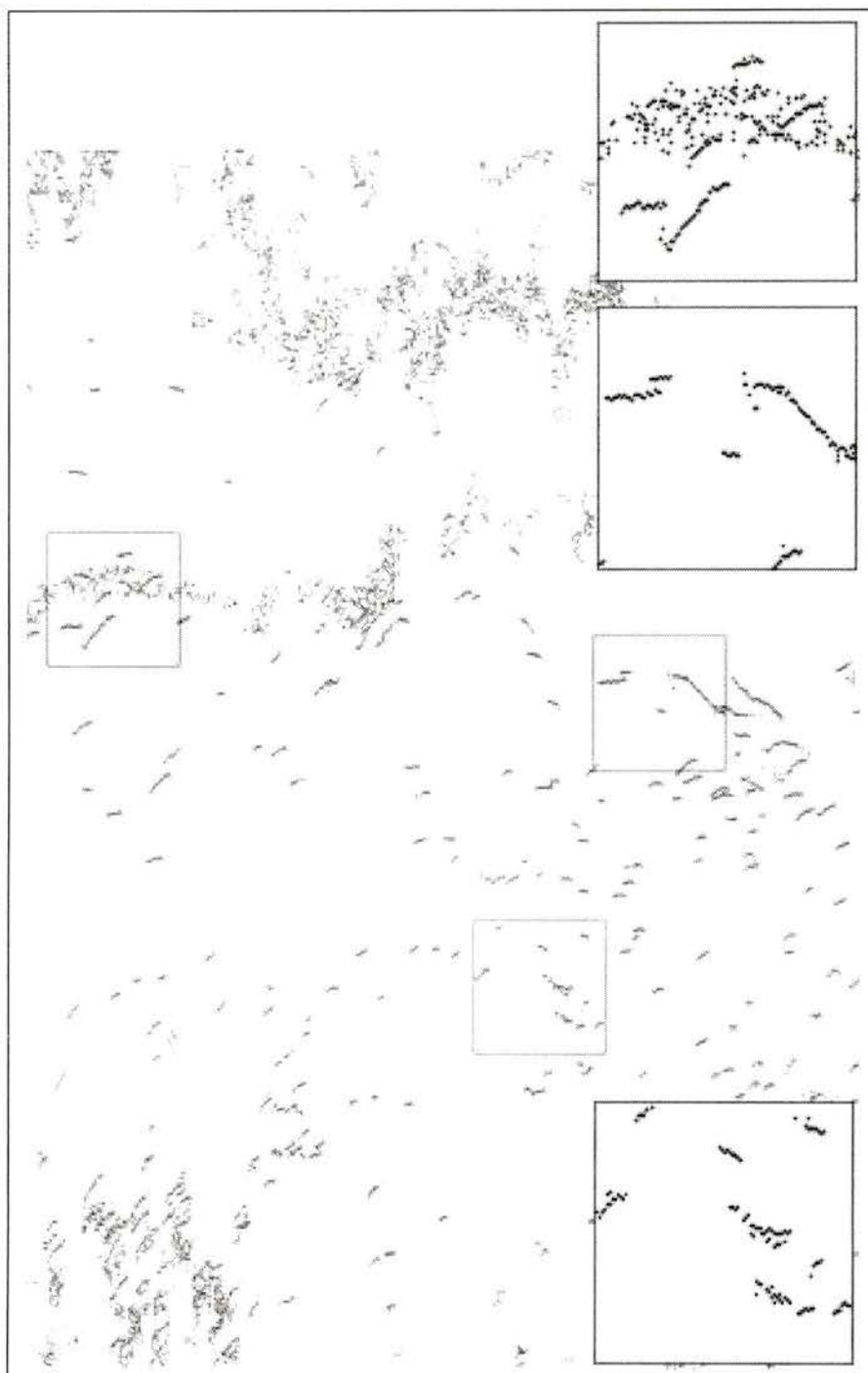


Figure 3.29: The dPeak-Streaks classification of image 3B6. All objects in the image belong to at least one particle track.

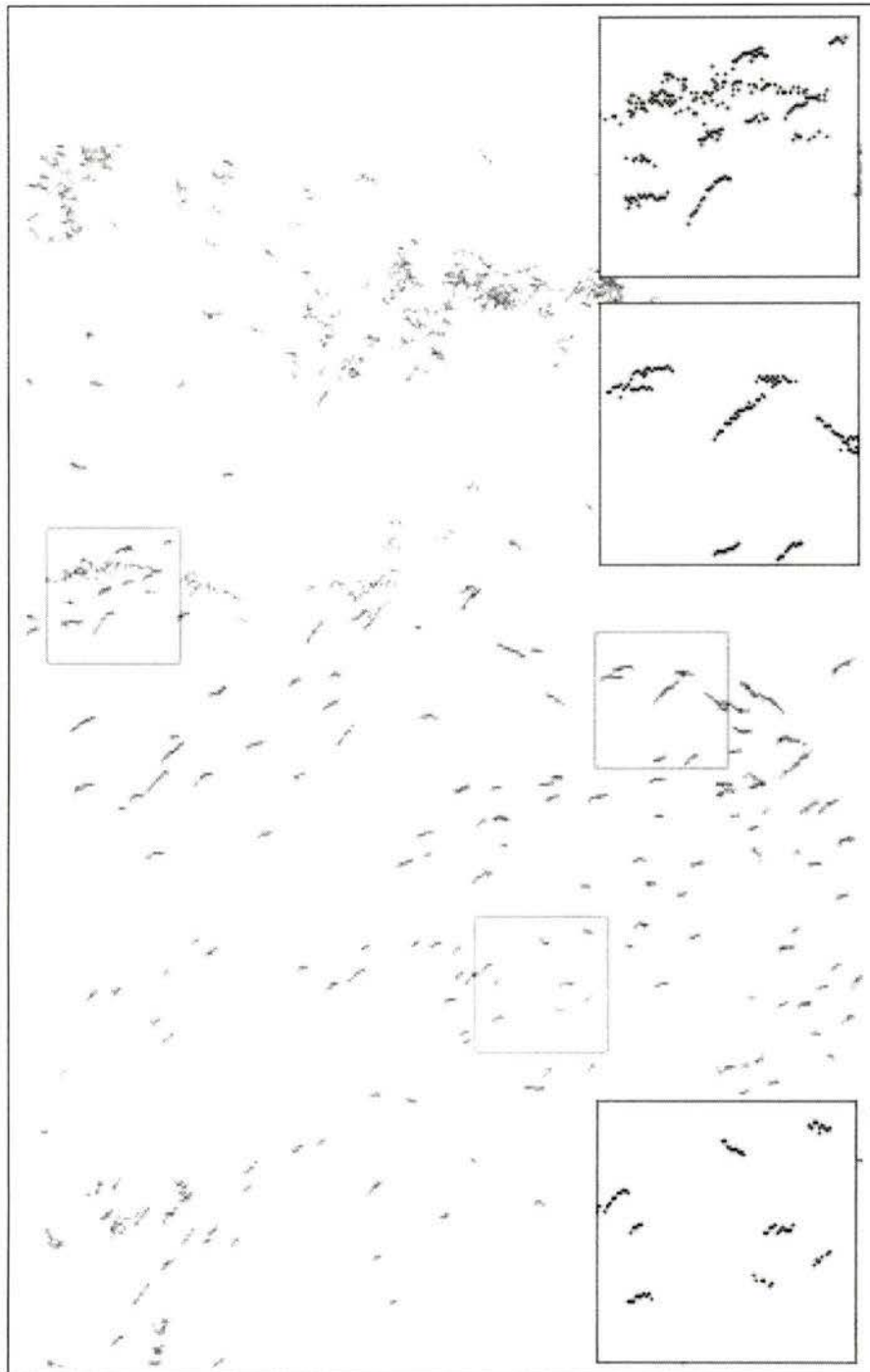


Figure 3.30: The Deconv-dPeak-Streaks classification of image 3B6.

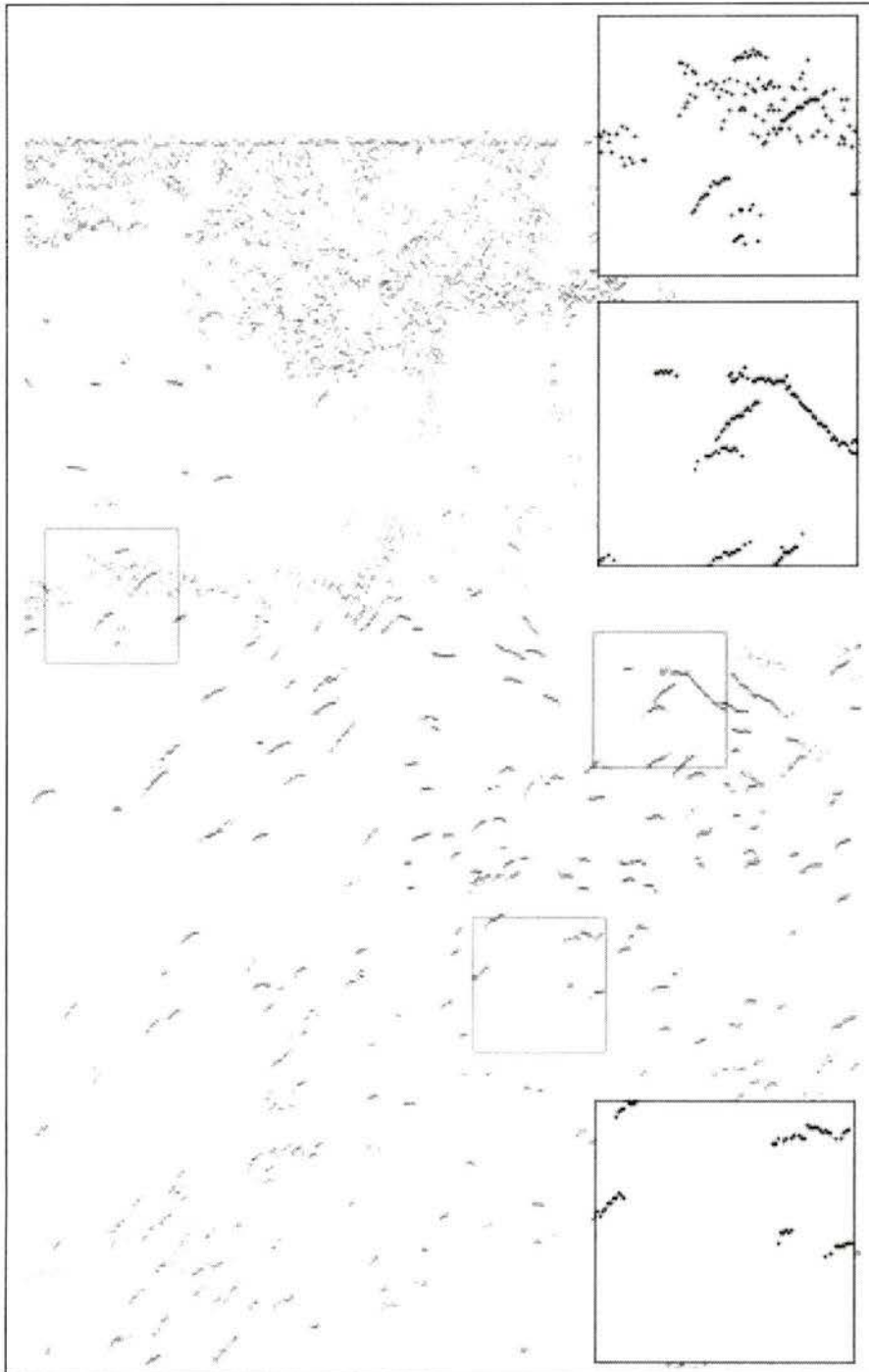


Figure 3.31: The Wavelet-dPeak-Streaks classification of image 3B6.

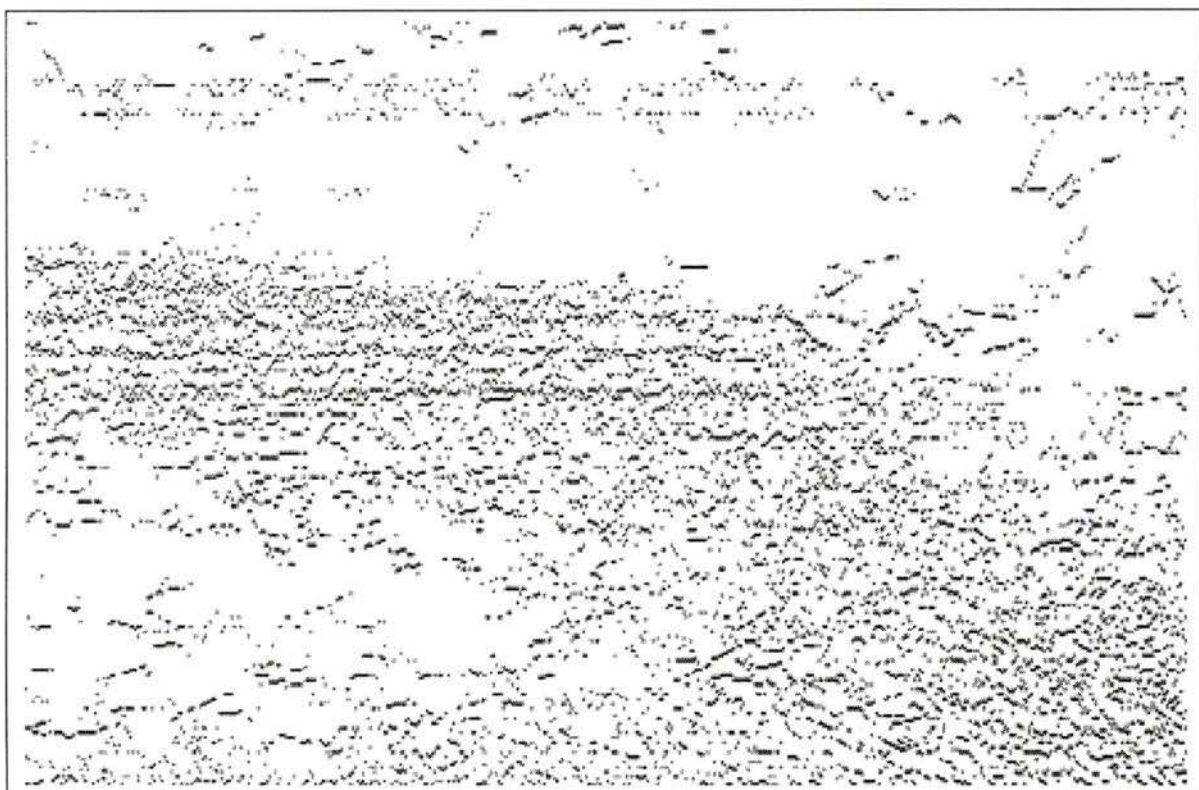


Figure 3.32: The Wavelet-dPeak-Streaks classification of image KL930.

optimal choice. The dPeak-Streaks classification performs well in the absence of noise, but is confused by the peaks which occur in the tails of the spikes. This effect is visible in the lower left corner of Figure 3.29. The Deconv-dPeak-Streaks algorithm performs on the sampled data about as well as it did on the synthetic data. Figure 3.30 shows the influence of the edge effects of the deconvolution.

Chapter 4

Discussion and Conclusion

Four questions were asked at the beginning of this investigation:

1. Can the biological reflectors be distinguished from the other signals in general acoustic images?
2. Can the biological reflectors be removed effectively from the images to allow further processing to access the fine detail.
3. Can the biological reflectors be further split into sub-classes?
4. Can the process be automated to act in an unsupervised approach?

The answer to the first question is, “It depends.” The collection of the acoustic data strongly influenced the results of the classification of the reflectors. For example, in the Ice Keel 85 data used in this investigation, the impulse response function of the equipment appeared to dominate the response functions of the reflectors in the water column. In this case, it was difficult to distinguish the biological reflectors from the other objects in the acoustic data using the signature of the reflectors. The reflectors all “looked” similar. A different phenomenon with a similar effect occurred in the WASP data. The image vectors

were derived through the integration of a number of sampled vectors in the hardware. This caused the responses of the reflectors to be binned, modifying the characteristic signatures.

The classification of the reflectors relied on the differentiation of the reflectors. The difficulty with the detection inhibited the classification. The second question can not be answered satisfactorily until the problems influencing the first are resolved.

In the process of determining the answers for the first and second questions, a number of algorithms were developed. The algorithms were presented with the aim of detecting and classifying biological reflectors in acoustic images. They provided a basis for manipulating acoustic data in three domains: the time domain, the frequency domain, and the wavelet domain. It was felt this multi-faceted approach would help surmount the difficulties that would be encountered in any one alone.

Four acoustic images were analysed. Two were synthetic and two were sampled from acoustic surveys. Each image brought some unique aspect to the analysis. The image analysis was split into two components: The first was the detection of objects in the acoustic image. This basic operation was implemented in all three domains in such a way to allow the results of the algorithms to be compared. The second component involved the classification of the objects.

The chief benefit to using the time domain for detection was simplicity. The disadvantage was the inability to separate the desired objects in the image from undesirables and noise. With that exception noted, the time domain detection performed well. It provided the basis by which the other techniques were measured.

The spatial frequency domain had the potential to provide powerful tools for separating the different classes of reflectors. Edge effects and the inability of deconvolution to handle dilation of the objects made it awkward to implement this operation effectively. Another potentially valuable approach was developed by Friedlander *et al* [8] but was not investigated in this thesis.

The wavelet domain fared better than the time domain in the face of a noisy signal. It also had the ability to separate the objects in question from objects with different characteristic time scales. The noise components of the image were not as distinct from the desirable objects as was expected. While this did cause difficulty, it is likely that a better set of wavelets would improve this operation. Perhaps a modification of the approach by Daubechies *et al* [5, 6] might yield wavelets which model exponentially decaying spikes efficiently in the wavelet domain. It is likely that these new wavelets would not describe an invertible wavelet transform, however.

More appropriate wavelets would solve two problems. First, the modelling and the detection of the objects would be improved, thereby increasing the utility of the wavelet domain for object classification. Second, the extraction of the objects could then be performed effectively in the wavelet domain.

These factors limited this investigation from realising the full potential of the wavelet domain techniques. Even in its inhibited form, the wavelet transform did serve to improve the time domain detection of the objects. Here, the issue of granularity arose. The lower hierarchies of the wavelet transform are essentially downsampled versions of the original image. Any reference from a wavelet hierarchy back to its estimated location in the time domain suffers from this problem.

The second component of the analysis was the classification of the reflector objects that were detected in the first component. Two approaches were taken with the classification. The first approach relied on the transient data present in each echo-return, each vector of the acoustic image.

The wavelet transform was used as a “filter” to separate objects with differing signatures. It seemed to be more effective on KL930 than with the 3B6 image. The reason for this was likely also related to the dominant equipment response function in the 3B6 image. Since each WASP image vector was integrated from a number of sampled vectors, it is likely that the response functions of the individual reflectors were also obscured, possibly to a lesser

extent than with 3B6.

The shape of each reflector object might also be determined by fitting a decay curve to the object, starting at the location determined in the detection stage. The coefficients of this fit should be examined for a natural grouping using a clustering algorithm. The most likely candidate is the GMVE clustering algorithm outlined in section 2.2.2. The most important factor that would likely affect a clustering routine implemented in this was the type of curve fitted to the decay of the objects. In the initial stages of this investigation, an exponentially decaying curve was found to be inappropriate. It caused the potential clustering of the object properties to be “smeared out.” A more appropriate fitted curve might have been a sum of exponential decays, as in the Particle Response Model. Another possibility is that the transient information was affected by the dominant response function of the sampling equipment. The clustering of the synthetic images did show more structure than the sampled images, but this could not be established adequately for this thesis.

The second approach to the classification made no assumptions about the transient information other than the locations of the objects had been previously detected. This approach used the persistent information, the locations of these objects with respect to time. In order to make statements about the time-related change in position of the reflector objects, the objects from each reflector had to be gathered together into a group. This was called a particle track. The Streaks algorithm was responsible for this operation. In general, the Streaks algorithm was a success. The detected objects were connected to visible particle tracks. A comparison with the original images shows a strong correlation in all cases.

The answer to third question on the list, above, was “No.” In this case, sub-classes of reflectors were not detected. That is not to say that this goal is unreachable using the techniques presented here. In fact, removing the constraints inhibiting these algorithms would likely improve both the detection and classification enough to yield a robust distinction within classes of reflectors.

It is expected that some of the methods presented in this study could be applied to an unsupervised classification of acoustic images. The most likely candidate for the unsupervised detection and classification of reflectors would be the wavelet transform. This multi-resolution approach seemed to have the least dependence on properties of the images. An example of this occurs in section 3.2. In all cases, the last two hierarchies were used to extract the “fish” reflectors from the images. In future applications of the wavelet transform, pending a successful search for more appropriate wavelets, the most appropriate hierarchies to use in the detection and classification could potentially be calculated from the acoustic velocity, sample rate, and estimated reflector sizes.

References

- [1] Addison, D.M., "Vanishing Point and Intersection Detection with Robust Clustering," unpublished, University of Victoria, 1995.
- [2] Addison, D.M., D.R. Topham, and C.J. Konzelman, "Acoustic Image Compression with the Wavelet Transform," *The Canadian Acoustical Association*, vol. 24, no. 3, pp. 49, 1996.
- [3] Chui, C., *An Introduction to Wavelets*, Academic Press, Toronto, 1992.
- [4] Daubechies, I., "Orthonormal Bases of Compactly Supported Wavelets," *Communications on Pure and Applied Mathematics*, vol. 41, pp. 909-996, 1988.
- [5] Daubechies, I., S. Jaffard, and J. Journé, "A Simple Wilson Orthonormal Basis with Exponential Decay," *S.I.A.M. Journal on Mathematical Analysis*, vol. 22, no. 2, pp. 554-572, 1991.
- [6] Daubechies, I., "Orthonormal Bases of Compactly Supported Wavelets II: Variations on a Theme," *S.I.A.M. Journal on Mathematical Analysis*, vol. 24, no. 2, pp. 499-519, 1993.
- [7] Ehrenberg, J., and T.C. Torkelson, "Application of Dual-Beam and Split-Beam Target Tracking in Fisheries Acoustics," *ICES Journal of Marine Science*, vol. 53, pp. 329-334, 1996.
- [8] Friedlander, B., and B. Porat, "Detection of Transient Signals by the Gabor Representation," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 2, pp. 169-179, 1989.
- [9] Frisch, M., and H. Messer, "The Use of the Wavelet Transform in the Detection of an Unknown Transient Signal," *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 892-897, 1992.
- [10] Jolion, J.M., P. Meer, and S. Bataouche, "Robust Clustering with Applications in Computer Vision," *IEEE Transactions on Pattern Analysis and Machine*, vol. 13, no. 8, p. 791, 1991.

- [11] Konzelman, C., "Introduction to Acoustics," *Mechanical Engineering 503 course notes*, University of Victoria, 1991.
- [12] Korn, G., and T. Korn, *Mathematical Handbook For Scientists and Engineers*, 2nd ed., McGraw-Hill, Toronto, 1968.
- [13] Oppenheim, A., and R. Schaffer, *Digital Signal Processing*, Prentice-Hall, Toronto, 1975.
- [14] Press, W.H., S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed., Cambridge University Press, Cambridge, 1992.
- [15] Strang, G., "Wavelets and Dilation Equations," *Siam Review*, vol. 31, pp 613-627, 1989.
- [16] Strang, G., and T. Nguyen, *Wavelets and Filter Banks*, Wellesley-Cambridge Press, Wellesley, 1996.
- [17] Topham, D.R., "Ice Keel 85," Institute of Ocean Sciences, Department of Fisheries and Oceans, Sidney, 1991.
- [18] Topham, D.R., P. Johnston, "Water Acoustic Structure Profiler (WASP) Deployment on Kootenay Lake," Institute of Ocean Sciences, Department of Fisheries and Oceans, Sidney, 1996.
- [19] Watkins, D., *Fundamentals of Matrix Computations*, John Wiley & Sons, Toronto, 1991.

Suggested Reading

1. Addison, D.M., "Vanishing Point and Intersection Detection with Robust Clustering," unpublished, University of Victoria, 1995.
2. Addison, D.M., D.R. Topham, and C.J. Konzelman, "Acoustic Image Compression with the Wavelet Transform," *The Canadian Acoustical Association*, vol. 24, no. 3, pp. 49, 1996.
3. Broadhead, M., "Broadband Source Signature Extraction from Underwater Acoustics Data with Sparse Environmental Information," *Journal of the Acoustical Society of America*, vol. 97, no. 2, pp. 1322-1325, 1995.
4. Burden, R., and D. Faires, *Numerical Analysis*, 3rd ed., PWS Publishers, Massachusetts, 1985.
5. Burdic, W., *Underwater Acoustic System Analysis*, 2nd ed., Prentice-Hall, Toronto, 1991.
6. Caughey, D., and R.L. Kirilin, "Blind Deconvolution of Echosounder Envelopes," *IEEE Acoustics, Speech, and Signal Processing*, pp. 3149-3152, May 1996.
7. Chen, J-F., J.A. Zagzabski, and E.L. Madsen, "Experimental Demonstration of the Frequency Dependence of the Effective Scatterer Number Density," *Journal of the Acoustical Society of America*, vol. 99, no. 4, pp. 1932-1936, 1996.

8. Churchill, R., and J. Brown, *Complex Variables and Applications*, 4th ed., McGraw-Hill, Toronto, 1984.
9. Chui, C., *An Introduction to Wavelets*, Academic Press, Toronto, 1992.
10. Colton, D., and P. Hahner, "Modified Far Field Operators in Inverse Scattering Theory," *S.I.A.M. Journal on Mathematical Analysis*, vol. 24, no. 2, pp. 365-389, 1993.
11. Daubechies, I., "Orthonormal Bases of Compactly Supported Wavelets," *Communications on Pure and Applied Mathematics*, vol. 41, pp. 909-996, 1988.
12. Daubechies, I., S. Jaffard, and J. Journé, "A Simple Wilson Orthonormal Basis with Exponential Decay," *S.I.A.M. Journal on Mathematical Analysis*, vol. 22, no. 2, pp. 554-572, 1991.
13. Daubechies, I., "Orthonormal Bases of Compactly Supported Wavelets II: Variations on a Theme," *S.I.A.M. Journal on Mathematical Analysis*, vol. 24, no. 2, pp. 499-519, 1993.
14. Denbigh, P., and Q. Smith, "Determination of Fish Number Density by a Statistical Analysis of Backscattered Sound," *Journal of the Acoustical Society of America*, vol. 90, no.1, pp. 457-469, 1991.
15. Ehrenberg, J., and T.C. Torkelson, "Application of Dual-Beam and Split-Beam Target Tracking in Fisheries Acoustics," *ICES Journal of Marine Science*, vol. 53, pp. 329-334, 1996.
16. Everitt, B.S., *Cluster Analysis*, Great Britain: Edward Arnold, 1993.
17. Faran, J., "Sound Scattering by Solid Cylinders and Spheres," *Journal of the Acoustical Society of America*, vol. 23, no. 4, pp. 405-417, 1951.

18. Feuillade, C., R.W. Nero, and R.H. Love, "A Low-Frequency Acoustic Scattering Model for Small Schools of Fish," *Journal of the Acoustical Society of America*, vol. 99, no. 1, pp. 196-208, 1996.
19. Friedlander, B., and B. Porat, "Detection of Transient Signals by the Gabor Representation," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 2, pp. 169-179, 1989.
20. Frisch, M., and H. Messer, "The Use of the Wavelet Transform in the Detection of an Unknown Transient Signal," *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 892-897, 1992.
21. Hall, M.V., "Dimensions and Units of Underwater Acoustic Parameters," *Journal of the Acoustical Society of America*, vol. 96, no. 6, pp. 3887-3889, 1995.
22. Jolion, J.M., P. Meer, and S. Bataouche, "Robust Clustering with Applications in Computer Vision," *IEEE Transactions on Pattern Analysis and Machine*, vol. 13, no. 8, p. 791, 1991.
23. Kelly, J.G., and R.N Carpenter, "Object Classification and Acoustic Imaging with Active Sonar," *Journal of the Acoustical Society of America*, vol. 91, no. 4, pp. 2073-2081, 1992.
24. Kiryati, N., Y. Eldar, and A.M. Bruckstein, "A Probabilistic Hough Transform," *Pattern Recognition*, vol. 24, no. 4, pp. 303-316, 1991.
25. Konzelman, C., "Introduction to Acoustics," *Mechanical Engineering 503 course notes*, University of Victoria, 1991.
26. Korn,G, and T.Korn, *Mathematical Handbook For Scientists and Engineers*, 2nd ed., McGraw-Hill, Toronto, 1968.

27. Love, R.H., "Dorsal-Aspect Target Strength of an Individual Fish," *Journal of the Acoustical Society of America*, vol. 49, no. 3, pp. 816-823, 1971.
28. Love, R.H., "Target Strength of an Individual Fish at any Aspect," *Journal of the Acoustical Society of America*, vol. 62, no. 6, pp. 1397-1403, 1977.
29. Lu, J., D.M. Healy, and J.B. Weaver, "Contrast Enhancement of Medical Images using Multiscale Edge Representation," *Optical Engineering*, vol. 33, no. 7, pp. 2151-2161, 1994.
30. Lukasová, A., "Hierarchical Agglomerative Clustering Procedure" *Pattern Recognition*, vol.?, no.?, pp. ??-??, 1979.
31. Madsen, E.L., M.F. Insana, and J.A. Zagzabski, "Method of Data Reduction for Accurate Determination of Acoustic Backscatter Coefficients," *Journal of the Acoustical Society of America*, vol. 76, no. 3, pp. 913-923, 1984.
32. Oppenheim, A., and R. Schafer, *Digital Signal Processing*, Prentice-Hall, Toronto, 1975.
33. Palmer, D.R., "Rayleigh Scattering from Nonspherical Particles," *Journal of the Acoustical Society of America*, vol. 99, no. 4, pp. 1901-1912, 1996.
34. Pierce, A., *Acoustics: An Introduction to its Physical Principals and Applications*, 2nd ed., Acoustical Society of America, New York, 1989.
35. Press, W.H., S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed., Cambridge University Press, Cambridge, 1992.
36. Proni, J.R., and J.R. Apel, "On the Use of High-Frequency Acoustics for the Study of Internal Waves and Microstructure," *Journal of Geophysical Research*, vol. 80, no. 9, pp. 1147-1151, 1975.

37. Shin, F.B., and D.H. Kil, "Full-Spectrum Signal Processing Using a Classify-Before-Detect Paradigm," *Journal of the Acoustical Society of America*, vol. 99, no. 4, pp. 2188-2197, 1996.
38. Stanton, T.K., R.D.M. Nash, R.L. Eastwood, and R.W. Nero, "A Field Examination of Acoustical Scattering from Marine Organisms at 70 kHz," *IEEE Journal of Oceanic Engineering*, vol. 12, no. 2, pp. 339-348, 1987.
39. Strang, G., "Wavelets and Dilation Equations," *Siam Review*, vol. 31, pp 613-627, 1989.
40. Strang, G., "Wavelets," *American Scientist*, vol. 82, pp 250-255, 1994.
41. Strang, G., and T. Nguyen, *Wavelets and Filter Banks*, Wellesley-Cambridge Press, Wellesley, 1996.
42. Stroustrup, B., *The C++ Programming Language*, 2nd ed., Addison-Wesley Publishing Company, Ontario, 1991.
43. Telfer, B.A., H.H. Szu, G.J. Dobeck, J.P. Garcia, H. Ko, A. Dubey, and N. Wither-
spoon, "Adaptive Wavelet Classification of Acoustic Backscatter and Imagery," *Optical Engineering*, vol. 33, no. 7, pp. 2192-2203, 1994.
44. Tewfik, A.H., D. Sinha, and P. Jorgensen, "On the Optimal Choice of a Wavelet for
Signal Representation," *IEEE Transactions on Information Theory*, vol. 38, no. 2,
pp. 747-765, 1992.
45. Topham, D.R., "Ice Keel 85," Institute of Ocean Sciences, Department of Fisheries
and Oceans, Sidney, 1991.
46. Topham, D.R., P. Johnston, "Water Acoustic Structure Profiler (WASP) Deployment
on Kootenay Lake," Institute of Ocean Sciences, Department of Fisheries and Oceans,
Sidney, 1996.

47. Wang, S., "Finite-Difference Time-Domain Approach to Underwater Acoustic Scattering Problems," *Journal of the Acoustical Society of America*, vol. 99, no. 4, pp. 1924-1931, 1996.
48. Watkins, D., *Fundamentals of Matrix Computations*, John Wiley & Sons, Toronto, 1991.
49. Weiss, L.G., R.K. Young, and L.H. Sibul, "Wideband Processing of Acoustic Signals using Wavelet Transforms. Part I. Theory," *Journal of the Acoustical Society of America*, vol. 96, no. 2, pp. 850-856, 1994.
50. Weiss, L.G., "Wideband Processing of Acoustic Signals using Wavelet Transforms. Part II. Efficient Implementation and Examples," *Journal of the Acoustical Society of America*, vol. 96, no. 2, pp. 857-866, 1994.
51. Weiss, L.G., and T.L. Dixon, "Wavelet-Based Denoising of Underwater Acoustic Signals," *Journal of the Acoustical Society of America*, vol. 101, no. 1, pp. 377-383, 1997.
52. Yang, T.C., and T.W. Yates, "Scattering from an Object in a Stratified Medium. I. Frequency Dispersion and Active Localization," *Journal of the Acoustical Society of America*, vol 96, no. 2, pp. 1003-1019, 1994.
53. Yang, T.C., and T.W. Yates, "Scattering from an Object in a Stratified Medium. II. Extraction of Scattering Signature," *Journal of the Acoustical Society of America*, vol 96, no. 2, pp. 1020-1030, 1994.
54. Ye, Z., and D.M. Farmer, "Acoustic Scattering from Swim-Bladder Fish at Low Frequencies," *Journal of the Acoustical Society of America*, vol. 96, no. 2, pp. 951-956, 1994.

55. Ye, Z., "Acoustic Scattering from Fish Swimbladders," *Journal of the Acoustical Society of America*, vol. 99, no. 2, pp. 785-792, 1996.
56. Yen, N., and L.R. Dragonette, "Application of Time-Frequency Analysis to Acoustic Response From an Insonified Object," *IASTED International Conference on Signal Processing and Digital Filtering*, 1990.
57. Ziomek, L., *Underwater Acoustics: A Linear Systems Theory Approach*, Academic Press, Toronto, 1985.

Vita

Surname: Addison

Given Names: Dean Michael

Place of Birth: Victoria, British Columbia, Canada

Date of Birth: 10 October 1967

Educational Institutions Attended:

University of Victoria

1988 to 1991

Malaspina College

1986 to 1988

Degrees Awarded:

B.Sc.

University of Victoria

1991

Honours and Awards:

Fessenden Prize in Underwater Acoustics

1997


Partial Copyright License

I hereby grant the right to lend my thesis to users of the University of Victoria Library, and to make single copies only for such users or in response to a request from the Library of any other university, or similar institution, on its behalf or for one of its users. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by me or a member of the University designated by me. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Title of Thesis:

A Toolset For Classifying Reflectors In Acoustic Images

Author



Dean M. Addison
30 March 1998