Applications of wavelet transforms to the suppression of coherent

noise from seismic data in the pre-stack domain.

Submitted for the degree of Ph.D.

by

Andrew J. Deighan B.Sc. (University of Glasgow), M.Sc. (University of Durham).

Department of Geology and Applied Geology,

University of Glasgow.

© Andrew J. Deighan. August 1997.

ProQuest Number: 11007656

ProQuest 11007656

Abstract

The wavelet transform, a relatively new mathematical technique, allows the analysis of non-stationary signals by using basis functions which are compact in time and frequency. The variables in the wavelet domain, scale (a frequency range), and translation (a temporal increment) can be associated with time-frequency, and so in the wavelet transform we have the potential to filter seismic signals in a pseudo time-frequency sense.

The one dimensional discrete multiresolution form of the wavelet transform can be effectively used to suppress low frequency coherent noise on seismic shot records. This process, achieved by the muting or weighting of coefficients in the wavelet transform domain, is demonstrated by suppressing low velocity, low frequency ground roll from land-based seismic data, the benefits of which are visible at both the shot and stack stages of the seismic processing stream.

The extension of this technique to the suppression of higher frequency coherent noise is limited by the octave band splitting of frequency space by the transform. The wavelet packet transform, an extension of the wavelet transform, allows a more adaptable tiling of the time frequency domain which in turn allows the suppression of noise containing high frequencies whilst minimising signal distortion. This technique is demonstrated to be effective in suppressing airblast from land based common receiver gathers, whilst minimising the distortion of reflected signals.

These filtering techniques can be extended to two dimensions, filtering data in the two dimensional wavelet and wavelet packet domains. This technique involves muting the transform coefficients in the wavelet/wavelet packet transform space which has four variables: temporal translation, offset translation, frequency scale and wavenumber scale. As for the one-dimensional case the two dimensional wavelet transform suffers from poor resolution due to the octave splitting of f-k space, but when used in combination with a velocity based shift such as normal moveout, can be used to filter data with minimal distortion to the residual signal. Extending the process to using the two-dimensional wavelet packet transform eliminates the shift requirement and leads to more effective filtering in the four variable transform space. The wavelet packet filtering technique is

effective in suppressing low velocity noise from land based seismic records showing visible improvement in both the common shot records and resultant stack.

The non-stationary properties of the wavelet transform allows the filtering across geophone arrays (that is, the common shot record) by the application of the transform in the offset domain. Filtering of the wavelet coefficients, in combination with a linear or hyperbolic shift applied before and removed after filtering, allows discrimination against linear noise on common shot records associated with first breaks and hyperbolic events on common midpoint records such as multiples. The use of a simple muting technique in the wavelet domain effectively suppresses these forms of coherent noise. Where the velocity contrast between signal and noise is high, noise suppression is possible whilst preserving reflector amplitudes. Where the velocity contrast is smaller, weighting of the wavelet coefficients (based on transforms of the input signal after translation) allows noise suppression whilst preserving the amplitude versus offset relationships of the primary signal. This is shown to be effective on synthetic, marine and land based data, with improvements observed on common shot records and resultant stacks.

The results of all these wavelet transform based filtering techniques are sensitive to the choice of wavelet transform kernel wavelet. The suitability of a kernel wavelet for filtering can be related to the frequency spectra of the kernel wavelet. A fast rate of frequency amplitude fall-off at the edge of a given scale of basis wavelet minimises frequency overlap between neighbouring kernel wavelet scales and so minimises contamination by noise associated with aliasing in the filtered signal, a process that is inherent in the transform process. A flat amplitude response across the frequency range of a given scale also leads to improved filtering results.

Declaration

This thesis is the culmination of research carried out between October 1994 and July 1997 in the Department of Geology and Applied Geology, University of Glasgow, under the supervision of Dr. Doyle R. Watts.

This thesis is the result of my own independant research and any published or unpublished material used by me has been given full acknowledgment in the text.

Andrew J. Deighan,

August 1997.

Acknowledgements

Notation

Let Z and R denote the set of integer and real numbers respectively.

The bar $\bar{x}$ represents the complex conjugate such that $\overline{(a+ib)} = (a-ib)$, where the symbol i represents the square root of -1.

The scalar product $\langle a,b \rangle$ of two vectors is defined as $\langle a,b \rangle = \sum_{i}^{n} a_i b_i$.

The inner product $\langle f,g \rangle$ of two functions is defined as $\langle f,g \rangle = \int_{-\infty}^{\infty} \overline{f(t)} g(t) dt$.

The modulus of a vector x is symbolised by $|x|$.

The length, or norm, of a vector $\|x\| = \left(\sum_{i}^{n} x_i^2 \right)^{\frac{1}{2}}$.

The Fourier transform of a function f is symbolised by $\hat{f}$.

The set of square summable sequences is denoted by $L^2(R)$.

The union of two sets is denoted by $\cup$ and the intersection of two sets by $\cap$.

The empty set is denoted by $\{0\}$.

If the set A is a subset of the set B, then this is denoted by $A \subset B$.

$\in$ represents 'is a member of' in set theory.

If two vectors x, y are orthogonal it is denoted by $x \perp y$.

The tensor product of two spaces V_0 and V_1 is denoted by $V_0 \otimes V_1$ and the tensor sum by $V_0 \oplus V_1$.

$\Leftrightarrow$ denotes equivalence.

Table of Contents

Table of Figures

List of Tables

1. Introduction

1.1 Introduction

Seismic data are non-stationary, mainly due to localised events such as head waves and ground roll which appear at different times on the seismic shot record. In addition, attenuation of seismic signals, which ideally have a wide frequency band width, is a frequency dependent phenomenon.

In seismic data processing the Fourier transform is used extensively to improve the quality of seismic data from the initial shot record to the final stacked section. The non-stationarity of the signal can degrade the performance of some seismic processing techniques such as filtering and deconvolution which rely on Fourier transform techniques. To minimise this effect in a conventional processing sequence, spherical divergence is applied to seismic shot records, correcting for amplitude variations due to attenuation, and time-variant spectral whitening or inverse Q filtering is applied to account for frequency attenuation. In deconvolution and filtering multigate processing (a form of windowed transform processing) is also used where operators are designed over several time gates to account for variations in the frequency of the source signal with time. These techniques attempt to fit seismic data to the stationary assumption of the Fourier transform and so improve the result of these processing techniques.

The last ten years have seen the emergence of wavelet transforms as an extension to the signal processor's toolbox. Wavelets, in the form of square integrable compact, band-limited functions have been present in the literature since early in this century in the form of the Haar wavelet (Haar, 1910). In the 1980's wavelet transforms were first constructed by Goupillaud, Morlet and Grossman (1985) for geophysical processing. The transform was based on a single prototype function and its scales and shifts. The concept of scale, a scaling operation dilating or contracting the prototype function, replaces the notion of frequency. Orthonormal wavelet bases (stable bases) for square integrable and other function spaces were discovered by Meyer (1990), Daubechies (1988), Battle (1987, 1988), and Lemarié (1988) amongst others. These constructions were formalised by Mallat (1989) and Meyer (1990) creating wavelet expansions called multiresolution analysis.

In this thesis we investigate applications of the discrete wavelet and wavelet packet transforms to the suppression of coherent noise from seismic shot and common midpoint (CMP) records. The increasing emphasis on analysis of pre-stack seismic data for the detection of hydrocarbons and determination of reservoir properties has placed further emphasis on the elimination of noise from seismic traces before the stacking process. In other words, it is essential that noise is eliminated from seismic shot records and CMP gathers, whilst preserving the amplitudes of the reflection signals. Coherent noise contaminates seismic shot and CMP records in a time variant fashion and so the wavelet transform presents itself as the ideal tool.

1.2 Wavelet Transforms

The initial development of the wavelet transform concept by Goupillaud, Grossman and Morlet (1989) was not followed up by the geophysical community. Mathematicians developed and fine tuned the concept and, apart from the early papers by Goupillaud, Grossman and Morlet, references on the wavelet transform in the geophysical literature before 1992 are very few and far between. The wavelet transform has many guises, in that, there are many different formats of wavelet transform many of which have impacted on the geophysical community in the last few years. These formats can be based into two classes, forward transforms and model building transforms.

Forward transforms, such as the continuous wavelet transform, discrete wavelet transform and wavelet packet transform, decompose a signal by band-pass filtering the signal at different bandwidths. These transforms are classed as forward transforms as the processes involve the application of filters to signals, independent of the signal. That is the filters are related to the kernel wavelet of the transform not to the signal. The spectral characteristics of the filters are related to the kernel wavelet and its dilations, and therefore are not unique, in that there are a multitude of possible kernel wavelets. The transform results in a series of wavelet coefficients associated with the filtering of the original signal at different bandwidths, frequency locations, and times. A limiting factor of this technique is that for there to be a stable inverse of the transform, the kernel wavelet used in the transform must obey certain conditions. Namely that the kernel wavelet must be band-limited, have finite energy and a zero mean.

Model building transforms, such as Matching Pursuit (Mallat and Zhang, 1993), apply a wavelet transform to a signal by building a model of the signal from a redundant database of all possible permutations (translations, dilations and modulations) of a kernel wavelet. This process involves performing a fast search through the database for the basis wavelet (or atom, as they are referred to) which best matches the input signal according to some predefined criteria. This atom is then subtracted from the signal to form the model and leave the remainder of the signal, called the residual. The model consists of the first basis wavelet and is gradually added to as the searching process is performed again, this time, however, on the residual. This process is performed until the residual reaches a predefined level and the transform consists of all the atoms contained in the model plus the residual. Unlike the forward transforms, the Matching Pursuit transform is dependant on the signal and the transform process is very much governed by the transform and the choice of first best matching atom. The benefit of this transform over the forward transform techniques is that it allows the use of the optimum kernel wavelet in terms of time-frequency compactness, whilst still allowing a stable inverse.

In this research we deal with the application of the discrete wavelet and wavelet packet transforms to seismic data processing with only reviews of applications of model based transforms. Previous application of wavelet transforms, to seismic data processing are summarised in the next section.

1.3 Previous Work

In recent years there has been an explosion in the number of papers published using wavelet transforms in fields outside geophysics, and now a great deal of attention is focussed on the wavelet transform in the context of oil exploration, especially in data compression.

The wavelet and wavelet packet transforms can give sparse representations of a signal, that is, in wavelet space most of the coefficients are zero. The greater the dimension of the signal and the transform the greater the percentage of zero coefficients due to the increase in coherent signal. This has led to the increasing use of the transform in data compression, and lossy data compression (Donoho *et al.* 1995, Luo and Schuster, 1992). Two and three dimensional discrete wavelet transforms in combination with Huffman coders can obtain compression ratios of up to 100:1 whilst only having a slight effect on the quality of the final stacked section. These wavelet

compression techniques are lossy techniques. Once the compression ratio goes above a certain ratio, the compression technique leads to loss of some data. Therefore, the trade off between compression ratios and data loss must be balanced. The sparsity in the wavelet domain has also led to investigations into the effectiveness of applying migration algorithms in the wavelet domain (Wu and McMechan, 1995; Dessing, 1995; Wang and Pann, 1996). Migration is a process which can be very computer intensive and the use of matching pursuit algorithms using databases of Ricker wavelets to perform Kirchoff migration have reduced processing time considerably (Wang and Pann, 1996). Seismic inversion techniques have also been investigated in wavelet space (Bunks *et al.*, 1995; Li and Ulrych, 1995) using wavelet transform multiscale techniques. Application of these techniques allow inversion of seismic data obtained from complicated earth models.

The use of frequency-time tools for geologic interpretation allows delineation and analyses of hierarchical structures, such as sedimentary deposits and for identifying geologic cyclicity which can appear on well logs as frequency-modulated signals. Discrimination between geological layers by their frequency-time plane patterns and the grading of reflections and logging boundaries can be performed by analysis of the sharpness of events in the time-frequency plane. Use of the continuous wavelet transform for such seismic attribute analysis was discussed by Li and Ulrych (1996) and Makarov *et al.* (1996) who used it to provide such scale-location properties. The use of the matching pursuit wavelet transform for analysis of seismic data was discussed by Chakraborty and Okaya (1995) and compared to the continuous wavelet transform, discrete wavelet transform and the wavelet packet transform. The matching pursuit algorithm allowed spectral localisation, with seismic reflections, direct and surface waves clearly identifiable. In borehole geophysics Li and Haury (1995) discussed the use of the continuous wavelet transform for identifying and analysing scaling problems in sonic logs. Grubb and Walden (1997) developed the use of the discrete wavelet transform in combination with Daubechies kernel wavelets as an attribute analysis tool with which to characterise or classify groups of seismic traces in reservoir studies.

The continuous wavelet transform has also been developed as a tool for zero-phasing seismic data and as a zero-phasing quality control tool (Mansar and Rodriguez, 1994). Interpretation of seismic sections is best performed on zero-phase data and seismic data are conventionally converted to zero-phase in the Fourier domain after the stacking process. The results of this zero-phasing process are not always satisfactory

due to the non-stationarity of the signal. The continuous wavelet transform using the complex Morlet wavelet (Morlet *et al.*, 1982) can be implemented to analyse the phase of events locally. Operators can then be designed to correct any phase deviations.

In the field of noise suppression, Alkemade (1993) developed a *f-x* deconvolution technique using the discrete wavelet transform which was demonstrated to be effective on real data sets. This used the temporally localised information given by the discrete wavelet transform to improve the *f-x* deconvolution technique. During the process of this thesis research, Faqi *et al.* (1995) used the non-symmetrical Daubechies wavelets and the discrete wavelet transform for filtering seismic data pre and post stack in combination with the linear radon transform. They concluded that the transform was a useful tool for the removal of localised noise bursts and power line interference. We will discuss this work more fully in chapter two where the use of the discrete wavelet transform as a pseudo time frequency filter is developed.

Schuster and Sun (1993) used the discrete wavelet transform with symmetrical spline wavelets to successfully suppress tube waves from VSP records and extended the work to suppressing ground roll on shot records. They used the discrete wavelet transform to decompose traces in the offset direction after a linear shift. This allowed an effective method of velocity filtering. We will discuss this work more in chapter five where it is developed further.

1.4 Summary of Thesis

In investigating applications of wavelet transforms to seismic data processing we develop and test techniques on synthetic and real data examples where appropriate. This chapter has briefly introduced the topic of wavelet transforms and reviewed previous applications of wavelet transforms to seismic processing techniques. In the next chapter we review the theory behind the one-dimensional wavelet transform and introduce the wavelet packet transform which can theoretically provide higher resolution in frequency than the wavelet transform. We also discuss the implementation of these transforms on discrete signals. In chapter three we develop the use of the discrete wavelet transform as a time-frequency filtering tool and use it to suppress low-frequency ground-roll from land-based seismic shot records. This technique suffers from poor resolution in frequency-time space at higher frequencies and so in chapter four we extend this filtering technique to the wavelet packet transform. We demonstrate this technique by suppressing airblast on land-based seismic records and we also discuss

limitations associated with this technique. The use of the wavelet and wavelet packet transforms can be extended to two dimensions, and we investigate the practicalities of extending these techniques in chapter five. Here we apply two-dimensional filters to land-based seismic records to suppress ground roll and guided waves and we demonstrate the subsequent improvement in stacked section.

Finally, in chapter six, we develop the use of the one-dimensional wavelet transform as a local filter, filtering across arrays of geophones/hydrophones. We apply this technique to the suppression of guided waves from land based seismic data, swell noise from marine data and multiple energy from CMP gathers.

2. Wavelet Theory

2.1 The Fourier Transform

In this chapter we look at the theory of several types of time frequency decomposition to build a basis, so to speak, for the application of some of these techniques to the filtering of seismic data. For a more rigorous and mathematical review of the theory behind wavelet transforms the reader is directed to the books by Daubechies (1992), Vetterli and Kovacevic (1995) and Chui (1992). Reviews of the topic are also present in the literature in Jawerth and Sweldens (1994), Alkemade (1993), and Strang (1989).

The initial topic of the Fourier transform is an essential starting point in any study of wavelet transforms, and we will develop this through windowed Fourier transforms to wavelet transforms and wavelet packet transforms.

In the early 19th century, Fourier introduced the use of harmonic trigonometric series for the decomposition of signals. The Fourier transform of a signal is defined as the inner product of the signal, $f(t)$, and a basis function $\phi_\omega(t)$:

$$\hat{f}(\omega) = \left\langle f(t), \phi_\omega(t) \right\rangle, \tag{2.1}$$

decomposing the signal into the sum of these basis functions. The inner product effectively measures the similarity between the basis function $\phi_\omega(t)$ and the signal $f(t)$. The basis functions comprise complex trigonometric functions which are combined into a complex exponential:

$$\begin{aligned} \phi_\omega(t) &= \cos(\omega t) + i \sin(\omega t) \\ &= e^{i\omega t}, \end{aligned} \tag{2.2}$$

where ω is angular frequency and t time. These basis functions have infinite extent in time/space and are correspondingly perfectly compact in frequency space. The transform results in the signal being decomposed into two spectra; an amplitude spectrum and a phase spectrum. The amplitude spectrum represents the frequency content of the signal, while the phase spectrum represents the phase relationship between each of the harmonic components. The corresponding inverse of the transform defined as

$$f(t) = \left\langle \hat{f}(\omega), \phi_t(\omega) \right\rangle, \tag{2.3}$$

reconstructs the signal from the amplitude and phase spectra which contain all the information required. Any temporal information, indicating signal frequency content with time, is contained in the phase spectrum, but can be problematic to extract when the signal contains non-harmonic features, that is, when the signal is non-stationary.

When dealing with signals in computers, continuous functions are unobtainable and so some form of discretisation is required. Applying the discrete form of the Fourier transform on a discrete signal containing N samples (which is therefore band-limited from 0 Hz to the nyquist angular frequency ω_{nyq} Hz) results in the angular frequency becoming discretised with a sampling rate corresponding to,

$$\Delta\omega = \frac{2.\omega_{nyq}}{N}. \qquad (2.4)$$

Therefore, in the discrete form, the Fourier transform and the corresponding inverse become

$$\hat{f}(m\Delta\omega) = \sum_{n=0}^{N-1} f(n\Delta t)e^{-i2\pi\frac{nm}{N}}, \qquad (2.5)$$

and,

$$f(n\Delta t) = \frac{1}{N} \sum_{m=-N/2}^{N/2-1} \hat{f}(m\Delta\omega)e^{+i2\pi\frac{mn}{N}}. \qquad (2.6)$$

The discrete form of the transform and the corresponding inverse recover the signal exactly, i.e.: the transform is orthogonal. The discrete basis functions for this form of the Fourier transform have become

$$e^{i2\pi\frac{m}{N}}, \qquad (2.7)$$

which have infinite extent in time, which leads to the incorporation of any temporal information into the phase spectrum, and is therefore, effectively lost. The Fourier transform, ideal for analysing stationary signals, may not be the ideal tool for non-stationary, as with seismic data.

2.2 The Windowed Fourier Transform

To overcome the limit on temporal information supplied by the Fourier transform, the transform can be used in combination with a temporal window which is applied to the function before analysis. The window function $\vartheta(t-\tau)$, which has a specified time width and amplitude decay at the edges, is used to determine local information about the signal around the centre of the time window τ. In the windowed

form, the Fourier transform is applied to the product of the signal, $f(t)$, and the complex conjugate of the window function $\overline{\vartheta(t-\tau)}$ and so the transform becomes

$$WFT(\omega,\tau) = \int_{-\infty}^{\infty} \overline{\vartheta(t-\tau)} f(t) e^{-i\omega t} dt .\qquad(2.8)$$

On the right hand side of the integral the order of the multiplication is interchangeable, and so the transform can be considered as the inner product of windowed basis function $\vartheta(t-\tau)e^{-i\omega t}$ and signal $f(t)$. These basis functions can be considered as window functions, $\vartheta(t-\tau)$, modulated by an exponential $e^{-i\omega t}$, and form the basis for the windowed Fourier transform. An example of one of these basis functions is shown in **Figure 2-1**.
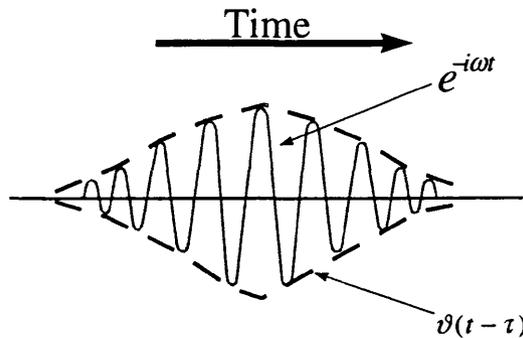
Figure 2-1 A modulated window basis function for the windowed Fourier transform

For practical applications, the continuous transform can be discretised with respect to the window functions centre time, τ, and angular frequency, ω using

$$\tau = n\tau_0 \text{ and } \omega = m\omega_0,\qquad(2.9)$$

where τ_0 and ω_0 are the temporal and angular frequency increments. This removes any redundancy in the transform and leads to the form

$$WFT_{m,n} = \int_{-\infty}^{\infty} f(t) \overline{\vartheta_{m,n}(t)} dt\qquad(2.10)$$

which is an inner product between the signal $f(t)$ and the discretised modulated window basis functions, $\vartheta_{m,n}(t) = \vartheta(t-n\tau_0)e^{jm\omega_0 t}$. From this discretisation, the temporal axis is divided into sections of width τ_0, while the frequency axis is subdivided into sections of width $\rho_0 = \dfrac{\omega_0}{2\pi}$. These widths in the time-frequency domain are related in a form similar to the Heisenberg uncertainty relation (Gabor, 1946), such

that, if the resolution (the subdivision size) along the time or frequency axes are increased there will be a corresponding decrease along the other axis as shown in **Figure 2-2**.
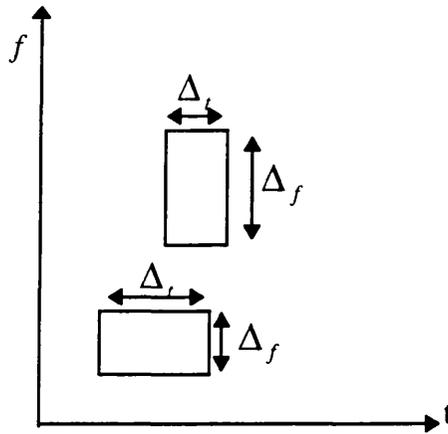
Figure 2-2 Constant area Heisenberg cells in frequency-time space.

In terms of time and frequency bandwidths, the uncertainty relation defined as

$$\Delta_t \Delta_\omega \geq \pi \tag{2.11}$$

where Δ_t is the time width of the cell, and Δ_f is the frequency bandwidth. To obtain optimum resolution in terms of temporal and frequency widths of the modulated window function, the inequality expressed in Equation 2.11 becomes an equality of its lowest possible value, π. This equality only holds for Gaussian functions

$$f(t) = \sqrt{\frac{\alpha}{\pi}} e^{-\alpha t^2} \tag{2.12}$$

where the window function $\vartheta_{m,n}(t)$ becomes

$$\vartheta_{m,n}(t) = \frac{1}{\sqrt{2\pi}} \frac{\tau_0}{\rho_0} e^{\frac{-(t - n\tau_0)^2}{2\rho^2}} e^{im\omega_0 t}. \tag{2.13}$$

This form of the windowed Fourier transform is known as the Gabor transform where the width of the window function is determined by ρ coupled with τ_0. To ensure reconstruction of the original signal $\tau_0 < 2\rho$. This window function leads to the subdivision of the time and frequency axes using a Gaussian window. The widths of the windows on the two axes are different, the time width being τ_0 and the frequency width ω_0.

2.3 Time-Frequency Tiling

Decomposition using Fourier techniques leads to a tiling of the time-frequency plane as shown in **Figure 2-3**. From this we can see that the Fourier transform lacks any temporal information. The stationarity of the Fourier basis leads to the assumption that the input data are stationary. When applied to signals that are non-stationary, such as seismic signals, this assumption is violated and the Fourier representation can only reveal the frequency content of a signal without any indication of the time varying properties of the signal.
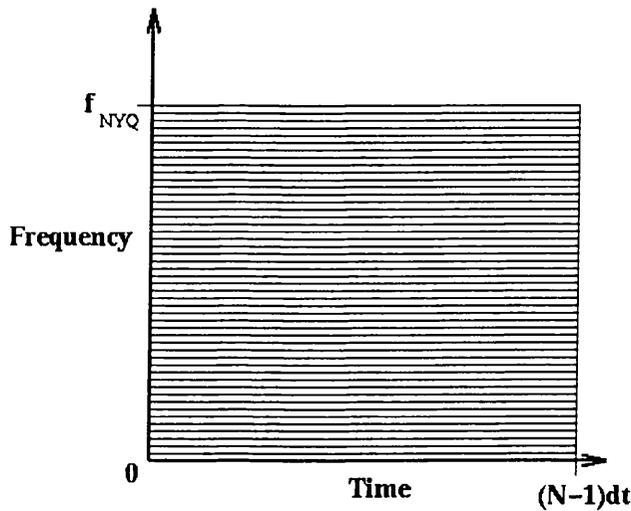
Figure 2-3 Tiling of frequency-time space by the discrete Fourier transform

The windowed Fourier basis allows a superior tiling of the time-frequency plane in terms of time-frequency representation, as shown in **Figure 2-4**. The drawback of this form of transform is the relationship between window size and the signal. **Figure 2-5** shows two windowed Fourier transforms (window lengths 500 ms and 50 ms respectively) of a synthetic seismic trace consisting of three zero phase events containing different spectral characteristics; the first has 4-8-30-40 Hz corner frequencies, the second 8-12-70-90 Hz and the third 4-6-15-20 Hz. From **Figure 2-5a** (500 ms window) we can see that for a signal with low frequency signal content a wide window function gives a satisfactory representation of the signal in the time-frequency plane whereas with a narrow window we lose resolution on the temporal axes. The opposite is true for a localised high frequency signal, a narrow time window giving the best time-frequency representation (**Figure 2-5b**, 50 ms window). When a signal contains both high and low frequency components which are localised in time, the

windowed Fourier transform fails as the window width on both the temporal and frequency axes are fixed.
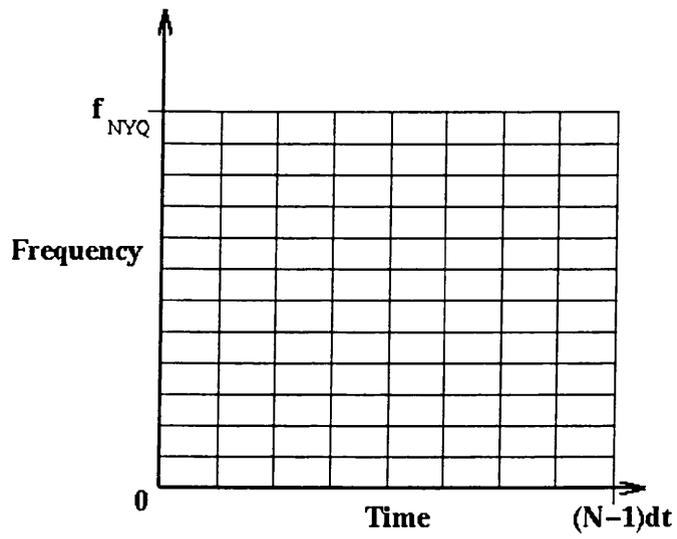
Figure 2-4 Tiling of frequency-time space by the windowed discrete Fourier transform
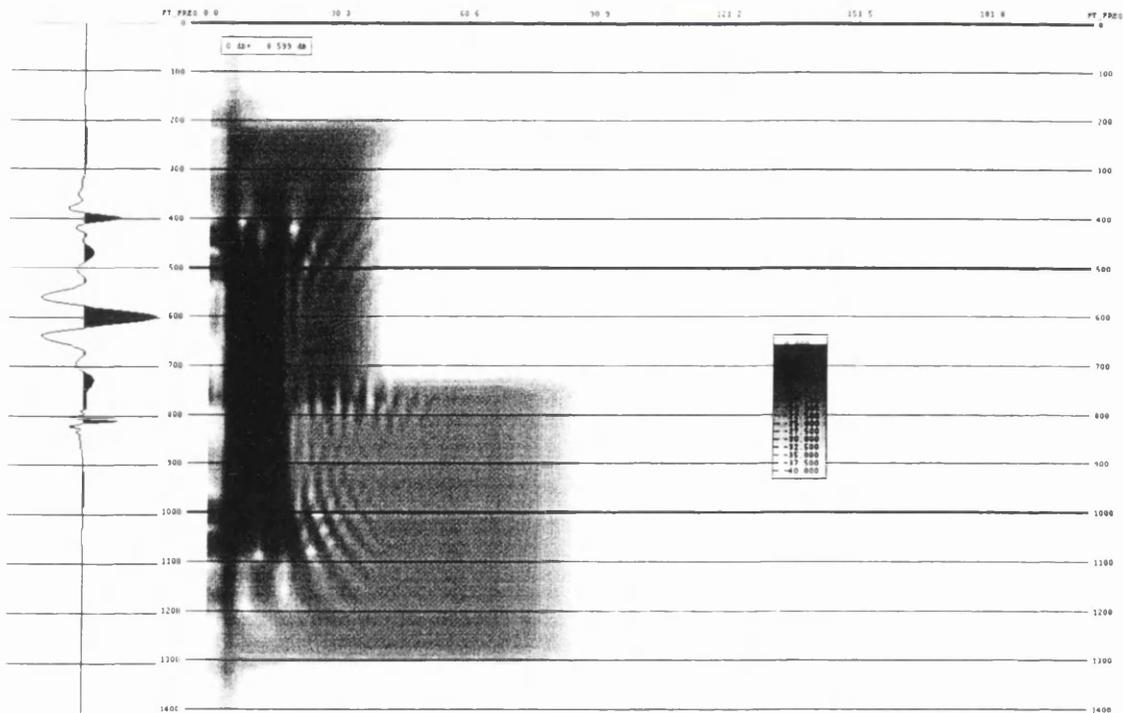
(a)
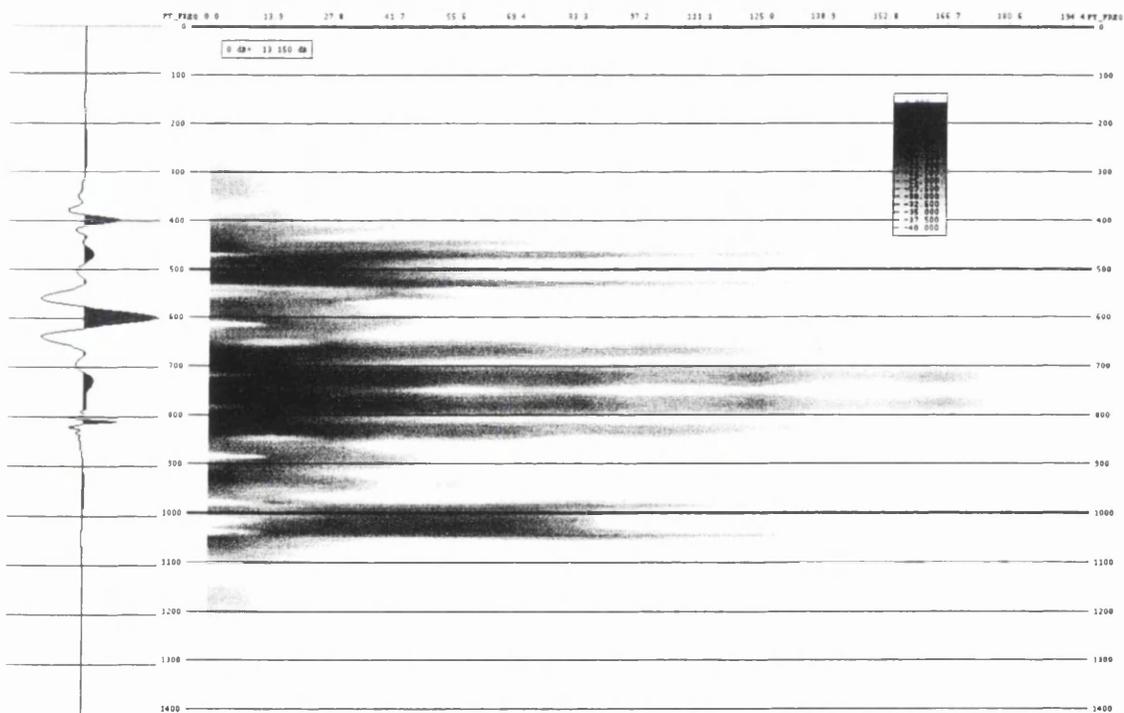
(b)

Figure 2-5 Windowed Fourier transforms of the trace shown using (a) a 500 ms and (b) a 50 ms window. The horizontal axis is frequency in Hz and the vertical time in ms. The maximum amplitude (grey) is 0 dB from the maximum, the lowest (white) -40 dB from the maximum.
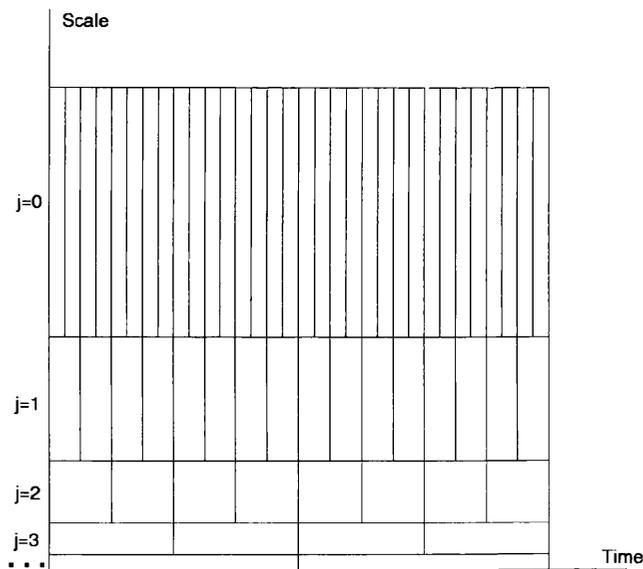
Figure 2-6 Tiling of time-frequency space by the discrete wavelet transform where the time-frequency window width adapts with the frequency being analysed. The scale axis is analogous to the frequency axis where a low scale corresponds to the highest frequency and largest frequency range.

2.4 The Wavelet Transform

Ideally for a signal that contains both high and low frequency time localised signals we would want a transform that allows an adjustment to the window length with frequency. Grossmann and Morlet (1984) introduced the concept of scaling analysis of functions through which the window changes width according to the frequency under analysis. This leads to a partitioning of the time-frequency plane as indicated in **Figure 2-6**, from which we can see that for low frequencies the temporal window has a large width and as the frequency increases the width of the temporal window decreases.

2.4.1 The Continuous Wavelet Transform

The continuous wavelet transform uses basis functions $\psi_{s,\tau}(t)$ called wavelets, which are defined as

$$\psi_{s,\tau}(t) = \frac{1}{\sqrt{|s|}} \psi\left(\frac{t-\tau}{s}\right) \tag{2.14}$$

where the function $\psi(t) \in L^2(R)$ is called the mother or kernel wavelet. The translation parameter τ controls the temporal position of the centre of the basis function (as τ_0 does with the windowed Fourier basis) and the parameter s is called the scale

parameter. The scale of a wavelet defines the temporal width of the basis wavelet, as the window length does with Fourier techniques. **Figure 2-7** shows a Morlet wavelet (Goupillaud *et al.*, 1985, Grossmann and Morlet, 1984) at several scales and translations and the corresponding Fourier amplitude spectra in **Figure 2-8**. Scale is analogous, although not directly, to frequency in Fourier analysis; the larger the scale the wider the basis wavelet is in time and so the lower the frequency, the narrower the wavelet, the smaller the scale and the larger the frequency and bandwidth. Scale more accurately corresponds to a frequency range as can be seen from **Figure 2-8**.
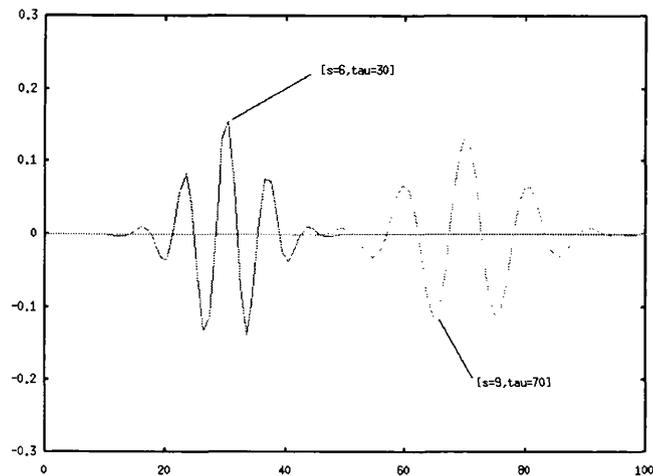
(a)
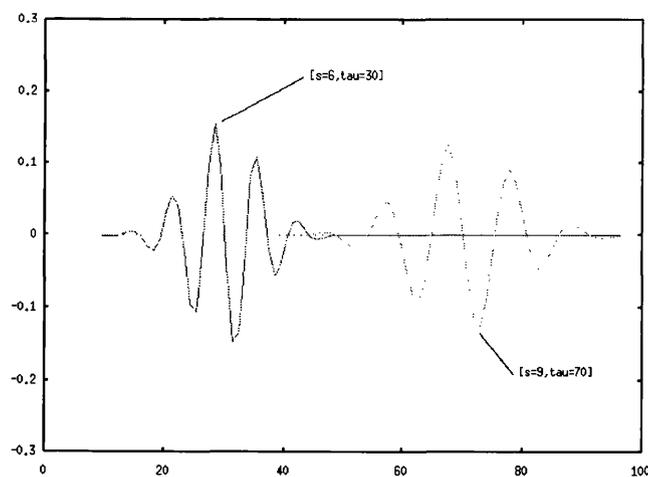
(b)

Figure 2-7 The time representation of the Morlet wavelet (a) real and (b) complex component with a scale factor s equal to 6 and 9 and at two translations with τ equal to 70 and 30.
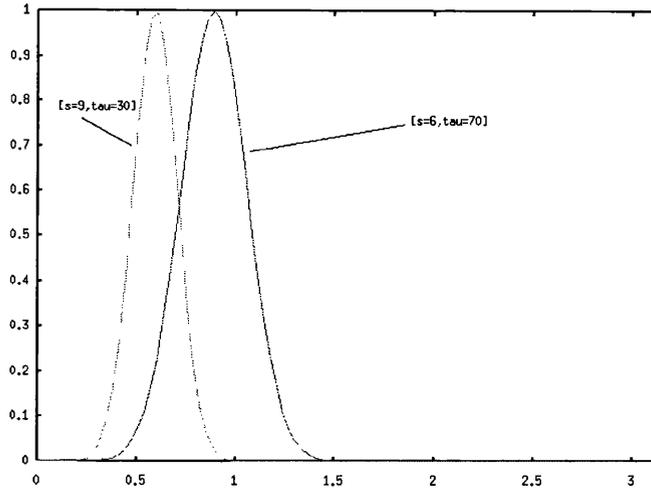
Figure 2-8 Frequency representation of the Morlet wavelet shown in Figure 2.7 showing dilation of the frequency spectrum with scaling

As for the Fourier transform, the wavelet transform is a projection which consists of the inner product of the signal $f(t)$ with the basis wavelets $\psi_{s,\tau}(t)$,

$$WT(s,\tau) = \left\langle f(t), \psi_{s,\tau}(t) \right\rangle. \tag{2.15}$$

Substituting for the basis function this leads to the transformation

$$WT(s,\tau) = \frac{1}{\sqrt{|s|}} \int_{-\infty}^{\infty} f(t) \overline{\psi\left(\frac{t-\tau}{s}\right)} dt. \tag{2.16}$$

For a given kernel wavelet function $\psi(t)$ to be admissible, it has to follow the following conditions (Daubechies, 1992):

- the kernel wavelet $\psi(t)$ must have finite energy, i.e.: the wavelet must be absolutely integrable and square integrable:

$$\int \psi(t) < \infty, \tag{2.17}$$

and

$$\int |\psi(t)|^2 dt < \infty. \tag{2.18}$$

- the kernel wavelet should be band limited and the low frequency behaviour of the Fourier transform sufficiently small around $\omega = 0$:

$$\int \left| \frac{\hat{\psi}(\omega)}{\omega} \right| d\omega < \infty, \tag{2.19}$$

This leads to the attainment of the inequality

$$C_\psi = \int_{-\infty}^{\infty} \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega < \infty, \tag{2.20}$$

where $\hat{\psi}(t)$ is the Fourier transform of $\psi(t)$. These conditions must be met by the kernel wavelet function otherwise reconstruction of the original function from its transform is not possible (Daubechies, 1992).

The corresponding inverse of the continuous wavelet transform is given by

$$f(t) = \frac{1}{C_\psi \sqrt{|s|}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} CWT(s,\tau)\psi(\frac{t-\tau}{s})\frac{d\tau ds}{s^2}. \qquad (2.21)$$

from which the reasoning behind the restrictions on the kernel wavelet become obvious.

2.4.2 The Discrete Wavelet Transform

In a basic discrete form, the wavelet transform of a signal would be evaluated for every translation step and scale increment determined by the sampling rate. In this form the wavelet transform would represent a highly redundant decomposition which would require a great deal of computational time. Adequate sampling of scale-translation space is required to remove redundancy from the transform and to maximise computational efficiency. To achieve this the scaling parameter s can be discretised such that $s = s_0^j$ ($j \in Z$) and $s_0 > 1$. The variable j is called the scaling index which in combination with the scaling factor, s_0, governs the width of the basis function in time. This transforms the basis functions to

$$\psi_{s_0^j \tau} = \frac{1}{\sqrt{|s_o|^j}} \psi\left(\frac{t-\tau}{s_0^j}\right). \qquad (2.22)$$

The transform becomes,

$$WT_{s_0^j \tau} = \frac{1}{\sqrt{|s_0|^j}} \int_{-\infty}^{\infty} f(t)\overline{\psi\left(\frac{t-\tau}{s_0^j}\right)} dt. \qquad (2.23)$$

Discretisation of the translation parameter τ, which governs the centre position of the basis wavelet in time, can be achieved in several ways. Mallat and Hwang (1992) discretised the translation parameter independently of the scale parameter, allowing the retention of local or temporal information. This is shown schematically in **Figure 2-9**. This creates a redundant transform where there is a constant number of samples in each scale which is of particular interest when estimating the local degree of irregularity of a feature, such as in edge detection. An alternative technique reduces the coverage of the temporal axis with scale by linking the number of samples to the value of the scaling parameter, altering the number of samples per scale (Mallat, 1989). For a wavelet at scale s_0^j the temporal width of the wavelet is proportional to s_0^{-j}, therefore, if the

translation parameter can be set such that $\tau = n\tau_0 s_0^j$, the resultant basis functions will become

$$\psi_{s_0^j n\tau_0}(t) = \frac{1}{\sqrt{|s_0|^j}} \psi\left(\frac{t}{s_0^j} - n\tau_0\right).$$ (2.24)

This leads to a sampling of the time-frequency plane as shown in **Figure 2-10**.
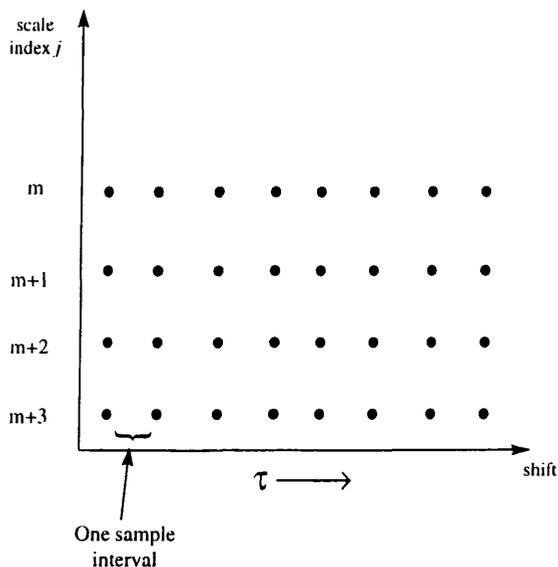
Figure 2-9 Sampling of the time-frequency plane by discretising the translation parameter independently of the scaling parameter
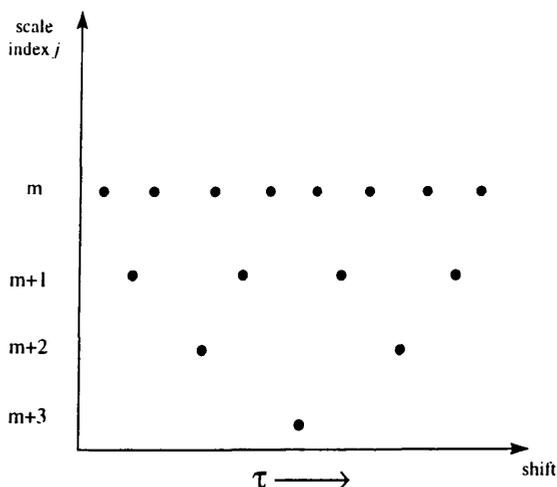
Figure 2-10 Sampling of the time-frequency plane by discretising the translation parameter in combination with the scaling parameter.

For stable reconstruction, the operator that maps a function $f(t)$ into wavelet coefficients $\langle \psi_{m,n}, f \rangle$ has to be bounded. That is, if the signal $f(t)$ has finite energy then the sum of the squares of the wavelet coefficients, $\sum_{m,n} |\langle \psi_{m,n}, f \rangle|^2$ has to be finite. Also, no signal $f(t)$ with $\|f\| > 0$ (that is, a signal with non-zero energy) should be mapped to zero. Therefore, for stable reconstruction, there are upper and lower bounds, called frame bounds, on the sum of the square of the coefficients, $\sum_{m,n} |\langle \psi_{m,n}, f \rangle|^2$. For a signal with $\|f\| > 0$, this limits the sum to be a finite, non-zero number. In mathematical terms, this is formalised by stating that any family of wavelets, $\psi_{s_0^j n \tau_0}$, with some decay in time and frequency, zero mean and for $s_0 > 1$ and $\tau_0 > 0$, form a frame if there exists two constants $A > 0$ and $B < \infty$ such that

$$A\|f\|^2 \le \sum_{j=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \left| \left\langle f, \psi_{s_0^j, n\tau_0} \right\rangle \right|^2 \le B\|f\|^2, \qquad (2.25)$$

where A and B are the frame bounds. For the original signal to be recovered exactly for a given discretised set of s and τ the basis functions $\psi_{s_0^j n \tau_0}$, A should equal B, forming what is called a tight frame (Daubechies, 1992). The value of A (or B) gives an indication of the degree of redundancy in the transform, with $A = B = 1$ indicating no redundancy and the wavelet basis constituting an orthonormal basis. If A does not equal B yet the remainder of the conditions hold, reconstruction is possible but there is an associated error in reconstruction. For $s_0 = 2$ and $t_0 = 1$, wavelets that conform with the admissibility conditions (equations 2.17 to 2.19) form an orthonormal basis (Daubechies, 1988) and the signal can be reconstructed exactly. A scaling factor of two allows fast implementation of the transform on a computer as multiplication by two is a simple shift operation on a digital number.

2.4.3 Multiresolution Analysis

The discrete form of the wavelet transform and its implementation were introduced by Mallat (1989), using the scale parameter, s_0^j and the translation parameter $\tau = n\tau_0 s_0^j$ with $s_0 = 2$ and $t_0 = 1$. The technique of applying the wavelet transform to a discrete time series, termed multiresolution analysis, splits a signal into a series of subspaces which represent a coarse approximation of the signal and the remaining detail signals.

More formally, a multiresolution analysis of $L^2(R)$ is defined as a sequence of embedded closed subspaces V_j of $L^2(R)$ ($j \in Z$)

$$\{0\} \cdots \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \cdots \subset L^2(R) \tag{2.26}$$

with the following properties (Vetterli and Kovacevic, 1995):

1. $\bigcup_{j \in Z} V_j = L^2(R)$. $\hspace{6cm}$ (2.27)

   The union of all subspaces should cover the space of square-summable sequences. This requires that the combination of the frequency ranges of the basis wavelets at every scale should completely cover the frequency space of the signal from 0 Hz to the Nyquist frequency, leaving no gaps.

2. $\bigcap_{j \in Z} V_j = \{0\}$ $\hspace{6.5cm}$ (2.28)

   The subspaces should not intersect.

3. $g(t) \in V_j \iff g(2t) \in V_{j+1}$, $\hspace{4.5cm}$ (2.29)

   The embedded subspaces are related by a scaling law which states that if the basis function g(t) is a member of a subspace V_j then after dilation by a factor of two, the scaled basis function is a member of the neighbouring subspace, V_{j+1}.

4. Each subspace is spanned by integer translates of a single function $g(t)$ such that

$$g(t) \in V_0 \iff g(t + 1) \in V_0. \tag{2.30}$$

   This states that if the function $g(t)$ is a basis function for the subspace V_0, then all translations of the function $g(t+n)$ are basis functions also.

A scaling function $\phi(t) \in V_0$ is then required such that its integer translates $\{\phi(t - \tau), \tau \in Z\}$ form an orthonormal basis for the space V_0. Therefore, from equation 2.29 we can derive a scaling function $\phi(2t - \tau)$ that will form an orthonormal basis for the subspace V_1. Since the space V_0 lies within the space V_1 as shown schematically in **Figure 2-11**, we can express any function in V_0 in terms of the basis functions of V_1. In particular,

$$\phi(t) = \sum_{\tau=-\infty}^{\infty} a_\tau \phi(2t - \tau) \tag{2.31}$$

in which $a_k, k \in Z$, is a square summable sequence (i.e.: is finite). The coefficients a_τ are called filter coefficients and it is often the case that only a finite number of these are non-zero. Equation 2.31 is referred to as the dilation equation. Defining

$$\phi_{s,\tau} = 2^{s/2}\phi(2^s - \tau)$$ (2.32)

then $\phi_{s,\tau}(t), k \in Z$ forms an orthonormal basis for the subspace V_s, where s, as before, is the scale parameter. If we take the difference between subspaces V_{s+1} and V_s, we can define a new subspace W_{s+1} such that it is the orthogonal complement of V_{s+1} in V_s,

$$V_s = V_{s+1} \oplus W_{s+1}, \quad V_{s+1} \perp W_{m+1}$$ (2.33)

where $\oplus$ represents a direct sum. From this it follows that the spaces W_s are orthogonal and that

$$\underset{j \in Z}{\oplus} W_j = L^2(R).$$ (2.34)

Figure 2-11, represents this concept schematically for basis wavelets which are sinc functions. The sinc function is a permissible basis wavelet (Jawerth and Sweldens, 1994) but is not practical due to its very slow decay in time, yet is ideal for illustration purposes. The kernel wavelet, $\psi(t)$, is defined as the function that forms the basis of the space W_0

$$\psi_{s,\tau} = 2^{s/2}\psi(2^s t - \tau) .$$ (2.35)

Therefore if $\{\psi(t - \tau), \tau \in Z\}$ form an orthonormal set, then it follows that $\{\psi_{s,\tau}, s, \tau \in Z\}$ form an orthonormal basis for $L^2(R)$.
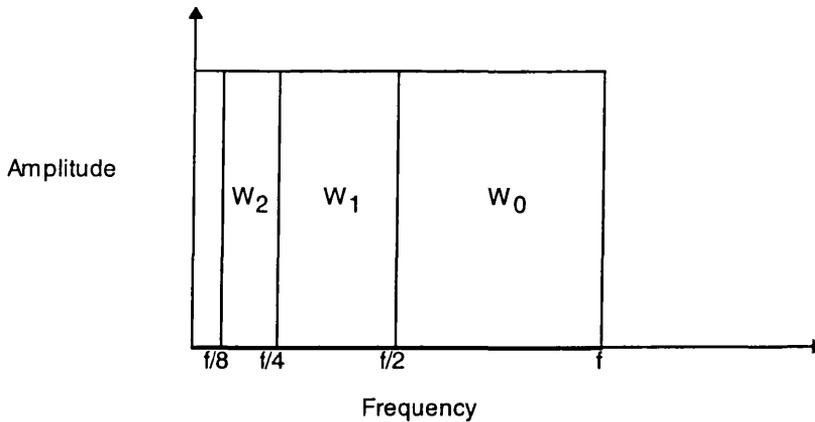
Figure 2-11 Partitioning of frequency space by a discrete wavelet transform using a sinc basis wavelet. The transform partitions frequency space into octave bands, each scale i of basis wavelet spanning an octave subspace W_i and each corresponding scaling function spanning the nested subspaces V_i (adapted from Vetterli and Kovacevic, 1995).

As the wavelet transform is a projection, let us denote the projection of f on a wavelet subspace, W_s as $Q_s f$ (**Figure 2-12**) and $P_s f$ the projection of f on the scaling subspace V_s. Then from equation 2.33, we have

$$P_s f = P_{s+1} f + Q_{s+1} f \, , \tag{2.36}$$

which shows that $Q_s f$ represents the detail that needs to be added to get from one level of approximation, $P_{s+1} f$, to the next finer approximation, $P_s f$.
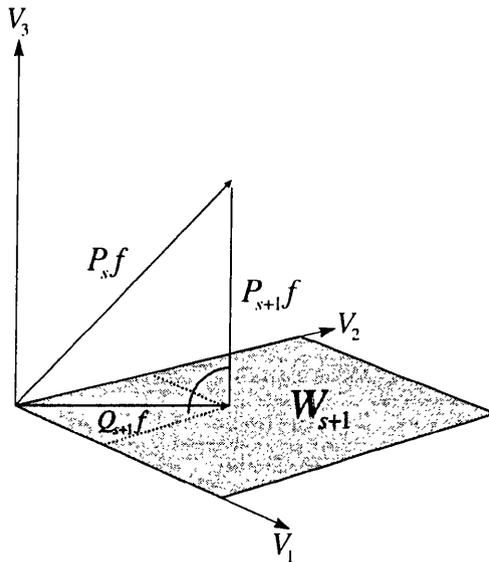
Figure 2-12 The wavelet transform is a projection of the vector $P_s f$ onto the wavelet space W_{s+1} giving the vector $Q_{s+1} f$ in the wavelet space and $P_{s+1} f$, the coarse approximation of $P_s f$.

Since the space W_0 is contained within the space V_1, we can express the wavelet function in terms of the scaling function at the next higher scale,

$$\psi(x) = \sum_{\tau=-\infty}^{\infty} b_\tau \phi(2t - \tau) \, . \tag{2.37}$$

The multiresolution decomposition takes the coefficients of an approximation to the original signal at scale s and decomposes them into

(1) coefficients, $c_{s+1,\tau}$ of $P_{s+1} f$, the approximation to the signal at the next coarser scale $s + 1$

(2) coefficients, $w_{s+1,\tau}$ of $Q_{s+1} f = P_s f - P_{s+1} f$ the detail component.

This process shown diagrammatically in **Figure 2-13** can then be repeated on the coefficients of the approximation to find the next set of coefficients $c_{s+2,\tau}$ and $w_{s+2,\tau}$ and so on. Therefore, we can see that the multiresolution decomposition breaks down the original $L^2(R)$ space into a series of orthogonal subspaces at different resolutions.

Equation (2.31) defined the filter coefficients which are used in the multiresolution decomposition. A wavelet, $\psi(t)$, is orthogonal to the scaling function and is defined by

$$\psi(t) = \sum_{\tau=-\infty}^{\infty} (-1)^\tau a_{N-1-\tau} \phi(2t - \tau) \tag{2.38}$$

where N, the number of scaling filter coefficients, is an even integer. The sets of coefficients $h = \{a_\tau\}$ and $g = \{(-1)^\tau a_{N-1-\tau}\}$ form a pair of quadrature mirror filters. These quadrature mirror filters (QMF) are a pair of half-band high pass and low pass filters which split the data into the smooth part and the detail part. **Figure 2-14** shows the QMF pairs and associated impulse response for a common basis wavelet.
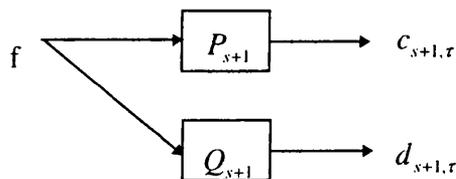
Figure 2-13 Schematic diagram of the decomposition of the signal f into the coarse approximation c and the detail signal d.

2.4.4 Transform Implementation

The discrete wavelet transform is implemented using a fast transform ($N\log(N)$ operations for a N sample signal) using the QMFs associated with the basis wavelet. The technique is indicated schematically in **Figure 2-15** and explained in the following text.

Each iteration of the transform step involves the splitting of the input signal into two output signals. One signal is filtered by the high-pass half band filter and the other by the low pass half band filter producing the projections indicated in equation 2.36. The output of the low pass filter gives the smoothed version of the input signal $P_{s+1}f$ and the resulting halving of the Nyquist frequency allows downsampling by two of the signal. The output of this step is referred to as scaling coefficients at that scale (scale 0, if it is the first iteration). The output of the high pass filter gives the detail signal $Q_{s+1}f$ which is the difference between the input signal and the original. This signal is

also downsampled by two and gives the wavelet coefficients at that scale (scale 0, if it is the first iteration). The same filtering and downsampling process is then applied to the scaling coefficients and constitutes the next iteration of the transform. This process is applied in a cascaded fashion until only one scaling coefficient remains.
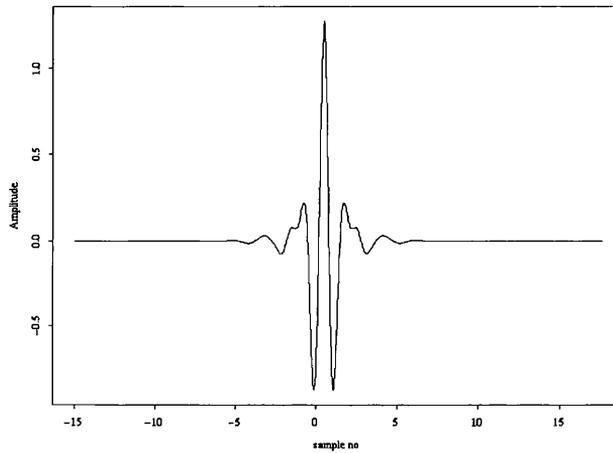
(a)
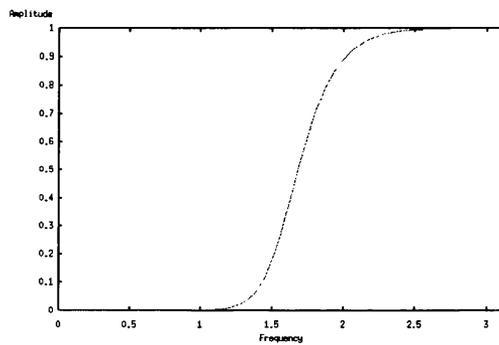
(b)
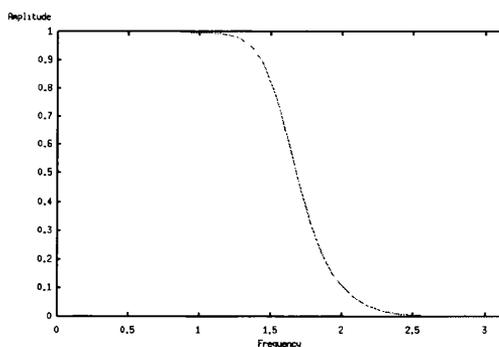
(c)

Figure 2-14 (a) A basis wavelet built from cubic splines (b) The associated high pass and (c) low-pass quadrature mirror filters.
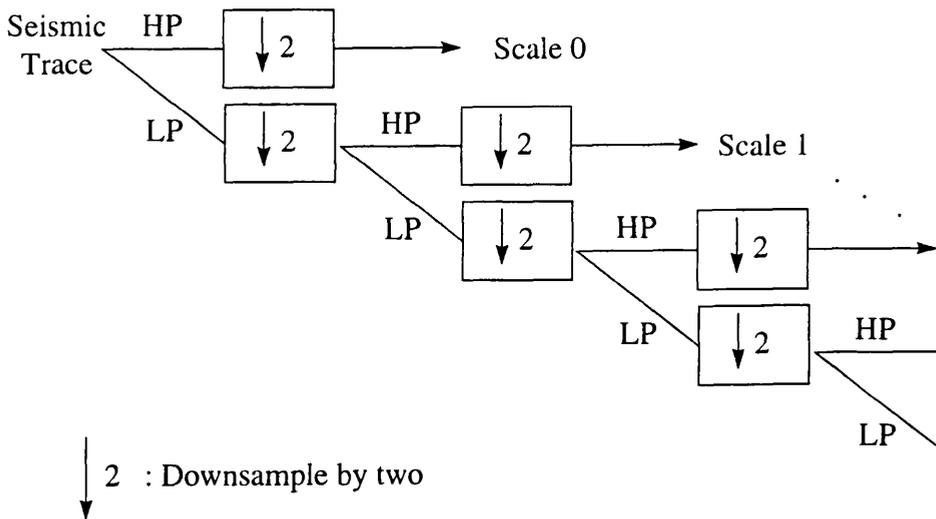
Figure 2-15 Schematic diagram showing the cascade process used in the 1-D discrete wavelet transform achieved using a pair of quadrature mirror filters HP (High pass half band) and LP (low pass half band) followed by downsampling by two.
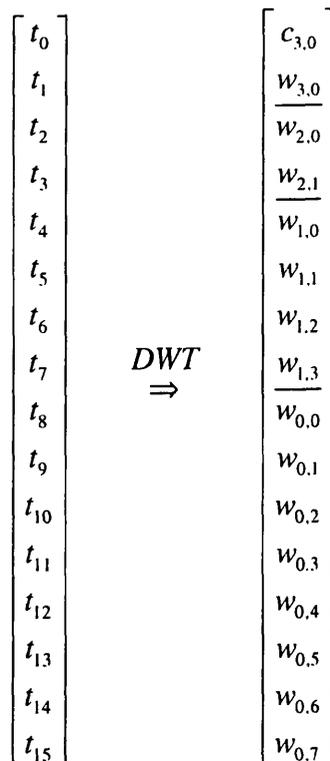
Figure 2-16 Transformation of the input signal by a discrete wavelet transform (DWT). Variable t_n represents the nth time sample, $w_{i,j}$ a wavelet coefficient at scale i and translation j within the scale, and $s_{i,j}$ a scaling coefficient at scale i and translation j within the scale

A requirement of this process is that the signal length is an integer power of two (as for the fast Fourier transform), and is required as a result of the downsampling process. The full transformation of a sixteen sample signal leads to a transformation of the input signal as represented in **Figure 2-16**. For our ideal boxcar high/low bandpass filters corresponding to a sinc basis wavelet this leads to a tiling of the time frequency plane shown earlier in **Figure 2-11**. However, from **Figure 2-17** which shows the frequency support of a wavelet basis function at two neighbouring scales we can see that there is overlap in frequency space between scales and so the boundaries shown in **Figure 2-6** are fuzzy. This frequency overlap between scales when combined with the downsampling in the transform process, leads to a form of aliasing which is accounted for in the inverse transform, allowing perfect reconstruction. If, however, the wavelet transform coefficients are altered in some way in wavelet space (in filtering, for example) then this aliasing may not be fully accounted for and aliased noise may be introduced.
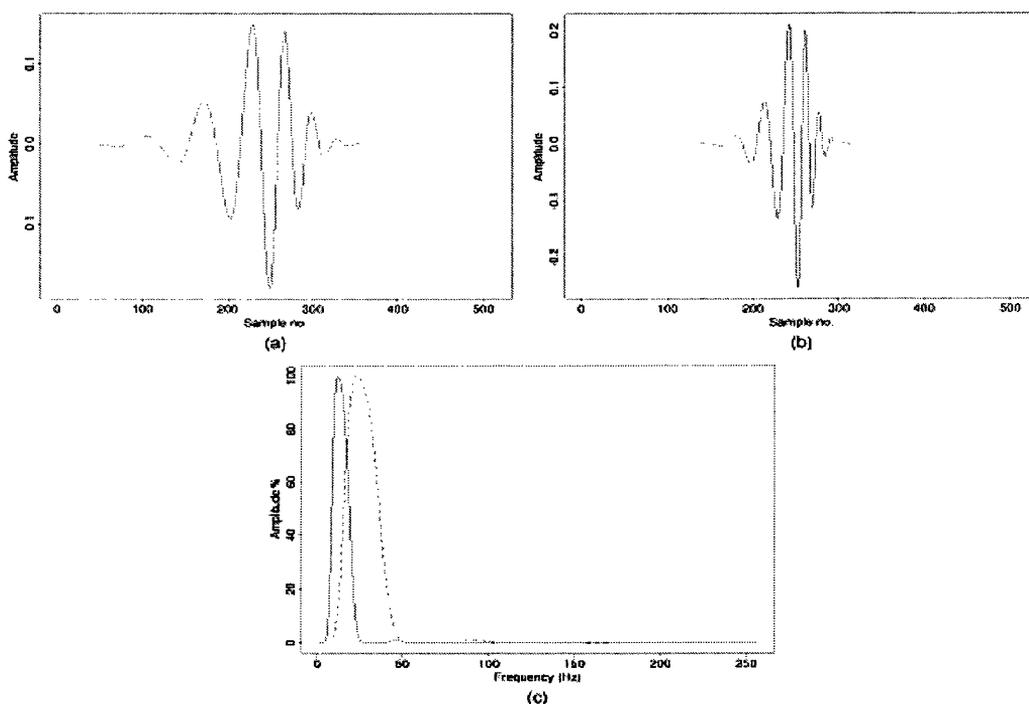
Figure 2-17 Daubechies 20 coefficient maximal phase wavelet at two adjoining scales in time and frequency space. Overlap of the frequency spaces occurs when the wavelet deviates from the boxcar frequency representation of the sinc wavelet.

2.4.5 Variation with Translation

Another important consequence of the downsampling process on the discrete wavelet transform of a signal, is translational variance. If we take two identical signals, shifting one by one sample with respect to the other, and take the discrete wavelet transform of the two signals, the resultant transformation of the shifted signal will not be the transform of the first shifted by one sample. The transform alters for shifted identical signals. The downsampling of frequency space to a coarse grid in the wavelet transform leads to the variation of the transform with translation.

2.5 The Wavelet Packet Transform

A simple but powerful extension of the wavelet transform is the wavelet packet transform (Coifman *et al.*, 1992). The wavelet packet transform takes the wavelet transform further by applying the quadrature mirror filters h and g to the wavelet coefficients as well as the scaling coefficients at each iteration as illustrated in **Figure 2-18**. This technique, referred to as the splitting trick, leads to several levels of resolution in frequency-time, each level being a orthonormal basis. In the z-domain after one iteration the basis functions are

$$W_0^1(z) = G(z), \text{ and } \quad W_1^1(z) = H(z).$$ (2.39)

At the next level, the basis functions become:

$$W_0^2 = G(z)G(z^2), \ W_1^2 = G(z)H(z^2),$$ (2.40)

$$W_2^2 = H(z)G(z^2), \text{ and } W_3^2 = H(z)H(z^2),$$ (2.41)

doubling the number of basis functions at each successive level. The wavelet packet decomposition is in the form of a linear weighted sum of these basis functions,

$$x(t) = \sum_f \sum_s \sum_{p=-\infty}^{\infty} A_{f,p}^s W_{f,p}^s(t)$$ (2.42)

where $A_{f,p}^s$ is the wavelet packet coefficient that is the inner product of the signal $x(t)$ with the wavelet packet $W_{f,p}^s(t)$ having selected values of the packet scale index f, the level index s, and the translation parameter p.
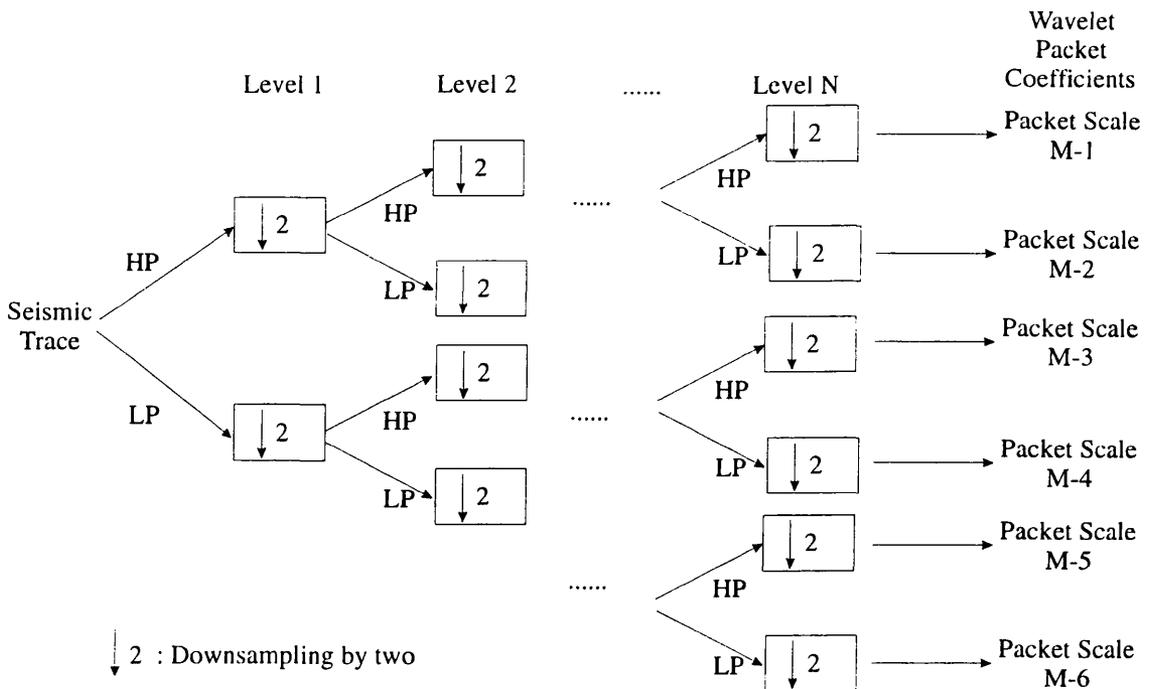
Figure 2-18 Schematic diagram showing the cascade process used in the 1-D discrete wavelet packet transform achieved using a pair of quadrature mirror filters HP (High pass half band) and LP (low pass half band) followed by downsampling by two. Comparing this to Figure 2-15 it is apparent that the wavelet packet transform applies the splitting trick to what were the wavelet subspaces.

As a result the wavelet packet transform tiles the time-frequency domain in a form shown in **Figure 2-19**. The wavelet packet transform results in an overcomplete expansion of the original signal in that for any signal of length N there are $N\log(N)$ wavelet packet coefficients. For a fixed level l the wavelet packets are orthonormal for s and p and so each level represents an orthonormal basis of the input signal. For example the second level from Figure 1 corresponds to the basis:

$$W = \left\{ W_{0,p}^2, W_{1,p}^2, W_{2,p}^2, W_{3,p}^2 \right\}_{p \in Z}.$$

(2.43)

Subsets can also be selected across levels to form orthonormal bases such as the wavelet basis:

$$W = \left\{ W_{0,p}^1, W_{2,p}^2, W_{3,p}^7, \ldots\ldots \right\}_{p \in Z}$$

(2.44)

which is represented by the shaded areas in **Figure 2-19**. From this array of coefficients an orthonormal basis can be selected according to some predetermined criteria such as minimum entropy, or minimising the number of non-zero coefficients. This allows varied representations of a single function which can be tailored to specific purposes such as data compression or time-frequency representation.
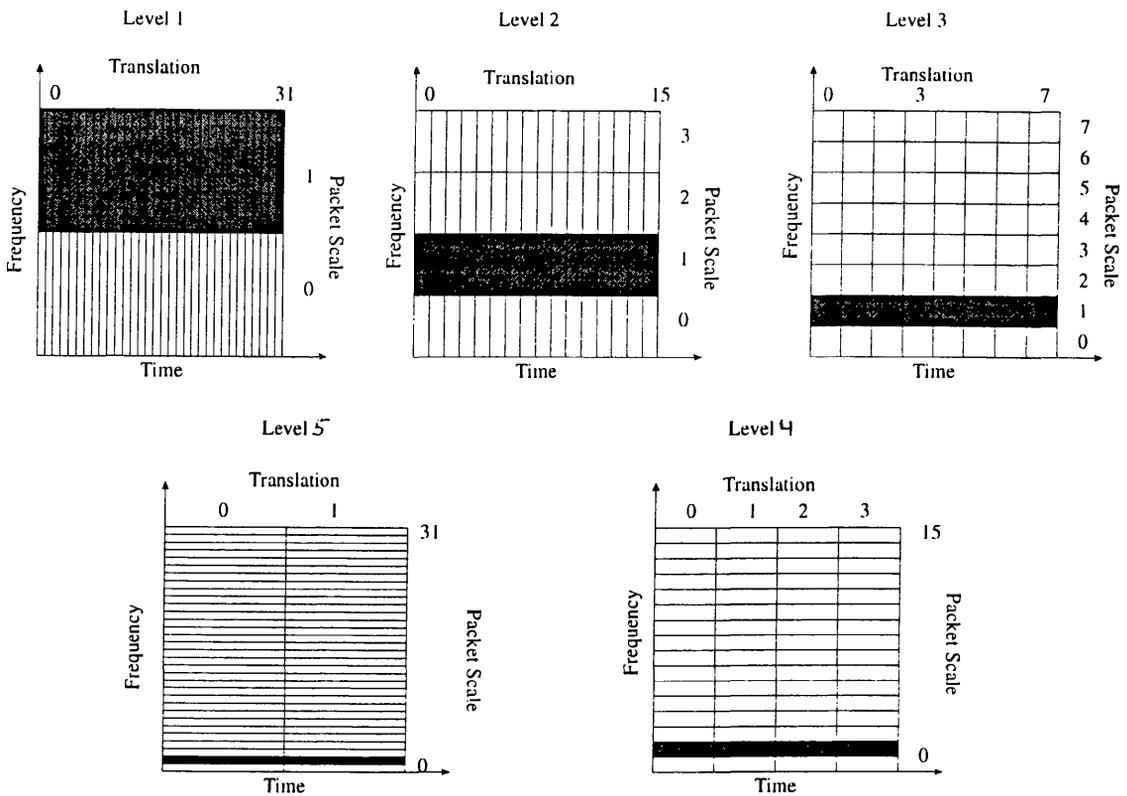
Figure 2-19 Tiling of time-frequency space by the wavelet packet transform. Each successive level leads to a theoretical increase in frequency resolution and a corresponding decrease in temporal resolution

As a result, the wavelet packet transform theoretically allows a higher resolution in certain areas of frequency space than the wavelet transform. This higher resolution in frequency is at the expense of resolution in time but also leads to a more flexible form of time-frequency representation. Although, as with the wavelet representation, is not exactly true time-frequency (the overlap between basis functions can be larger for wavelet packet bases) this representation is a useful tool for displaying the transform.

The wavelet packet transform allows more flexibility in frequency-time resolution than the wavelet transform at the expense of the extra processing time required for the extra iterations. The wavelet packet transform can be done in-situ like the wavelet transform if the transform is given prior information on the basis required. Otherwise the transform needs to generate all levels of the wavelet packet transform and then select the desired basis from the resulting decomposition. **Figure 2-20** shows the impulse response for a wavelet packet generated using the same quadrature mirror filters used for **Figure 2-14**.
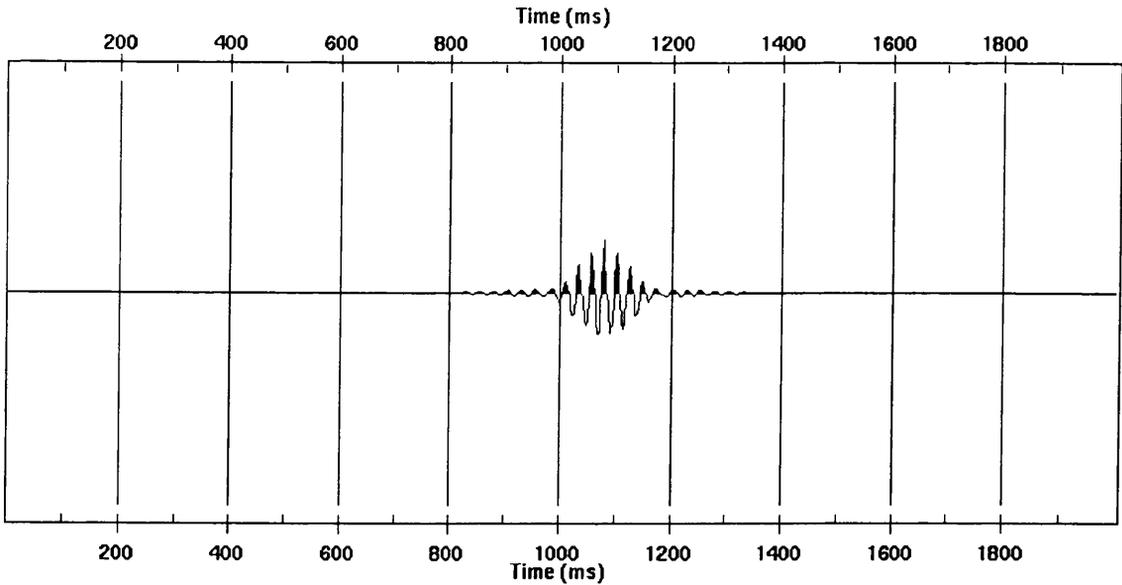
Figure 2-20 Impulse response of a wavelet packet built using the basis wavelet shown in Figure 2-14.

2.6 Summary

We have seen that the wavelet transform decomposes a signal in a time-scale sense using dilations and translations of a single prototype wavelet. Further decomposition of scales of the wavelet transforms leads to the wavelet packet decomposition which allows a flexible sampling of the time-frequency plane. These techniques can be implemented by fast efficient algorithms on digital signals, allowing perfect signal reconstruction for certain basis wavelets.

These time-scale decompositions allow effective analysis of non-stationary signals and in the following chapters we investigate the use of these techniques as time-scale filters where the scale is analogous to frequency. Chapter three investigates the use of wavelet transforms for one-dimensional time-scale filtering of seismic data.

3. Wavelet Transform Filtering

3.1 Introduction

In this chapter we investigate the use of the discrete wavelet transform as a tool for filtering seismic data in the scale-translation domain. Filtering of seismic data in the scale-translation domain of the wavelet transform can be thought of as a form of time-frequency filtering, in that the transform decomposes time-frequency space. This analogy is not exact for several reasons; scale is a frequency range, there is overlap between adjacent scales in the frequency domain (as discussed in Chapter 2), the translation increment is linked to scale and so is scale dependant. We will investigate the effects of these deviations from the time frequency concept in terms of filter performance. The discrete wavelet transform is not unique, in that there is an infinite choice of possible kernel wavelets that could be used in the decomposition process. The time-frequency properties of any kernel wavelet used for wavelet decomposition will influence the values of the wavelet transform coefficients, thus influencing the performance of wavelet transform based filters. We therefore also investigate the influence of choice of kernel wavelet on filter performance and determine the principal properties that influence filter performance.

We then demonstrate the use of the wavelet transform as a pseudo time-frequency filter for the suppression of ground-roll energy from a land-based seismic survey. Ground roll contaminates seismic data in a time-varying fashion and so the wavelet transform presents itself as the ideal tool. The scale-translation properties of the wavelet transform allows filtering of specific time-frequency zones contaminated by ground-roll, leaving the remainder of the data unaltered. Although there is not perfect signal/noise separation in the wavelet transform domain, filtering in the transform domain allows noise suppression whilst minimising deterioration of the signal component, as each coefficient represents a limited temporal area. We then proceed to demonstrate the effect of scale-translation filtering on subsequent stacks and make comparisons with bandpass and *f-k* filtering.

3.2 Choice of Kernel Wavelet

In Chapter 2 we introduced the scalogram representation of the wavelet domain

to visualise transformed data, dividing frequency-time space into a series of cells known as Heisenberg cells. We have seen that the frequency and time support of the basis wavelets extends beyond the boundaries of these rectangular cells, the degree of which depends on the kernel wavelet used in the decomposition process. The optimal kernel wavelet for compactness in both time and frequency is the Gaussian (Chakraborty and Okaya, 1995) which, when used as kernel wavelets, are referred to as Gabor atoms. Implementing the fast wavelet transform using this kernel wavelet is not possible as it does not allow stable reconstruction of the original signal. The non-linear Matching Pursuit algorithm (Mallat and Zhang 1993) is required to provide a stable forward and inverse transform. Therefore, we limit the study of kernel wavelets to those that are permissible in terms of the fast wavelet transform, that is those that satisfy equations 2.17 to 2.19.

For the purposes of filtering data in the wavelet domain, we must determine the zone of influence, in the temporal and frequency domains, of any wavelet coefficient at a given scale in the transform domain for any given kernel wavelet. In the temporal domain, the zone of influence of the wavelet at a given scale is governed by the temporal support of that wavelet. That is, the length in time of the impulse response of the wavelet at the given scale. Vetterli and Kovacevic (1995) derived that if a kernel wavelet $\psi(t)$ is compactly supported on the interval $\left[-n_1, n_2\right]$ then the basis wavelet at a given scale m and translation n, $\psi_{m,n}(t)$, is supported on the interval $\left[(-n_1 + n)2^m, (n_2 + n)2^m\right]$. Therefore, at scale m a wavelet coefficient with translation n will influence samples of the signal within the temporal zone:

$$(-n_1 + n)2^m \leq t \leq (n_2 + n)2^m. \tag{3.1}$$

Similarly, in the frequency domain, if the support of the kernel wavelet $\psi(t)$ is $\left[\omega_{min}, \omega_{max}\right]$, at scale m, the support of basis wavelet $\hat{\psi}_{m,n}(\omega)$ will be $\left[\dfrac{\omega_{min}}{2^m}, \dfrac{\omega_{max}}{2^m}\right]$. Therefore, at scale m, a wavelet coefficient will influence frequency components of the original signal within the range

$$\frac{\omega_{min}}{2^m} \leq \omega_0 \leq \frac{\omega_{max}}{2^m}. \tag{3.2}$$

These relationships are indicated diagrammatically in **Figure 3-1**.
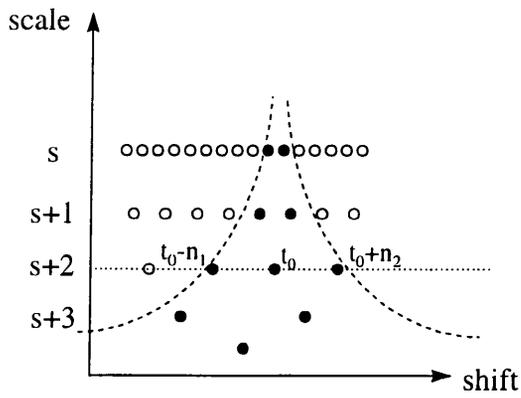
(a)
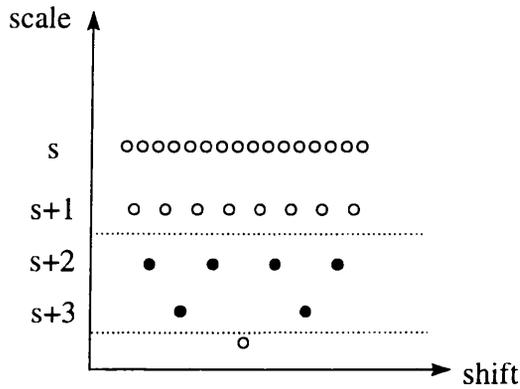
(b)

Figure 3-1 (a) Region of wavelet coefficients which will be influenced by the value of the function at t_0. (b) Region of influence of the Fourier component. Adapted from Vetterli and Kovacevic, 1995.

Figure 2.17 demonstrated that the frequency support of a basis wavelet $\psi(t)$ at two adjoining scales can overlap, this is also true for the optimum, Gabor atoms. Unless a given kernel wavelet consists of a perfectly compact function in the frequency domain the support of the kernel wavelet $[\omega_{min}, \omega_{max}]$ can be subjective. The frequency at which the amplitude becomes negligible, can be dependant on the signal being analysed, and is therefore hard to define. To overcome this problem we define the support $[\omega_{min}, \omega_{max}]$ as the frequencies for the ideal basis wavelet in frequency space, the sinc function, keeping in mind that there is frequency overlap between neighbouring scales. This overlap in combination with the downsampling process inherent in the fast transform process leads to a form of aliasing in the transform domain which is accounted for perfectly in reconstruction. If, however, the wavelet coefficients of the seismic data are filtered (and so their amplitudes altered), this aliasing may not be accounted for fully in reconstruction and may introduce noise associated with the aliasing process into any filtered signal. Subsequently, the choice of the $[\omega_{min}, \omega_{max}]$ corresponding to the sinc function as the limits of the kernel wavelet support in the frequency domain during the filtering process is justifiable if we minimise frequency overlap between neighbouring scales.

To do this we must minimise the frequencies outside of the ideal frequency support region, and so choose a basis wavelet that has a suitably fast fall off in frequency space. Ideally, the wavelet must also have a flat response across the band of interest ($[\omega_{min}, \omega_{max}]$) as in the ideal sinc function case. **Figure 3-2** to **Figure 3-4** show several Daubechies basis wavelets and their corresponding frequency representation at a set scale. Marked on the figures are the limits of the Heisenberg cells from the scalogram representation. From this we can see that the four coefficient kernel wavelet is not suitable for filtering due to frequency amplitude response across the support range, which is not constant, and also when compared to the ideal frequency support limits indicated on the diagram. For the Daubechies kernel wavelets, as the number of coefficients of the associated quadrature mirror filters increases, the smoother the wavelets become in time and the sharper the amplitude fall-off in frequency space at scale boundaries. This minimises the quantity of energy outwith the main frequency band. However, the amplitude response across the frequency range remains curved and the extent of the wavelet outside the limits of the Heisenberg cell in time increases. A repercussion of requiring a flat response and rapid fall-off in frequency space is, that the resultant number of quadrature mirror filter coefficients associated with the basis

wavelet increases with rapidity of the fall-off and so the temporal support of the wavelet increases. This in turn increases the number of calculations required in the transform process. Subsequently, we must make a trade off between speed of transform and the rate of frequency amplitude fall-off.
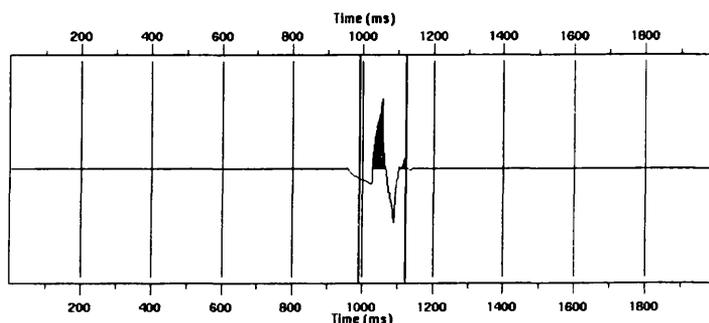
(a)
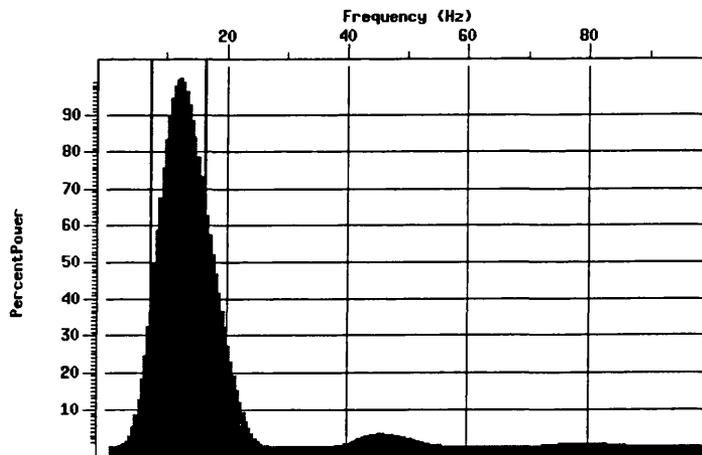
(b)

Figure 3-2 (a) The time and (b) frequency form of the four coefficient Daubechies wavelet at one scale. The time and frequency limits of the Heisenberg cell associated with this scale are marked on the figure as vertical dark lines.
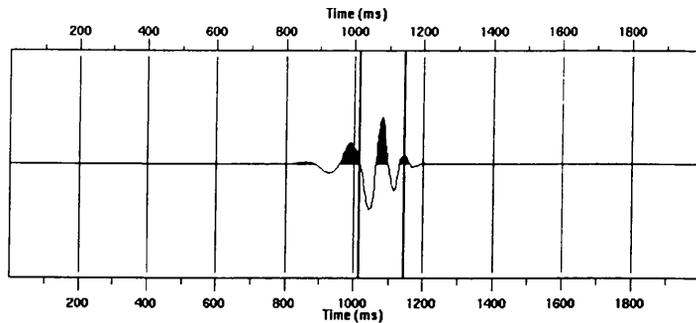
(a)
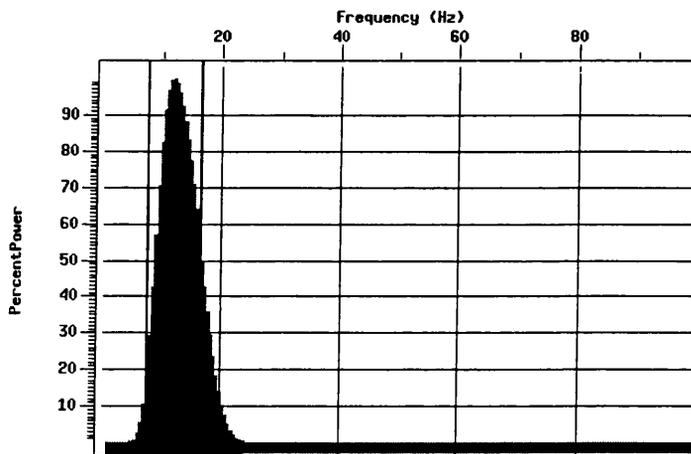
(b)

Figure 3-3 (a) The time and (b) frequency form of the twelve coefficient Daubechies wavelet at one scale. The time and frequency limits of the Heisenberg cell associated with this scale are marked on the figure as vertical dark lines.
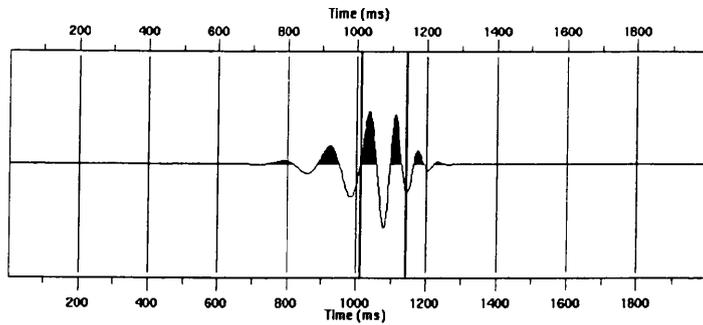
(a)
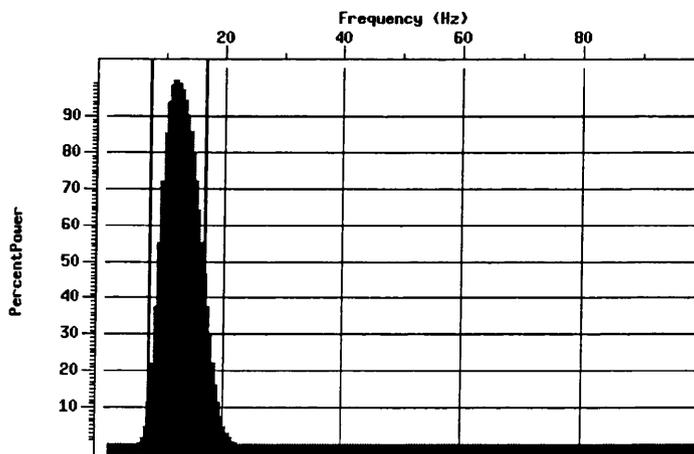
(b)

Figure 3-4 (a) The time and (b) frequency form of the twenty coefficient Daubechies wavelet at one scale. The time and frequency limits of the Heisenberg cell associated with this scale are marked on the figure as vertical dark lines.

A further requirement of the kernel wavelet is that it should have linear phase, i.e.: it should be symmetrical. For operations in the wavelet domain, deviations from linear phase will lead to phase distortions in the signal and subsequent shifts of reflected events (Chui, 1992). Daubechies (1992) showed that it was impossible for a perfectly compact basis wavelet to be symmetrical whilst the associated quadrature mirror filters, h and g, were of the same length. However, a subset of the Daubechies wavelets exists which have the minimum possible deviation from zero phase. Examples of these wavelets are shown in **Figure 3-5** and **Figure 3-6.** From this we can see that the frequency support of these wavelets is the same as for the most asymmetrical Daubechies wavelets with the same number of coefficients.

The Battle-Lemarié family of basis wavelets (Battle, 1987, Lemarié 1988) are admissible kernel wavelets which can be designed to have linear phase and suitably fast fall-off in frequency space, but are not perfectly compact. In other words, the basis wavelets have large temporal support yet have suitably fast decay in time and frequency to be a permissible basis (Mallat, 1989). The fast decay in time means that the kernel wavelet is represented by a large number of coefficients but in practice can be approximated a small number of filter coefficients without introducing large errors into the signal. Obviously, the larger the number of coefficients used, the more accurate the reconstruction, but the slower the algorithm. This trade off for temporal support is similar to the trade off required in the frequency domain.

In terms of a one dimensional wavelet transform using a cubic spline Battle-Lemarié kernel wavelet (**Figure 3-7**), twenty four coefficients is sufficient to transform and reconstruct the trace with an error of 0.5%. From the frequency support of the wavelets we can see that these wavelets are also suitable in terms of minimising frequency overlap, when compared to Daubechies wavelets. The quintic spline member of the Battle-Lemarié family (**Figure 3-8**) is also a possible kernel wavelet, yet has a wider temporal support in the time domain. In the frequency domain, the amplitude fall-off is quicker thus minimising any noise associated with transform aliasing by a greater degree than for the cubic kernel wavelet. The amplitude response is also flatter. The wider temporal support results in more coefficients for the filters slowing the filtering process, and also results in the wavelet extending beyond the limits of the Heisenberg cell. Forty eight filter coefficients is the minimum number required without introducing large errors in the transform, inverse transform process. Bi-orthogonal wavelets (Cohen *et al.*, 1992), where the quadrature mirror filters h and g have different lengths, allow perfectly compact wavelets in time with symmetrical properties. **Figure 3-9** shows a bi-orthogonal basis wavelet pair and their corresponding frequency

supports. This bi-orthogonal wavelet basis is called the Pseudocoiflet basis (Daubechies, 1992) as it is the bi-orthogonal basis closest to the orthogonal Coiflet basis wavelets. From this figure we can see that the amplitude of the frequency response is curved across the frequency band where we would prefer it to be flat and that again, the rate of amplitude fall-off in the frequency domain is dependant on the number of coefficients in the corresponding filters.
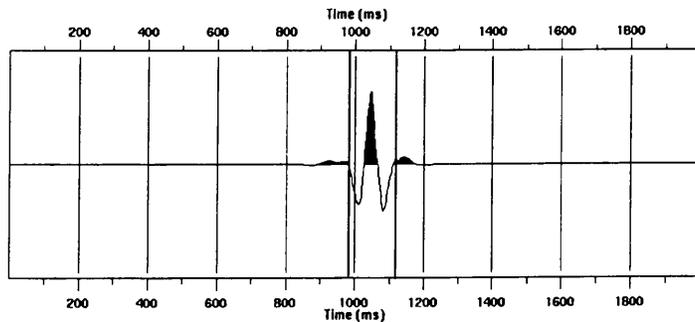
(a)
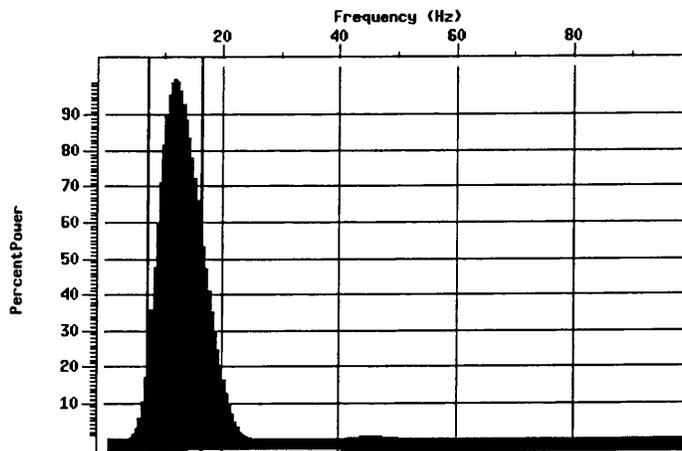
(b)

Figure 3-5 (a) The time and (b) frequency form of the eight coefficient least asymmetric Daubechies wavelet at one scale. The time and frequency limits of the Heisenberg cell associated with this scale are marked on the figure as vertical dark lines.
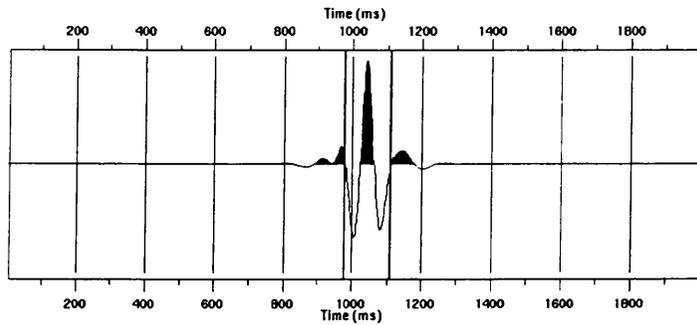
(a)
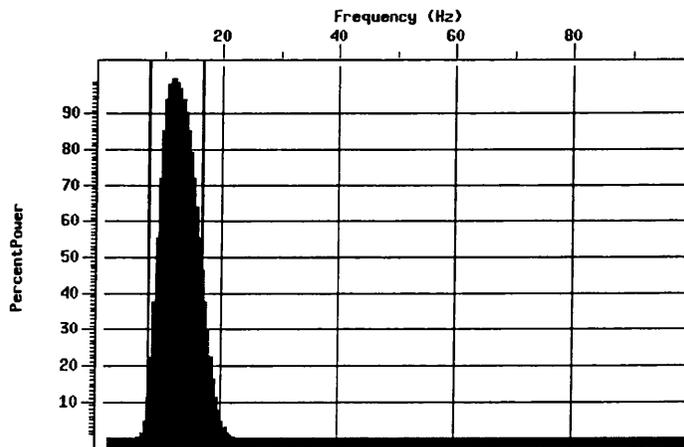
(b)

Figure 3-6 (a) The time and (b) frequency form of the twenty coefficient least asymmetric Daubechies wavelet at one scale. The time and frequency limits of the Heisenberg cell associated with this scale are marked on the figure as vertical dark lines.
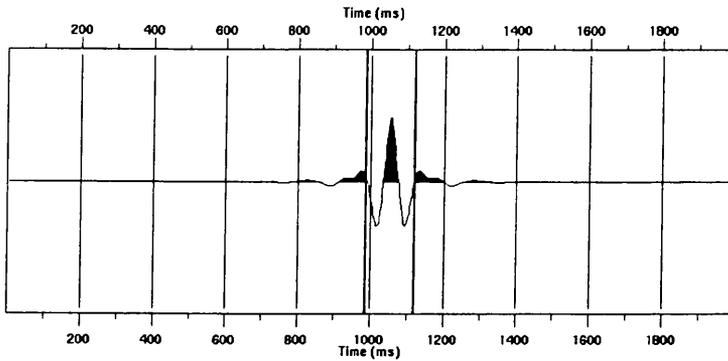
(a)
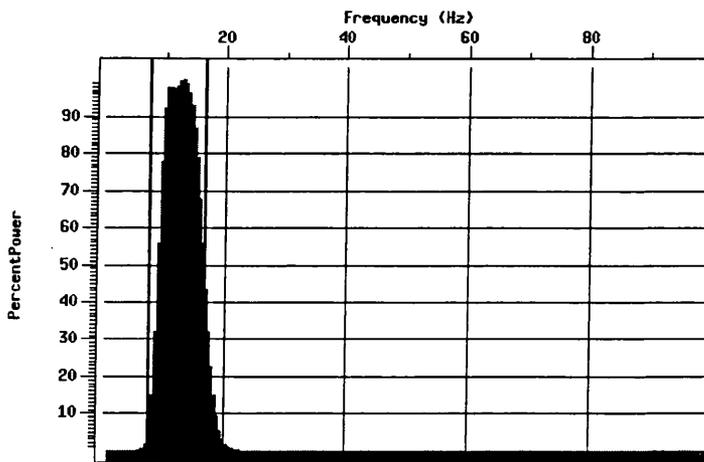
(b)

Figure 3-7 (a) The time and (b) frequency form of the cubic spline Battle-Lemarié wavelet at one scale. The time and frequency limits of the Heisenberg cell associated with this scale are marked on the figure as vertical dark lines.
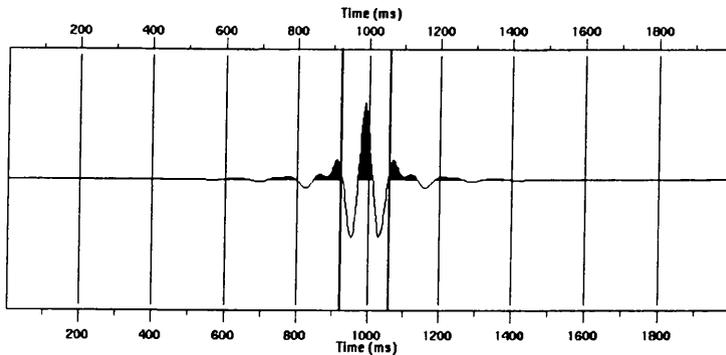
(a)
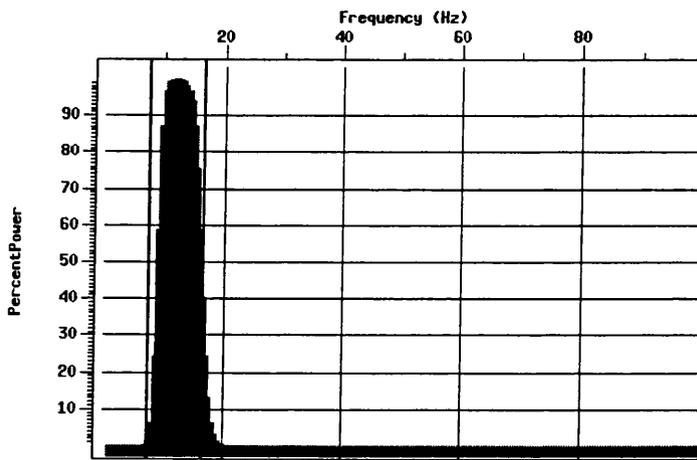
(b)

Figure 3-8 (a) The time and (b) frequency form of the quintic spline Battle-Lemarié wavelet at one scale. The time and frequency limits of the Heisenberg cell associated with this scale are marked on the figure as vertical dark lines.
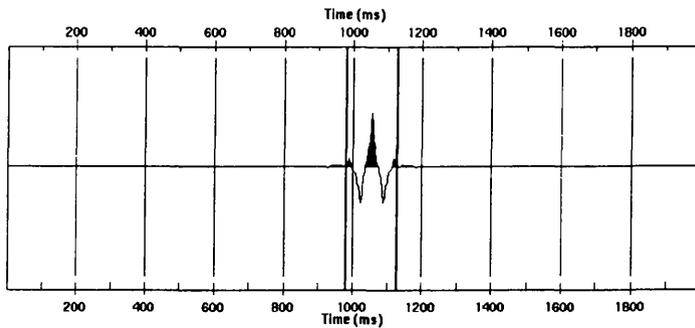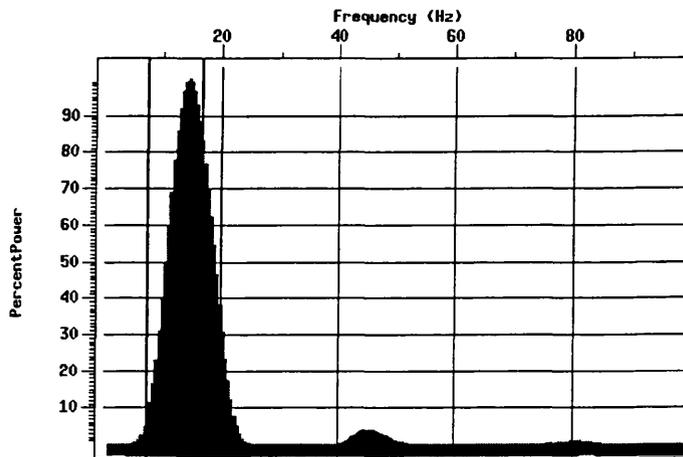
Figure 3-9 (a) The time and (b) frequency form of the Pseudocoiflet bi-orthogonal synthesis wavelet at one scale. The time and frequency limits of the Heisenberg cell associated with this scale are marked on the figure as vertical dark lines.
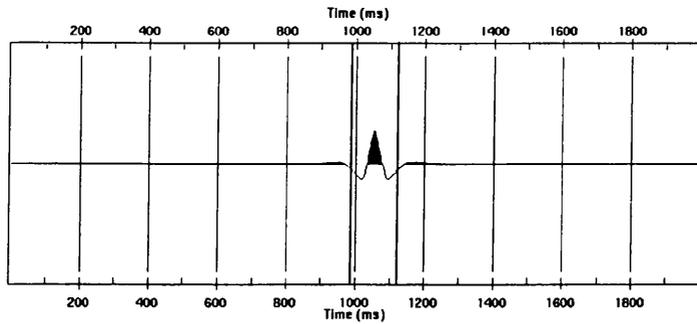
(a)
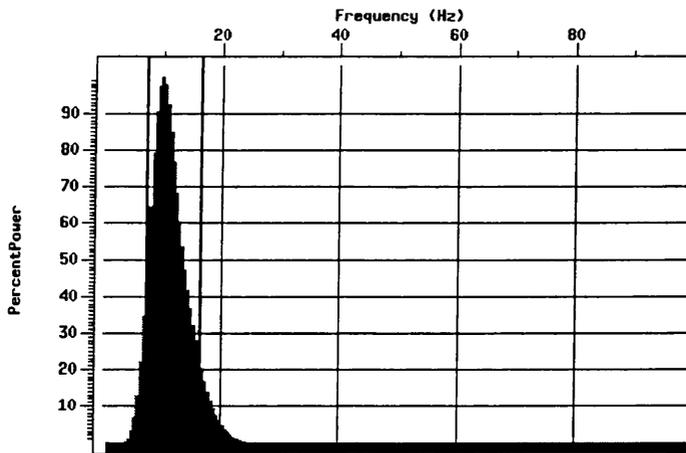
(b)

Figure 3-10 (a) The time and (b) frequency form of the Pseudocoiflet bi-orthogonal analysis wavelet at one scale. The time and frequency limits of the Heisenberg cell associated with this scale are marked on the figure as vertical dark lines.

To investigate the effect of kernel wavelet choice on any filtering process we create a model signal by adding the two signals shown in **Figure 3-11** to create the signal shown in **Figure 3-12j**. The two signals contain band limit wavelets with 35-45-50-60 Hz and 2-4-8-16 Hz corner frequencies. Each signal contains two versions of the same signal one of which is zero phase and one minimum phase. The signals have relatively good separation in frequency domain as shown in **Figure 3-13**. The two zero phase events have the same arrival time on the trace, as do the two minimum phase signals. The purpose of this synthetic signal is to investigate the effect of basis wavelet choice on the performance of any filtering operation in the wavelet domain. The wavelet transform is used to filter the trace by applying the transform and zeroing out the range of scales corresponding to the low frequency signals. All the wavelet coefficients at these scales are zeroed, that is, we are not using the temporal properties of the wavelet transform, making the process similar to bandpass filtering. This lets us evaluate the effect of the frequency support of the basis wavelet on the filtering process.
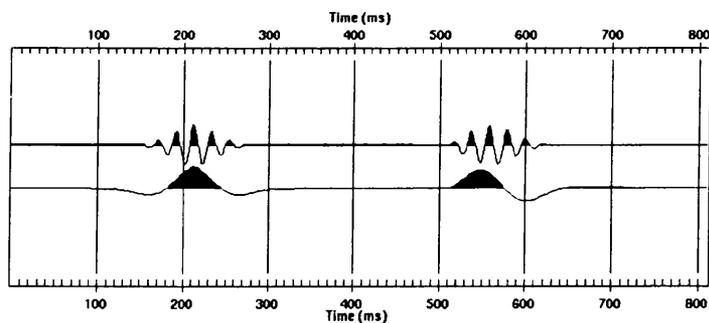
Figure 3-11 Two signals consisting of 35-45-50-60 Hz Ormsby wavelet (upper trace) with zero and minimum phase and a 2-4-8-16 Hz Ormsby wavelet (lower trace) with zero and minimum phase.

Figure 3-14 shows the scalogram representation of the signal using the Battle-Lemarié quintic spline wavelet which, of the basis wavelets shown previously, had the least frequency overlap between scales. From this we can see that the signals are well separated in the wavelet domain apart from a few coefficients in scale 3 separating the main energy areas on the diagram.

The signal has been filtered using the basis wavelets shown in **Figure 3-2** to **Figure 3-10**. The resultant traces are shown in **Figure 3-12** with the ideal output signal. The associated frequency spectra are shown in **Figure 3-16** to **Figure 3-23**. The scales equal to and greater than 3 were removed. From this we can see that each of the kernel wavelets performed well in removing the low frequency signal from the composite signal with all the kernel wavelets apart from the Battle-Lemarié wavelets

introducing varying degrees of phase shifts. From this we can see that the performance of the signal filtering is dependant on the basis wavelet used, with the Battle-Lemarié wavelets producing the best results.

Examining the frequency spectra of the signals after filtering with the ideal output spectra shown in **Figure 3-15** to **Figure 3-23** we can see that the different kernel wavelets lead to different frequency representations. The Daubechies coefficients with fewer coefficients (12 and less) introduce notches into the frequency spectrum which is undesirable and may lead to problems in subsequent processing steps. The 20 coefficient Daubechies wavelets, bi-orthogonal wavelets and the Battle-Lemarié wavelets leave smoother spectra. The suppression of the spectral peak corresponding to the low frequency signal is maximum for the Battle-Lemarié quintic spline wavelet. This filtering performance has been evaluated where we have removed as much of the low frequency signal as possible in the wavelet domain, that is the scales that correspond to the main frequencies of the low frequency signal and the scale (scale 3) where there is a slight overlap in the signals. If, however, we perform the filtering process again, retaining scale 3 we can evaluate the performance of the filtering process when there is residual energy left from the low frequency signal.

Figure 3-24 shows this in the same manner as **Figure 3-12**. From this we can see that the different kernel wavelets produce different results, with the Battle-Lemarié wavelets producing the best results in terms of minimising the residual noise present and any phase shifts introduced. This performance must be due to the frequency supports and symmetrical properties of these wavelets as temporal properties were not used.
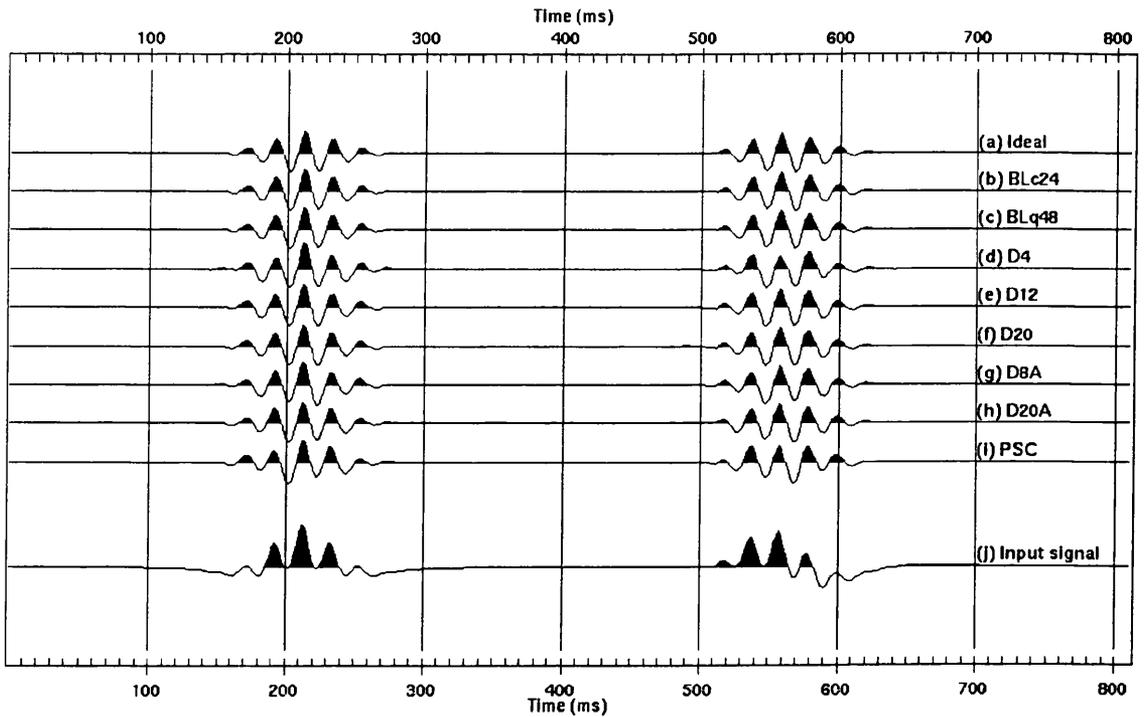
Figure 3-12 Wavelet transform filtering of a model signal to ascertain the influence of choice of basis wavelet on the performance of the filter. Signal (j) is the input signal, and signal (a) is the ideal output. The remaining traces are (j) after filtering with (b) cubic spline Battle-Lemarié (c) quintic spline Battle-Lemarié (d) 4 coefficient Daubechies (e) 12 coefficient Daubechies (f) 20 coefficient Daubechies (g) 8 coefficient least asymmetric Daubechies (h) 20 coefficient least asymmetric Daubechies and (i) bi-orthogonal Pseudocoiflet basis wavelets.
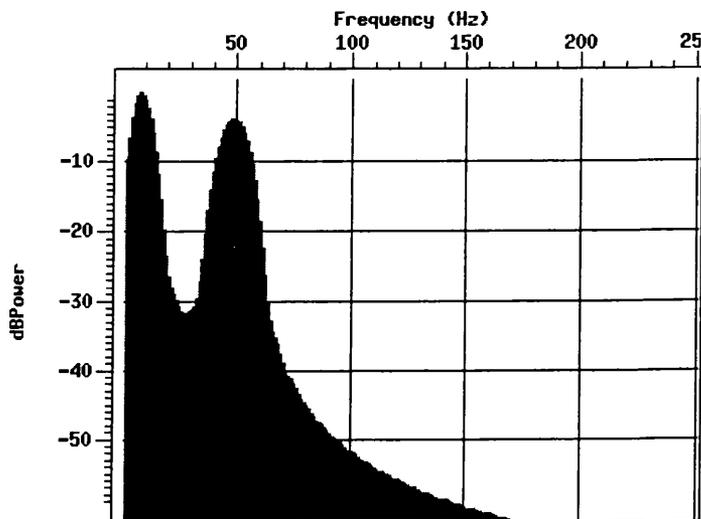
Figure 3-13 Frequency analysis of the input signal.
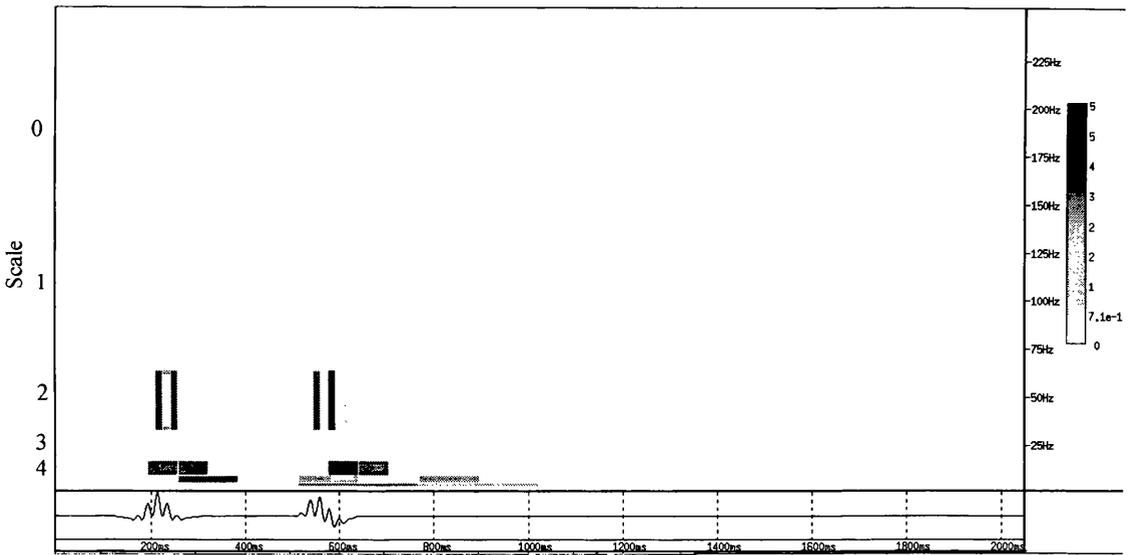
Figure 3-14 Scalogram of the wavelet transform of the synthetic signal shown in Figure 3-12. The Battle-Lemarié kernel wavelet was used in the decomposition.
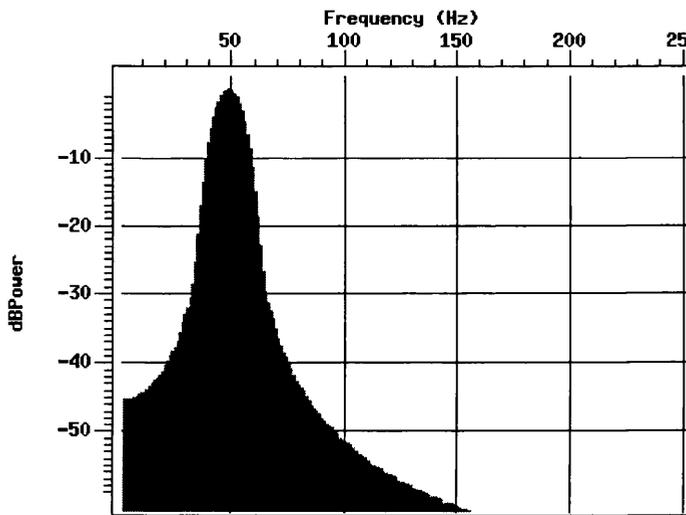
Figure 3-15 Frequency analysis of the ideal filtered signal.
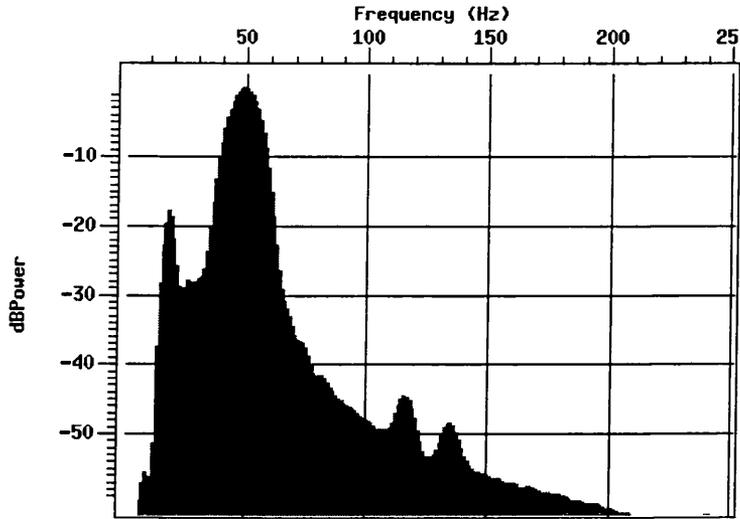
Figure 3-16 Frequency analysis of the signal after wavelet transform filtering using the cubic spline Battle-Lemarié basis wavelet.
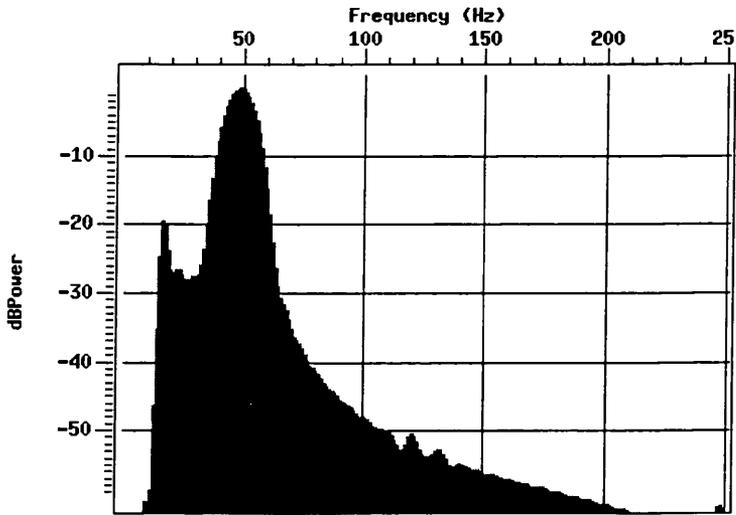
Figure 3-17 Frequency analysis of the signal after wavelet transform filtering using the quintic spline Battle-Lemarié basis wavelet.
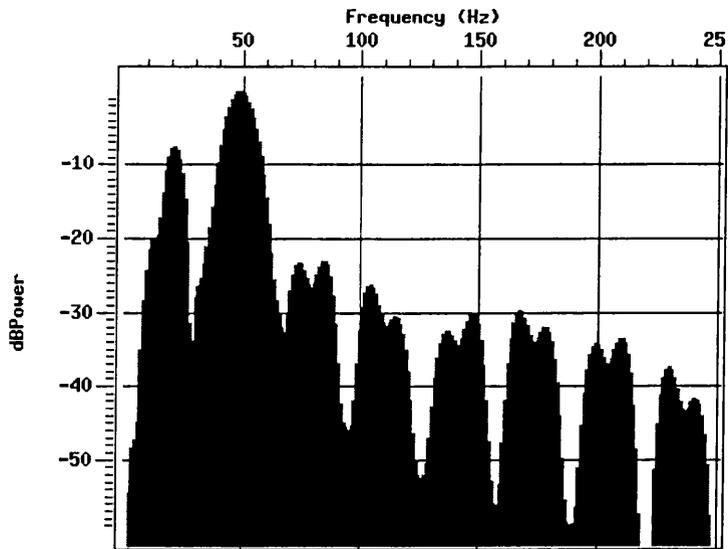
Figure 3-18 Frequency analysis of the signal after wavelet transform filtering using the 4 coefficient Daubechies basis wavelet.
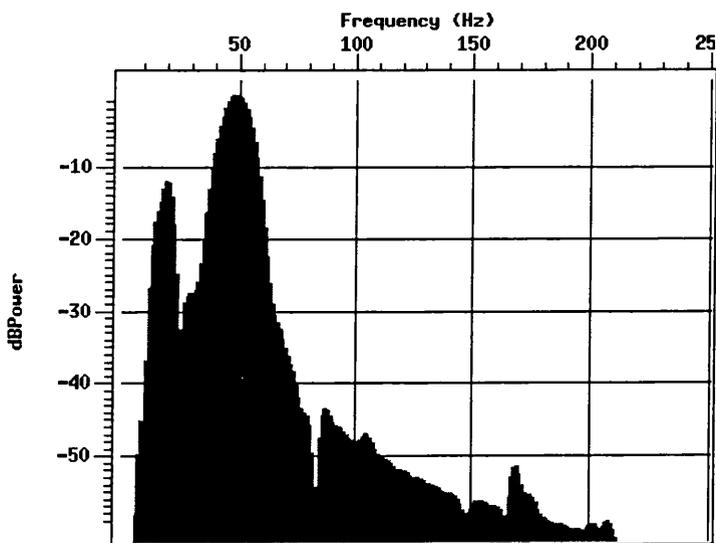
Figure 3-19 Frequency analysis of the signal after wavelet transform filtering using the 12 coefficient Daubechies basis wavelet.
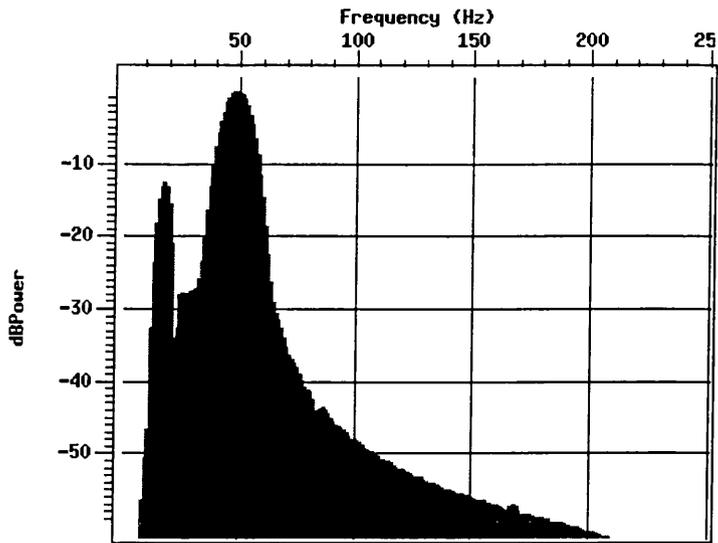
Figure 3-20 Frequency analysis of the signal after wavelet transform filtering using the 20 coefficient Daubechies basis wavelet.
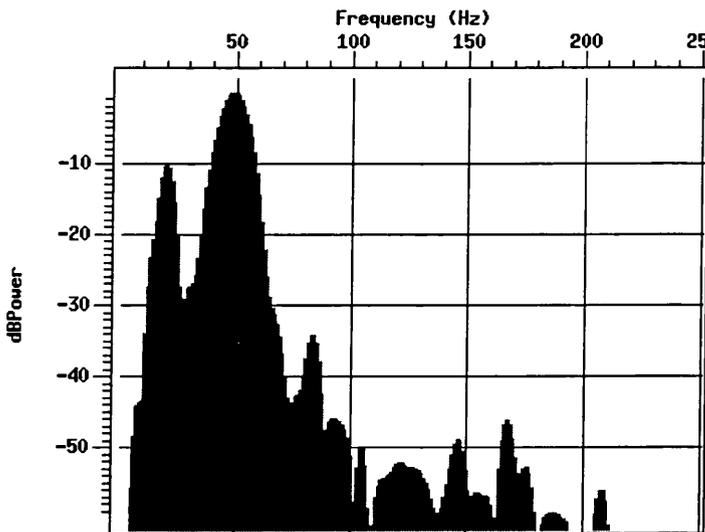
Figure 3-21 Frequency analysis of the signal after wavelet transform filtering using the 8 coefficient least asymmetric Daubechies basis wavelet.
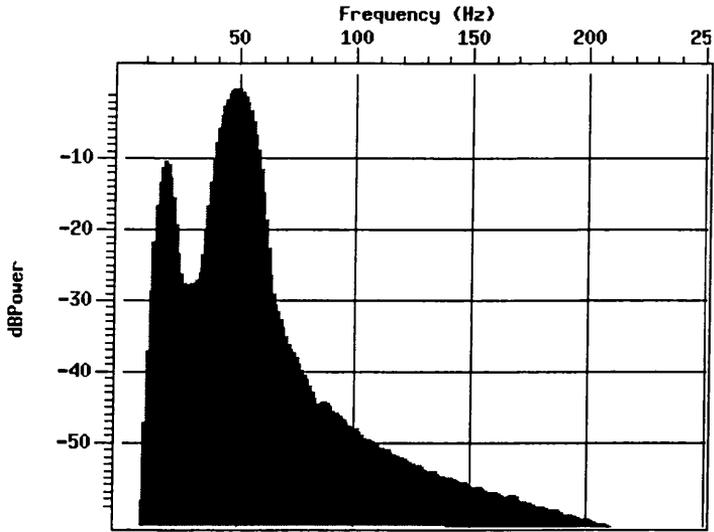
Figure 3-22 Frequency analysis of the signal after wavelet transform filtering using the 20 coefficient least asymmetric Daubechies basis wavelet.
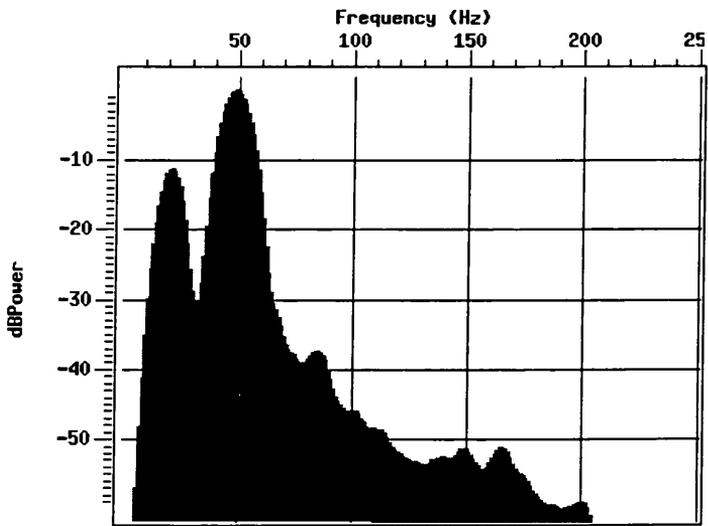
Figure 3-23 Frequency analysis of the signal after wavelet transform filtering using the bi-orthogonal Pseudocoiflet basis wavelet.
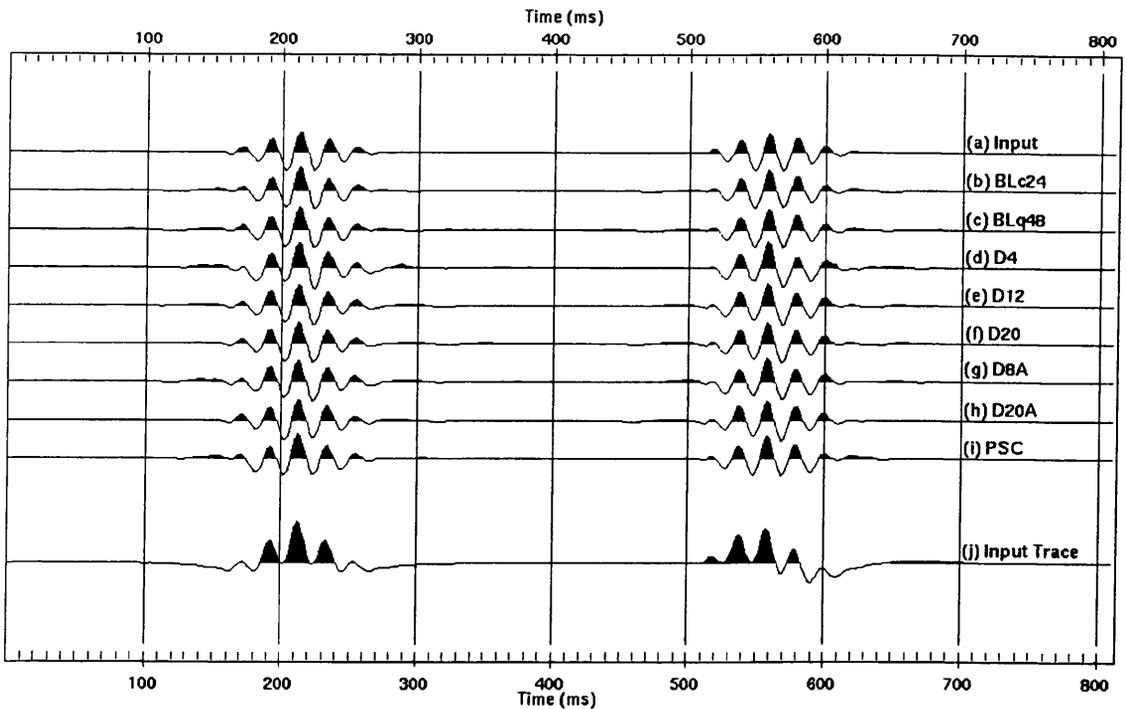
Figure 3-24 Wavelet transform filtering of a model signal to ascertain the influence of choice of basis wavelet on the performance of the filter. One fewer scale has been removed than in Figure 3-12. Signal (j) is the input signal, and signal (a) is the ideal output. The remaining traces are (j) after filtering with (b) cubic spline Battle-Lemarié (c) quintic spline Battle-Lemarié (d) 4 coefficient Daubechies (e) 12 coefficient Daubechies (f) 20 coefficient Daubechies (g) 8 coefficient least asymmetric Daubechies (h) 20 coefficient least asymmetric Daubechies and (i) bi-orthogonal Pseudocoiflet basis wavelets.

3.3 Edge Effects

As we discussed in Chapter 2, a requirement of the fast discrete wavelet transform is that the signal length must be an integer power of two. Therefore, when transforming a signal which is not the appropriate length, it must be padded out. Simple padding with zeroes is not acceptable as this may introduce a sharp discontinuity into an otherwise smooth signal. The discontinuity will have a wide bandwidth in frequency space and so will result in wavelet coefficients spanning all scales. This is indicated schematically in **Figure 3-25**. Therefore, to circumvent discontinuities, some form of edge handling process is required. This can be implemented in two ways: cosine tapering and mirroring of the signal. Cosine tapering involves tapering the edge of the trace gradually to zero by multiplying the edges of the signal by a cosine taper function of a specified length. Mirroring the data involves mirroring the samples about the last sample so that sample (n+1) is the same as sample (n-1), where n is number of samples in the original unpadded signal. The mirroring

process is implemented until the array is an integer power of two in length. These processes are also indicated schematically in **Figure 3-25**.
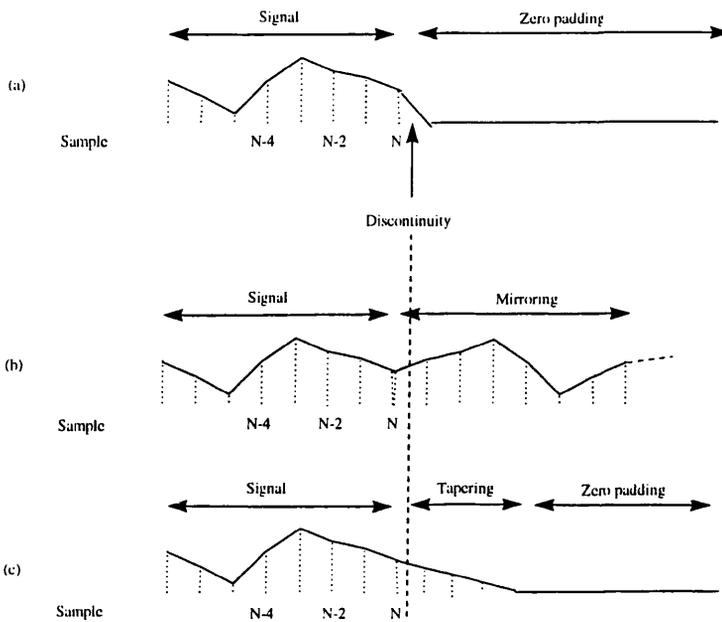
Figure 3-25 Padding of a signal to the next integer power of two number of samples by (a) simply filling with zeroes, (b) mirroring the data about the last sample and (c) tapering the signal with a cosine taper gradually taking the signal to zero plus filling with zeroes.

In addition to signal padding, the transform requires the definition of a edge handling technique. The discrete transform is a series of inner products of the signal and scaled/translated basis wavelets. When the translation parameter approaches the end of the recorded signal there must be a technique for dealing with the multiplications required in inner products which overrun the end of the signal. At the edges of the signal the transform can be designed to treat the signal as periodic; that is to the transform the end of the data appears as:

$$[\quad \ldots \quad f(n-3) \quad f(n-2) \quad f(n-1) \quad f(n) \quad f(0) \quad f(1) \quad \ldots\ldots \quad],$$

where $f(x)$ is the value of the signal at sample x, and the signal has n samples. Again, however, this may lead to the introduction of abrupt discontinuities in signals where there is abrupt amplitude differences between the first and last sample, as can be the case with seismic data. This would lead to the introduction of edge effects in the reconstructed signal, after any subsequent filtering. To minimise edge effects due to discontinuities between the beginning and end of a signal the data at the ends of the signal can be treated so that there is no discontinuity between the beginning and end of

the trace, that is, ramp the values at the end of the trace so that they reach the value of the beginning of the trace, so eliminating any discontinuity. This is indicated schematically in **Figure 3-26.**

When filtering seismic data in time, the edge effects may not be considered to be a major problem as for seismic data, the edges of the traces may be muted, or not of importance. The edge effect will be scale variant in that high scale edge effects will contaminate long time periods and low scale edge effects will contaminate smaller time periods corresponding to the temporal support of the basis wavelets at a given scale. This edge-effect problem and the data length limitation can be eliminated using a more advanced variation of the algorithm such as the one contained in Taswell and McGill (1994).
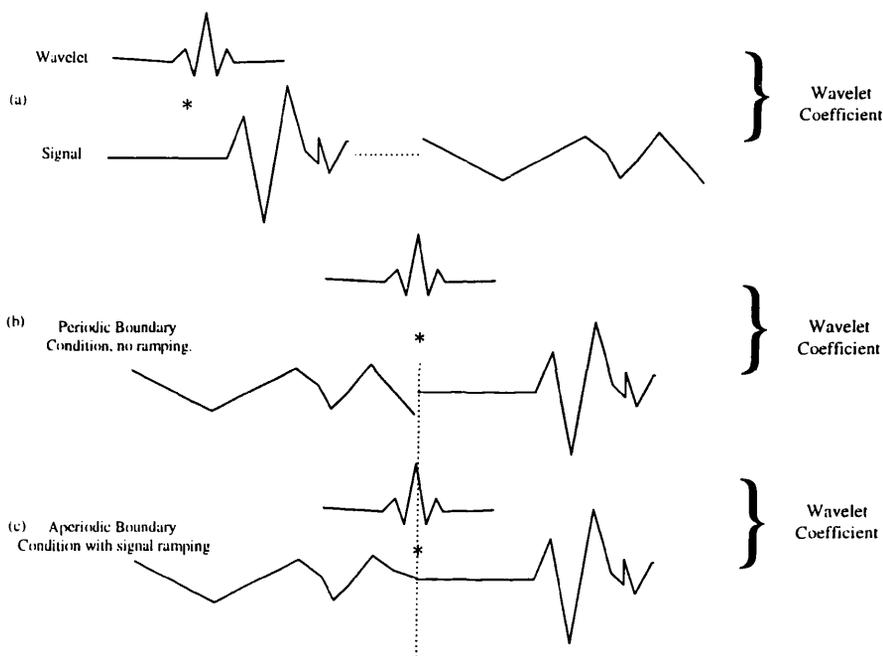
Figure 3-26 (a) A schematic of the wavelet transform showing the inner product (*) of the basis wavelet at a single scale with a finite length signal. When the inner product reaches the end of the signal the boundary condition is set to treat the signal as (b) periodic which due to large amplitude differences can lead to edge effects during any subsequent processing. (c) These effects can be minimised by treating the edges so that discontinuities are avoided.

3.4 Time-Scale Representation

As we have seen previously, the one dimensional discrete wavelet transform, transforms a signal into a two dimensional space providing a mechanism for analysing data at different translations and scales. Higher scales have a more compact basis

function in time allowing analysis with small translation steps, while lower scales have longer basis functions requiring larger translation steps. This leads to the scalogram representation (**Figure 2.6**) of the discrete wavelet where the cell boundaries do not fully represent the support of the basis functions in time and but they give an indication of the time-frequency extent of the corresponding basis function. This representation is dynamic in that it is the same for each kernel wavelet, unlike the temporal and frequency support. **Figure 3-27** shows the decomposition of a 10-50 Hz chirp signal over 1024 samples (2 ms sampling interval) using the kernel wavelet shown in **Figure 3-7**. The grey scale indicates the square of amplitude, in decibels of each constituent wavelet coefficient at that scale and translation. Note the localisation of the decomposition with the high scale (corresponding to low frequencies) dominating the early translations of the decomposition and the lowest scale (corresponding to the highest frequencies) at the later times.

As mention in Section 2.4.5 a property of the discrete wavelet transform is that it is translation variant, meaning that if the input signal is translated by one sample the entire transform changes. This is demonstrated in **Figure 3-28** which shows the wavelet transform for the signal in **Figure 3-27** after translation by one sample and in **Figure 3-29** which shows the difference between the two transforms. From the amplitudes on the difference scalogram and the reconstructed trace resulting from applying the inverse transform to the difference coefficients, we can see that the translational variance affects higher frequencies to a greater degree.
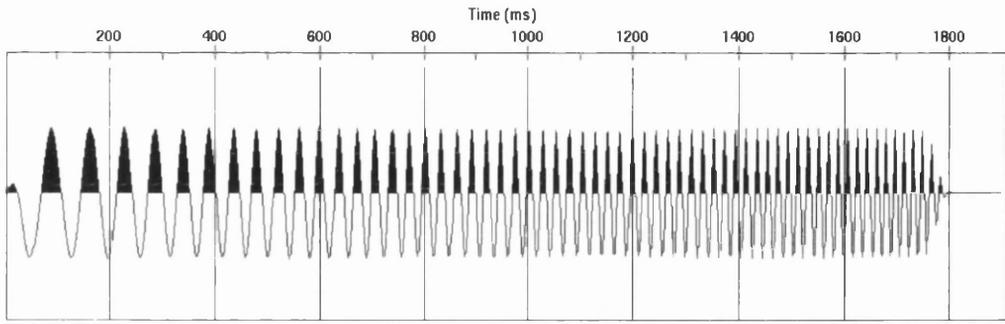
(a)
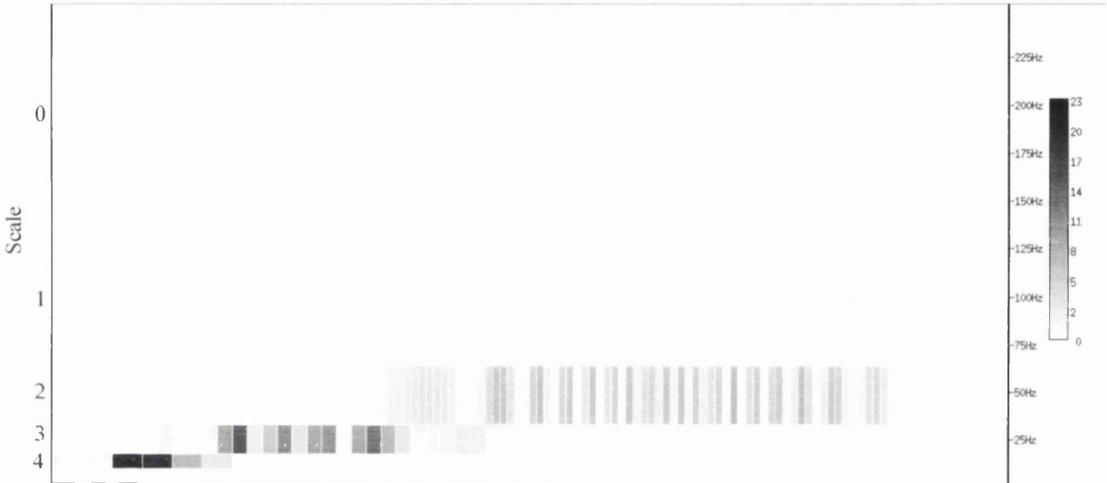
(b)

Figure 3-27 (a) 10-50 Hz linear sweep and (b) the corresponding scalogram representation after wavelet transform using a cubic spline Battle-Lemarié basis wavelet.

Figure 3-28 Scalogram representation of the signal shown in Figure 3-27(a) after shifting the signal by one sample.

Figure 3-29 The difference between the two scalogram representations showing the translational variance of this form of the discrete wavelet transform.

3.5 Time-Frequency Resolution Considerations

From the scalogram representations shown in the previous section we can see that the wavelet transform has good time resolution at lower scales and good frequency resolution at higher scales. Therefore, when using the wavelet transform to filter a signal, considerations must be made for these resolutions and the type of noise that has to be filtered from a signal. The poor time resolution at higher scales leads to a coarse time sampling which would make the filtering process similar to Fourier techniques if only the highest scales were modified. This resolution increases as the scales decrease, with a corresponding decrease in frequency resolution. When filtering coherent noise from seismic records, we usually suppress a frequency range, which would mean that the wavelet transform would suffice. However, we must ensure that the coherent noise frequency range and the wavelet transform scale frequency ranges correspond, thus minimising any signal that may be lost during the filtering process. This principle is demonstrated in **Figure 3-30** to **Figure 3-36**, which shows the filtering of two forms of coherent noise from a signal in the wavelet domain.

Figure 3-30 shows the construction of a synthetic signal (c) from a low frequency signal (a) and a higher frequency signal (b). The wavelet transforms of these signals are shown in **Figure 3-31** to **Figure 3-33** from which we can see that there is signal overlap in the wavelet domain. Removing the scales containing frequencies above and including 15 Hz (**Figure 3-30(d)**) and scales containing frequencies above and including 25Hz (**Figure 3-30(e)**) in an attempt to suppress the higher frequency components leads to heavy distortion of the original signal. Alternatively if we filter a

signal (**Figure 3-34(c)**) to remove signal which is more localised in the wavelet domain we should get less distortion of the remaining signal, when there is signal overlap. This is demonstrated in **Figure 3-34** in which a signal is built from the two signals **Figure 3-34(a)** and **Figure 3-34(b)**, with their corresponding scalograms shown in **Figure 3-35** to **Figure 3-37**. The noise which is well localised in the wavelet domain is efficiently suppressed minimising the loss of signal information.

The coarse time sampling of the discrete wavelet transform at lower frequencies is another property we must consider. This does not seem to be a problem at first, as low frequency signals have large temporal widths and high frequency signals low temporal widths. However, if the low frequency signal has a temporal position between the sampling times of two coefficients then we have to filter a larger portion of the time axis than is otherwise necessary. This limitation is built into this form of the transform and cannot be overcome without resorting to another form of transform. This will be discussed more when we consider the wavelet packet transform as a filtering tool.
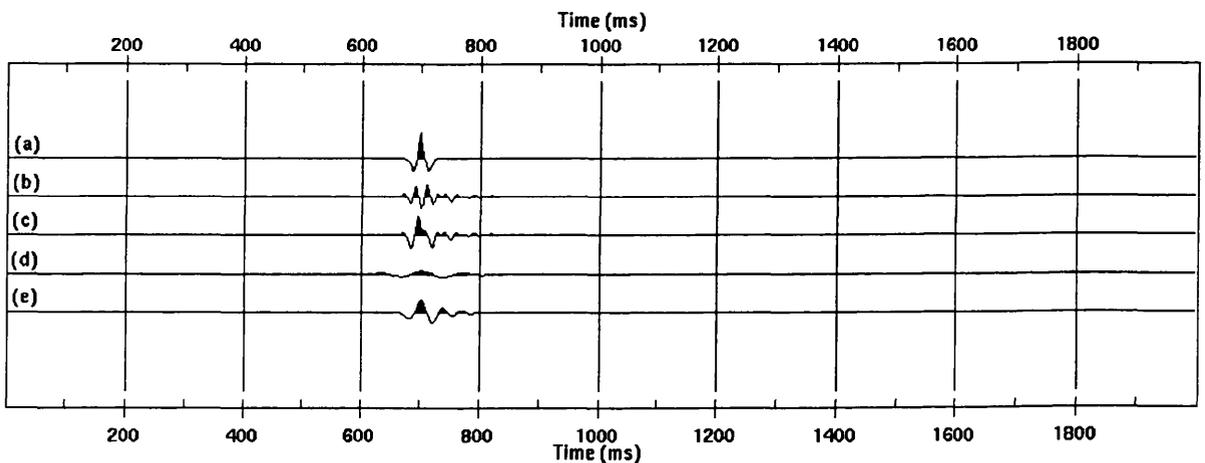
Figure 3-30 Two signals (a) and (b) superposed to form signal (c). (d) shows the signal (c) after filtering with a wavelet transform using a cubic spline Battle-Lemarié basis wavelet. Scales centred above and including 15 Hz were muted. (e) shows the signal after filtering, muting scales over and including 25 Hz.
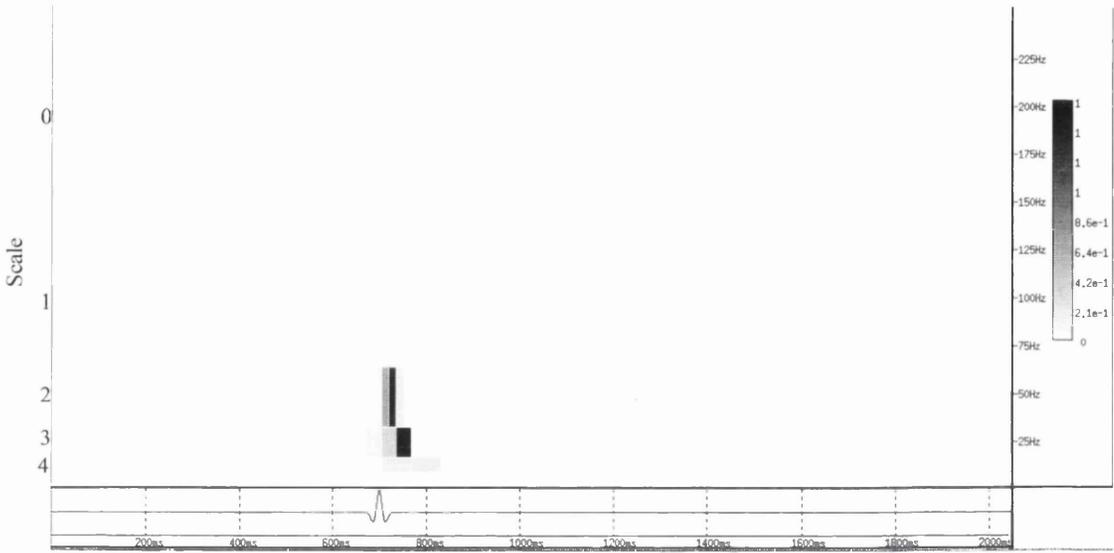
Figure 3-31 Scalogram representation of the wavelet transform of the signal in **Figure 3-30(a)** *showing the span of the signal in the wavelet domain.*
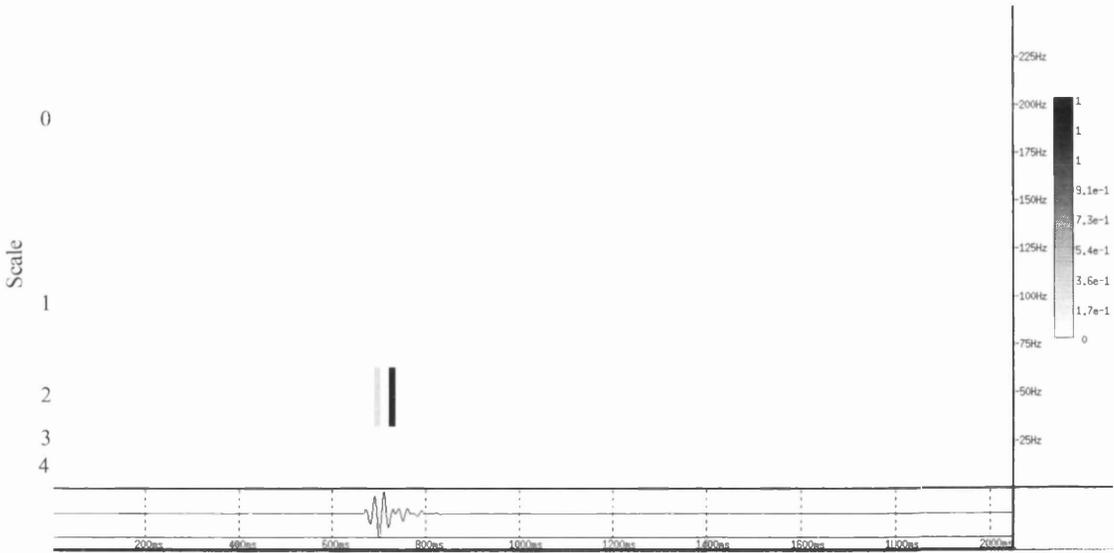
Figure 3-32 Scalogram representation of the wavelet transform of the signal in Figure 3-30(b) *showing the span of the signal in the wavelet domain.*

Figure 3-33 Scalogram representation of the wavelet transform of the signal in Figure 3-30(c) showing the span of the signal in the wavelet domain.
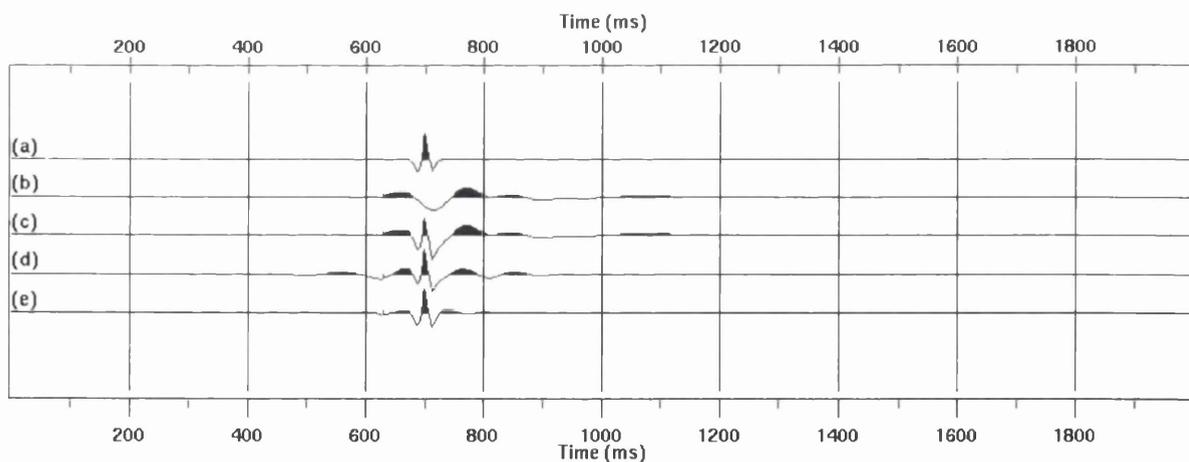
Figure 3-34 Two signals (a) and (b) superposed to form signal (c). (d) shows the signal (c) after filtering with a wavelet transform using a cubic spline Battle-Lemarié basis wavelet. Scales centred below and including 8 Hz were muted. (e) shows the signal after filtering, muting scales below and including 16 Hz.
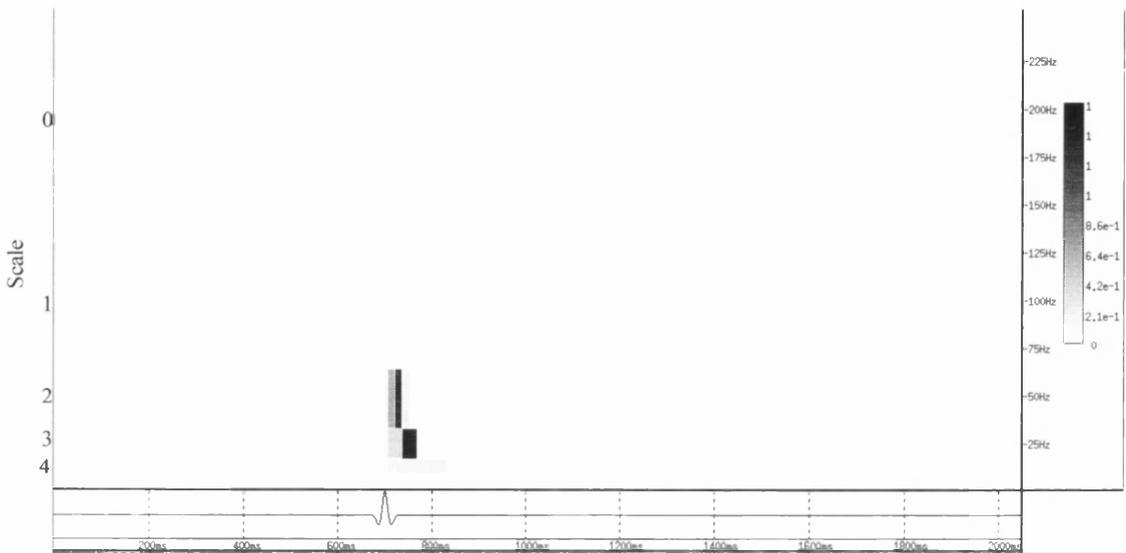
Figure 3-35 Scalogram representation of the wavelet transform of the signal in Figure 3-34(a) showing the span of the signal in the wavelet domain.
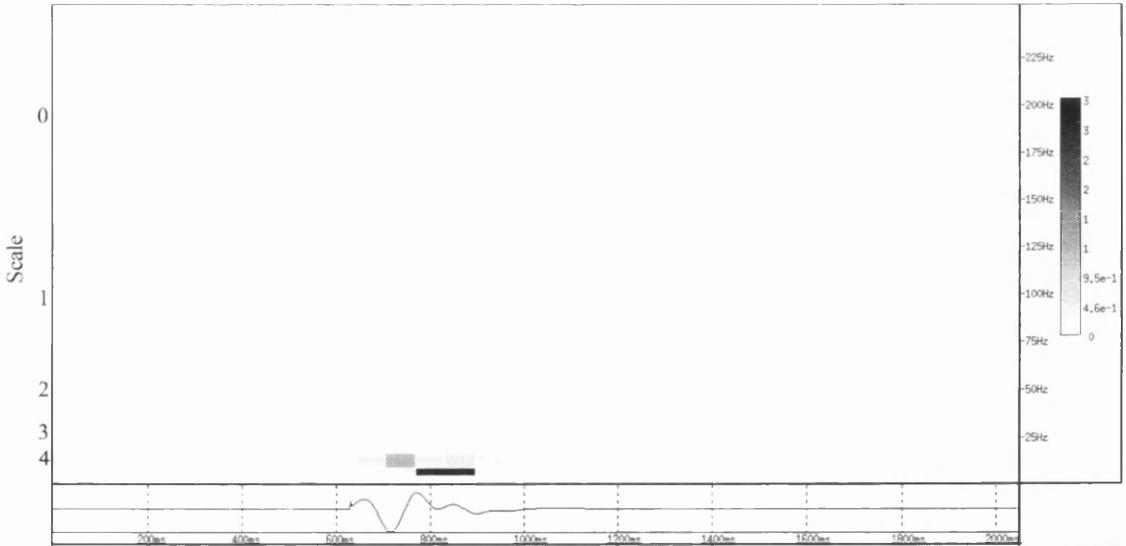
Figure 3-36 Scalogram representation of the wavelet transform of the signal in Figure 3-34(b) showing the span of the signal in the wavelet domain.
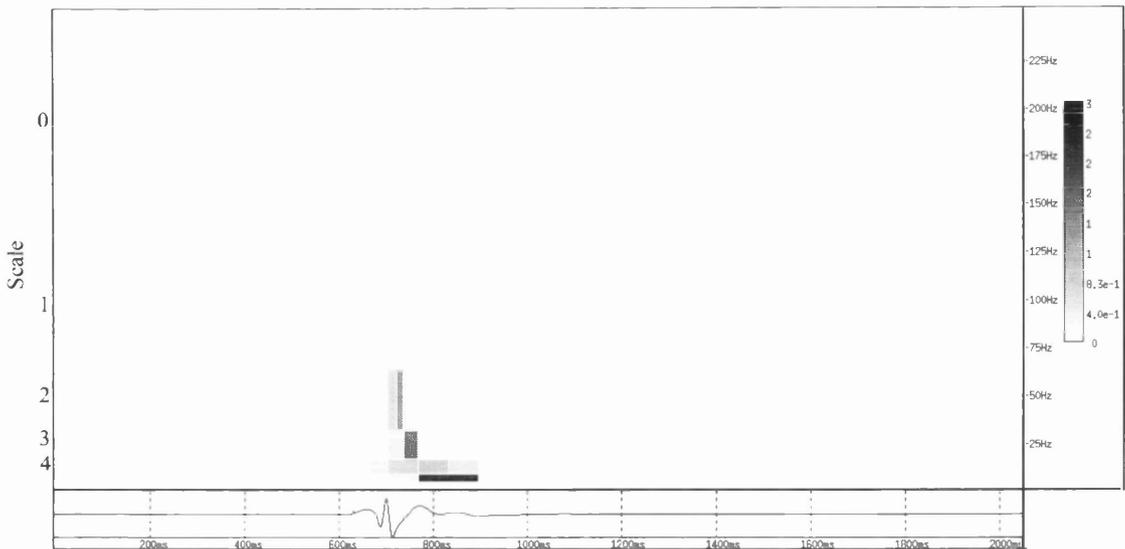
Figure 3-37 Scalogram representation of the wavelet transform of the signal in Figure 3-34(c) showing the span of the signal in the wavelet domain.

3.6 Ground Roll Suppression

Surface waves are a common source of unwanted noise in the seismic record. Seismic sources generate various types of surface waves depending on the near-surface environment and nature and position of the source (Dobrin, 1957; Al-Husseini *et al.*, 1981). Ground-roll is a surface wave whose vertical component is composed of dispersive Rayleigh waves whose different frequency components travel at different velocities leading to long complex wave trains that change as the length of the path travelled increases (Beresford-Smith and Rango, 1988). This can dominate near-source traces on seismic records. Surface waves are such a problem in land seismic acquisition that the design of acquisition parameters are dominated by the need to suppress them.

Current processing methods of eliminating such surface waves from seismic records include frequency filtering, windowed frequency filtering, radon and *f-k* filtering. Elimination of the ground-roll through frequency filtering results in the filtering of the entire trace in the time-domain when ideally, to preserve the frequency content of reflectors, only the affected areas of a trace need to be filtered. To overcome this, data can be processed using time windowing and the short-time Fourier transform (Nawab and Quatieri, 1988). In Chapter 2, it was observed that any windowed transform processing of traces requires careful window design, and should ideally be tailored for frequency; large time windows for low frequencies and a small window for higher frequencies. When a range of frequency values are to be processed, a compromise is required and the so performance of the filter is therefore frequency

dependant.

Two dimensional techniques, such as *f-k* filtering, lead to signal distortion and spatial correlation of background noise producing seismic sections of "wormy" appearance, especially in gathers with low fold. The design of optimal two dimensional filters is non-trivial and their implementation is computationally expensive (Oppenheim and Schafer, 1975). Windowing in the time-offset domain followed by *f-k* filtering has been used as a method to avoid signal deterioration (Beresford-Smith and Rango, 1989) yet this method is still computationally expensive and of limited use for low fold of coverage surveys. Again, as in the one dimensional case, a compromise has to be reached between window design and the filtering parameters. The fundamental element of all these techniques is the Fourier transform, which we have seen uses orthogonal basis functions which have perfect localisation in frequency but infinite extent in time, assuming the signal is stationary (Bracewell, 1986; Chui, 1992).

Faqi *et al.* (1995) introduced the use of the discrete wavelet transform as a form of time-frequency filter to remove locally coherent noise from seismic records. They define locally coherent noise is defined as aliased airwaves, high power-line interference and severely aliased or scattered surface waves. High power-line interference, which appears as a constant single frequency signal superimposed on a seismic trace, can be considered as being almost stationary, as the frequency does not change with time, and it is present on the entire trace. This would indicate that the wavelet transform is not the ideal tool for high power-line interference. It would not have the required resolution in frequency space, and is also a local transform in time whereas this form of noise is global. Airwaves, which consist of air-coupled Rayleigh surface waves are indeed highly spatially aliased due to their low velocity (the speed of sound in air) and are problematic to remove from seismic records as a result. However, as we have seen, the discrete wavelet transform splits the frequency domain into octaves and so will have poor frequency resolution at higher frequencies. Airblast is a source of high frequency noise (Mooney and Kaasa, 1962) on seismic sections and so again, it would seem that the wavelet transform is ill suited to airblast suppression as it would also degrade the data by removing a great deal of reflected signal frequencies as we demonstrated in Section 3.5.

In the paper, Faqi *et al.* (1995) used the most asymmetric Daubechies 20 coefficient wavelet (**Figure 3.6**) as the transform kernel wavelet. This may not be the best choice of wavelet from the criteria determined earlier, but distortions would be localised. The frequency and time resolution of the discrete wavelet transform is optimal for filtering at the lower end of the frequency spectrum and so it has been

investigated as a tool for the localised suppression of ground roll.

3.7 Filtering Technique

The main objective of any form of filtering is to remove noise, in this case associated with surface waves, from a signal whilst minimising the distortion of the signal retained after filtering. Ideally, to filter data without distorting the retained signal, we require perfect separation of retained signal and noise in the transform domain. Typically, seismic reflection signals are wide-band and ground-roll mainly contains low frequencies (Yilmaz, 1987) and so signal separation in the frequency domain is not achieved. Therefore, filtering of any signal based on frequency will lead to the removal of low frequency content of the signal associated with reflections as well as any ground roll. In the case of bandpass filters, based on the Fourier transform, any filtering will remove frequencies outwith the band over the whole trace. The aim of filtering seismic data with a discrete wavelet transform is to localise the filtering process to areas only affected by the ground roll and so minimise the loss of any low frequency signal component from the signal. This method of filtering is a form of adaptive-window frequency filtering in that this will be a windowed process, only filtering the data in specific areas, but unlike windowed Fourier techniques, uses a window function that adapts to the frequency of the signal that is being filtered.

Filtering in the wavelet domain can be performed using several techniques. The simplest technique involves zeroing wavelet coefficients in the time-scale areas corresponding to ground-roll energy in the traces. This would effectively suppress the ground-roll in a time-frequency sense. This filtering operation is applied on a trace by trace basis, with the wavelet coefficients being zeroed in the appropriate areas for each trace. The appropriate coefficients to be filtered can be determined using the effective support of the kernel wavelet and Equations 3.1 and 3.3. No tapering of wavelet coefficients is required as the compact nature of the basis functions provides a natural taper minimising any Gibbs ringing, which can be observed with Fourier based filters. Although the filter may be regarded as two dimensional, in that it zeroes coefficients in the time-scale space, it is primarily a one dimensional filter in the sense that it is a trace by trace operation.

An alternative to simply zeroing the coefficients in the wavelet domain, is to apply a weighting to the coefficients. Perfect separation of signal from noise cannot be achieved using this technique as there is overlap between signal and noise in the time-frequency domain, therefore, to try and preserve the low frequency component of the

signal, a weighting can be applied to the wavelet coefficients to discriminate against the component corresponding to the ground-roll. The main problem with this principle is to determine the weight to apply to the coefficients.

Source frequency attenuation with travel time on seismic records is measured by a quantity called the quality factor Q (Yilmaz, 1987). An infinite Q means that there is no attenuation, and the factor can change in depth and in the lateral direction. If we assume a constant Q (Kjartannson, 1979) we can use this technique to estimate a weight to apply to the wavelet coefficients. The Q-value governs the amplitude change of a frequency component f on a seismic trace between two times t_1 and t_2 according to:

$$A_1(f) = A_2(f)e^{-\frac{\pi f}{Q}\Delta t},\tag{3.3}$$

where $A_1(f)$ is the amplitude of the frequency component at t_1 and $A_2(f)$ is the amplitude at t_2. The exponential term applies the attenuation to the amplitude according to the Q-value, an example of which is shown in **Figure 3-38**.
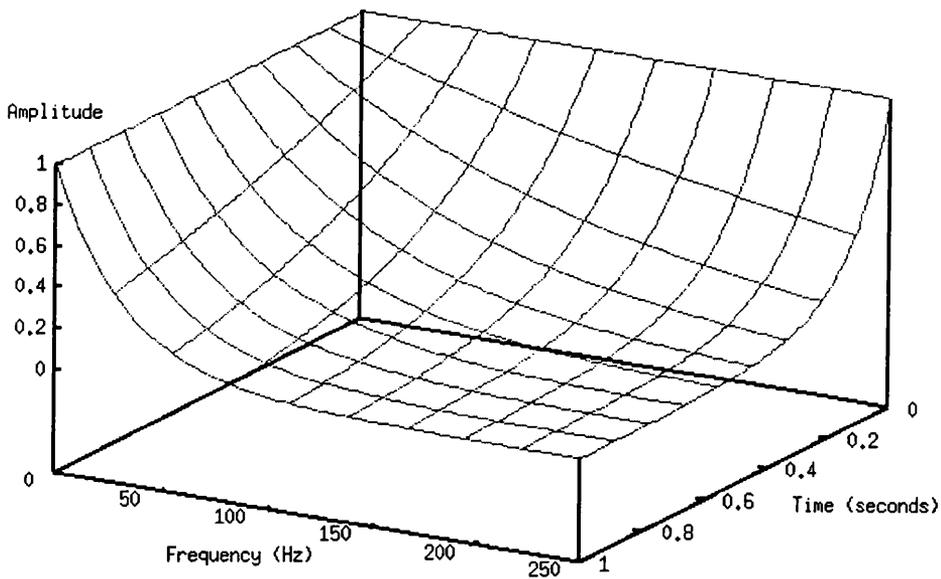
Figure 3-38 The attenuation in frequency-time due as modelled by a Q value of 100. An amplitude equal to one means no attenuation.

Although the Q concept is based on Fourier techniques, if we use a basis wavelet that is well localised in the frequency domain (minimise the energy outside the Heisenberg cell boundaries), the wavelet transform represents a good approximation to a frequency-time distribution. From this we can estimate a Q-value for the seismic

signal and by comparison with the wavelet transform of the signal, estimate the component of the coefficients to be filtered that represent signal. **Figure 3-39** shows a three dimensional view of the modulus of the discrete wavelet transform of a seismic trace containing ground roll and an estimated Q-value attenuation function for the trace. We can use the wavelet coefficients from the zone which is not to be filtered (assumed to represent the signal only), in combination with the Q-value (to extrapolate the value of these coefficients at a lower frequency) to estimate the value of the wavelet coefficients to be filtered which correspond to the reflected signal and not the ground roll. The zones of the transform which are above the Q-envelope/surface are shown in **Figure 3-40**

This technique will not perform well with a kernel wavelet that is not well localised in the frequency domain, such as the 4 coefficient Daubechies wavelet. The purpose of this technique is to provide an alternative to muting the coefficients and so preserve signal bandwidth during the filtering process. The process is by no means the optimum weighting procedure, but gives an indication of whether this is a viable process. We will evaluate the performance of this technique in the next section.
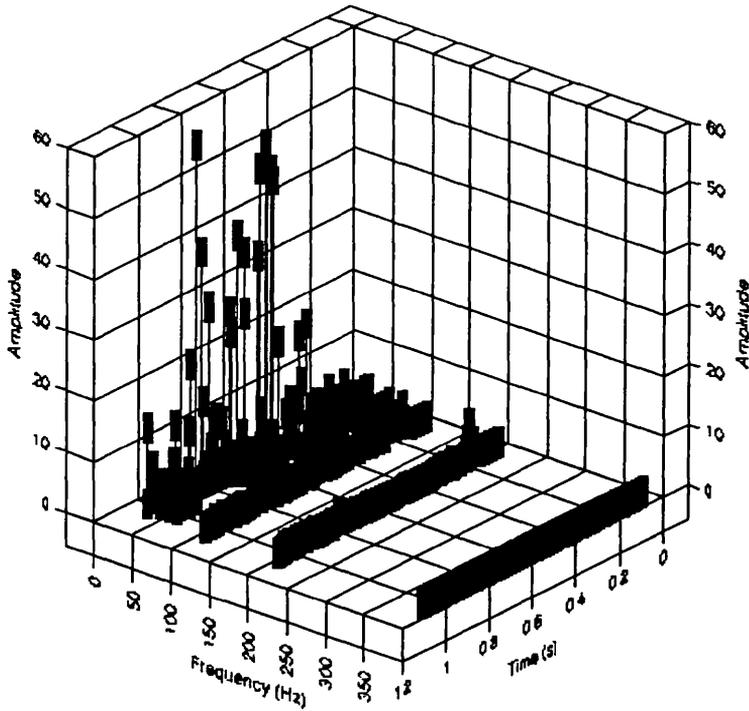
(a)
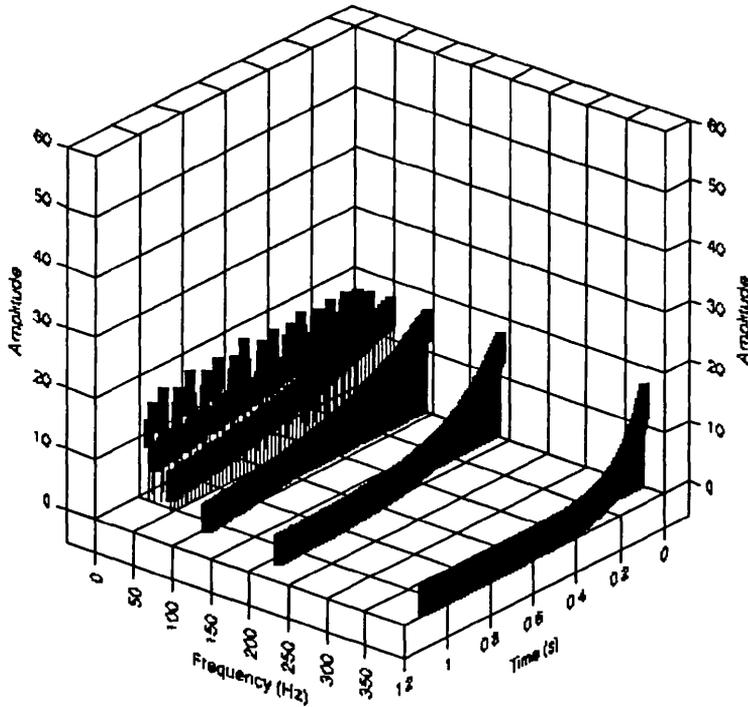
(b)

Figure 3-39 Wavelet transform coefficients of (a) the signal in Figure 3-42 and (b) of the estimated Q value at the same points. Using the weighting filter technique, where the coefficients are above the Q surface, and within the area to be filtered, a weighting will be applied.
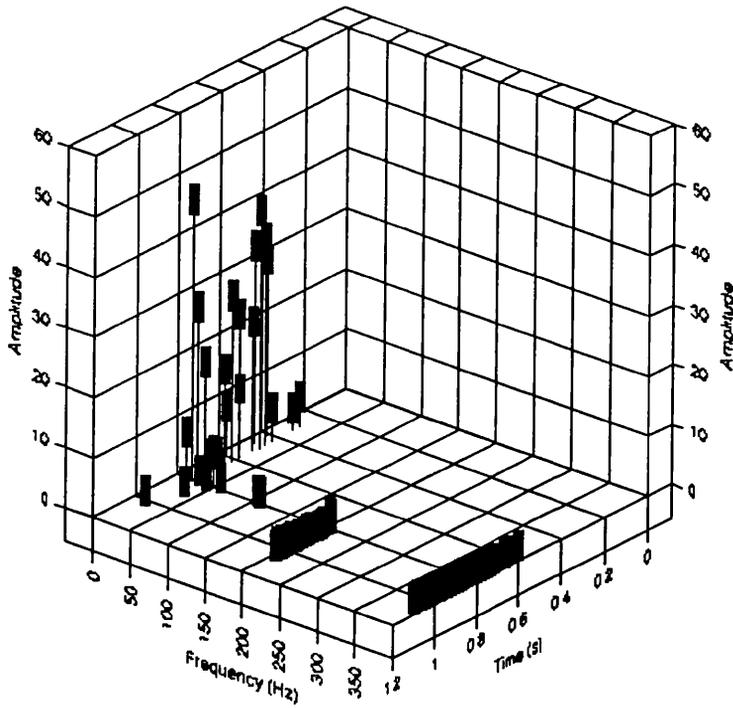
Figure 3-40 The amplitudes of the wavelet coefficients above the Q-surface. This is generated by taking the Q-value away from the absolute value of the wavelet transform coefficient. The majority of the energy is in the low frequency areas, which in this case correspond to ground roll.

3.8 Wavelet Filtering of Land-Based Seismic Data

The wavelet filtering techniques were tested on shot records from land-based seismic data collected from a site near Glasgow. The seismic line was shot to provide a three component data set to test the effectiveness of polarisation filtering methods. Nineteen, eight ounce gelignite sources at a depth of two meters were shot into 16 groups of three, 14 Hz tri-phones with one meter spacing, and thirty meters group spacing. The three component data were collected using a Geosource MDS-10 with a sample interval of 1 ms. No attempt was made to suppress ground-roll in the field acquisition. Attempts to suppress the ground-roll with polarisation filtering were unsuccessful, and this analysis uses only the vertical component data set. A sample shot record is shown in **Figure 3-41** and can be seen to be heavily contaminated by ground-roll.
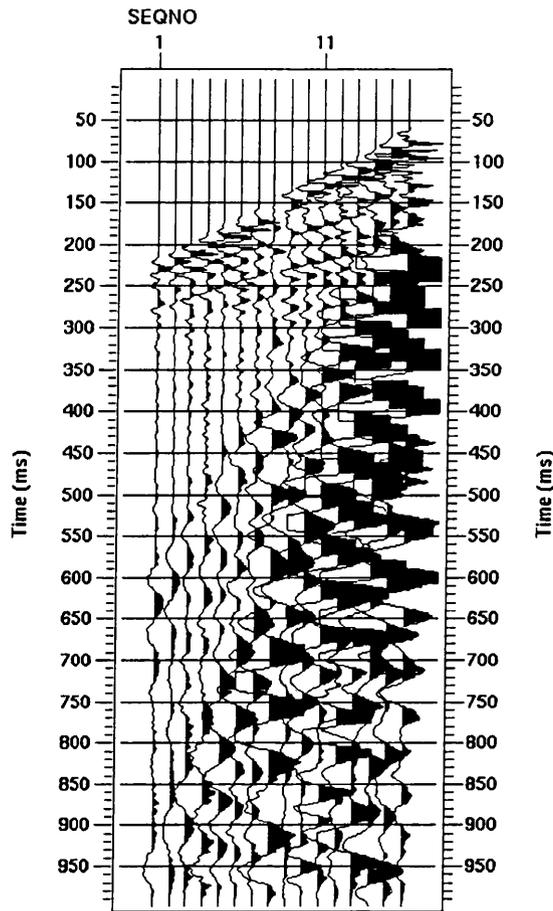
Figure 3-41 A field record from the Robroyston survey showing heavy contamination by ground roll.

Using the time-scale representation, the shot energy can be observed in a domain similar to time-frequency. **Figure 3-42** shows the time-scale representation of trace 12 from the shot record in **Figure 3-41**. Scale zero represents the basis function, containing the highest frequencies, and scale 9 the basis function containing the lowest frequencies. The ground-roll is first discernible on this trace at approximately 300 ms. We can see from this time-scale representation that the ground-roll energy is concentrated in scale 4 and higher scales at times greater than 275 ms.

The filtering of a whole shot-record using the wavelet techniques is illustrated in **Figure 3-43** where it is assumed that for ground-roll the start time for the zeroing of the coefficients increases with offset while the scales to be filtered were kept constant from trace to trace.

The result of applying the muting procedure to all the traces in the shot record is shown in **Figure 3-44**. **Figure 3-45** shows the result of applying the weighting procedure in the wavelet domain to the data. For comparison purposes, a band-pass filtered shot record and a *f-k* filtered record are shown in **Figure 3-47**. The *f-k* filtered

filtered shot record and a *f-k* filtered record are shown in **Figure 3-47**. The *f-k* filtered record contains large amplitude artefacts, particularly after 650 ms , which are not apparent on the bandpass filtered record.
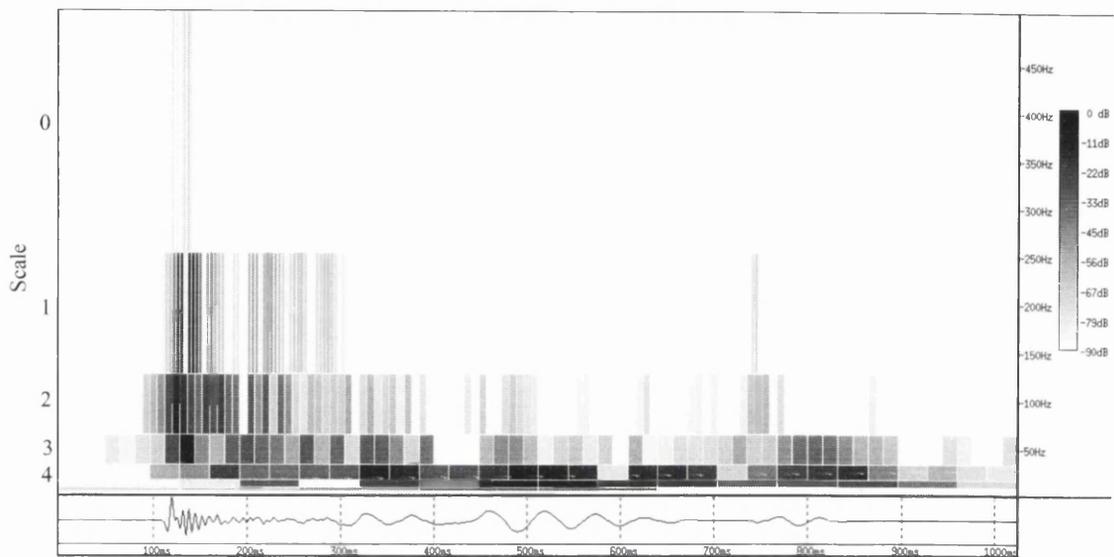
Figure 3-42 Scalogram representation of the wavelet transform of trace 12 of Figure 3-41. The decomposition used the Battle-Lemarié cubic spline kernel wavelet.

Figure 3-46 shows a spectral analysis of the shot record before and after filtering by the wavelet transform technique. The large amplitude at low frequencies due to the ground roll is apparent on **Figure 3-46(a)**. From **Figure 3-46(c)** we can see that the shot record filtered by the weighting technique has a wider bandwidth than the record filtered by the muting technique, due to retention of lower frequency signals by this technique.

The performance of the *f-k* filter is effective in removing the ground roll but produces a shot record of wormy appearance, degrading the quality of the record. This can be attributed to the low number of traces in the record combined with spatial aliasing and is a common problem with the performance of *f-k* filters (Peacock, 1982). The band-pass suppresses the ground-roll effectively but, also removes the corresponding frequencies from reflectors outwith areas affected by the ground-roll and so reduces reflector bandwidth unnecessarily. **Figure 3-44** and **Figure 3-45**, the wavelet transform filtered records, show considerable improvement particularly for the event at ~400ms and produce comparable results with the bandpass filter in the area previously contaminated with ground roll.
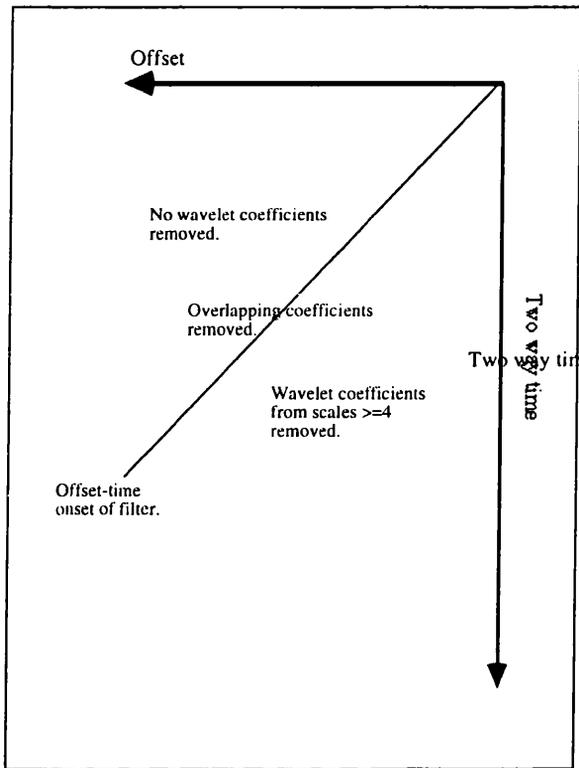
Figure 3-43 Schematic diagram showing technique for filtering shot record using the wavelet transform technique.
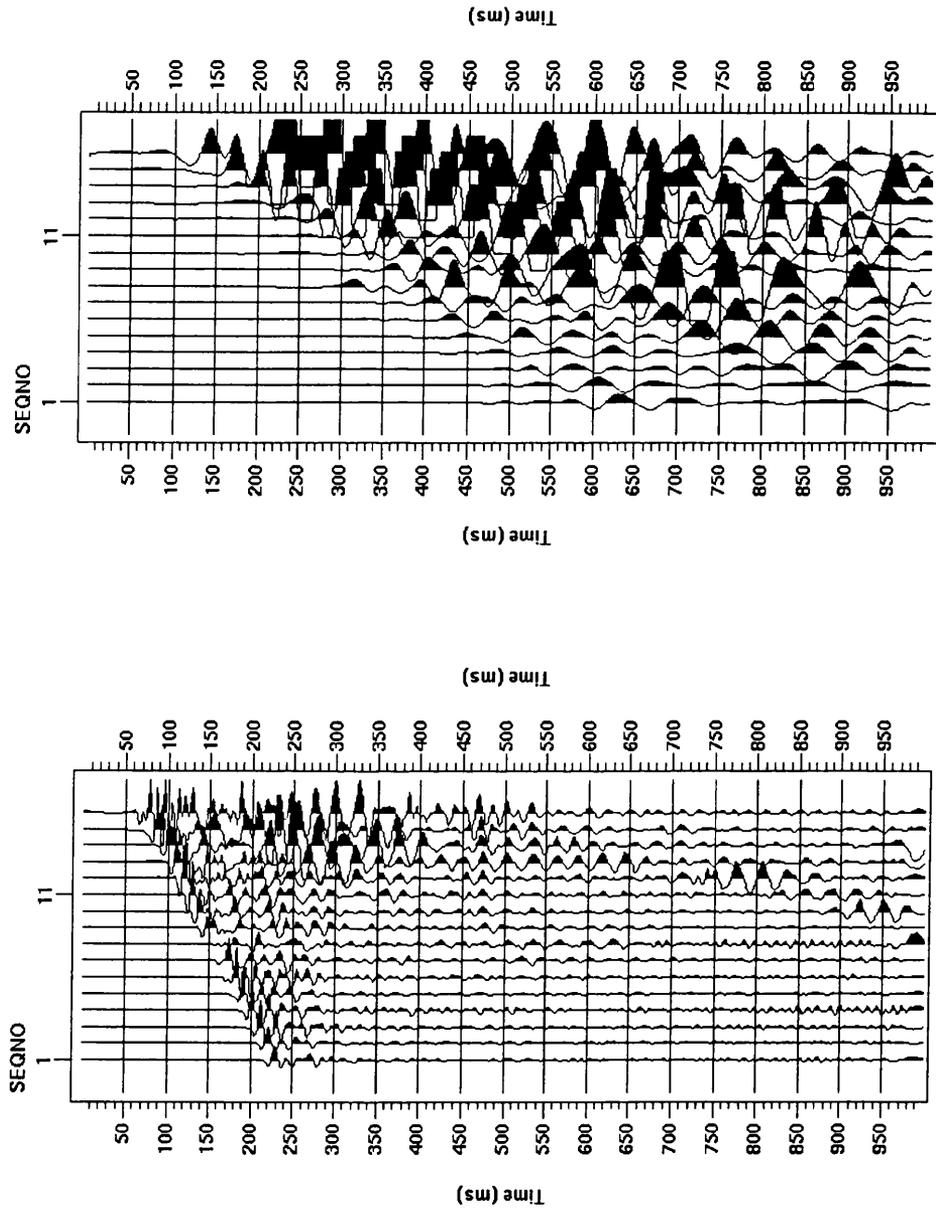
Figure 3-44 (a) Figure 3-41 after filtering using the wavelet transform technique. The coefficients were muted in the wavelet domain. (b) The difference section showing the noise removed from the record.

Chapter 3: Wavelet Transform Filtering
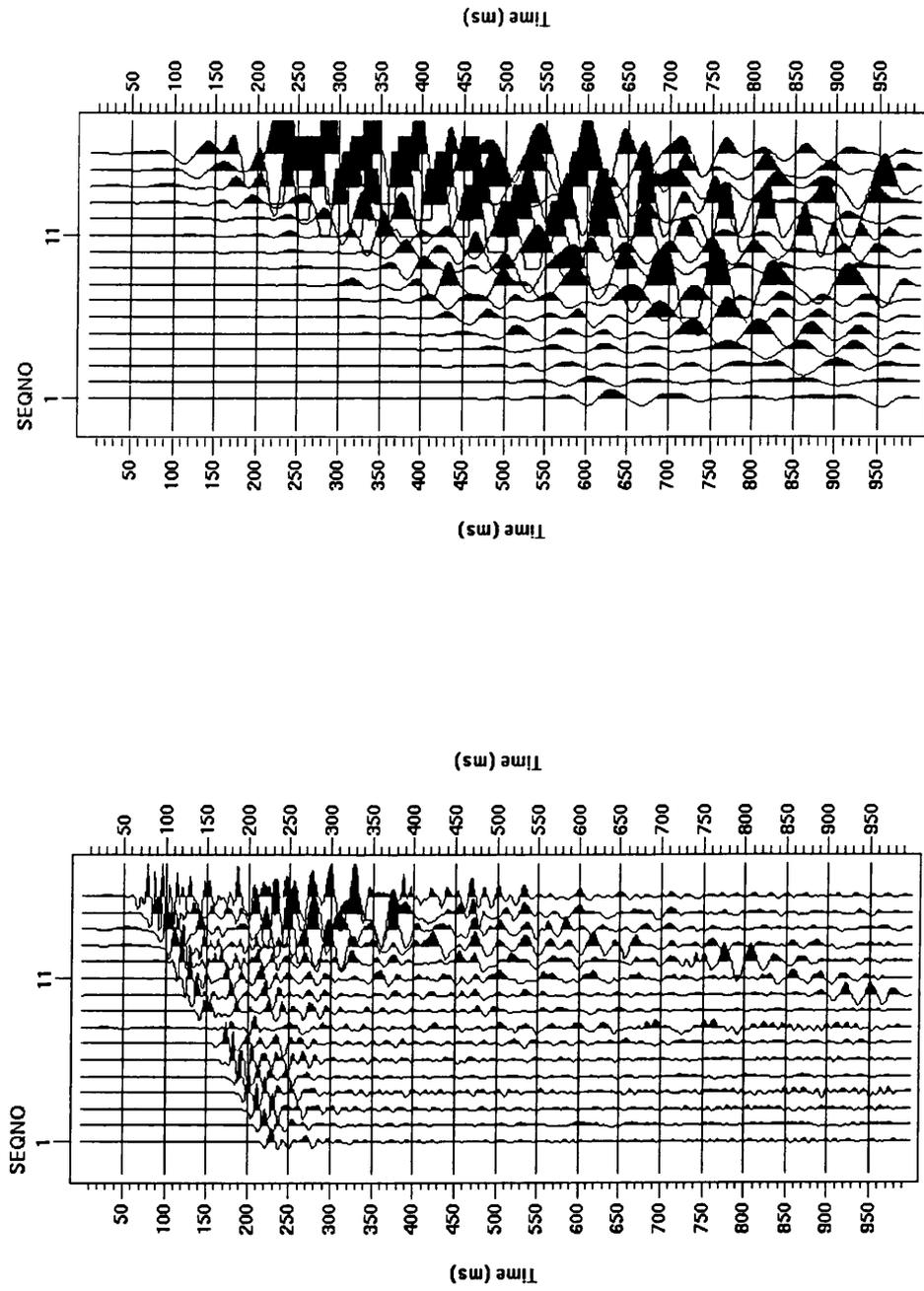
Figure 3-45 (a) Figure 3-41 after filtering using the wavelet transform technique. The coefficients were weighted in the wavelet domain using a constant Q=80. (b) The difference section showing the noise removed from the record.

Chapter 3: Wavelet Transform Filtering
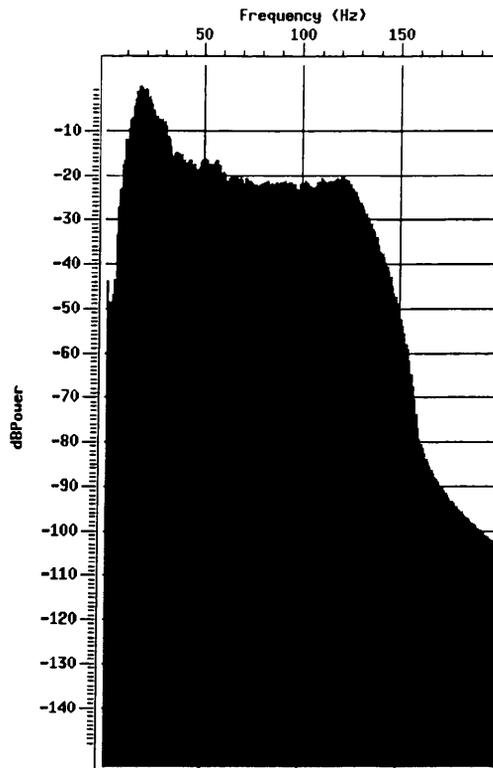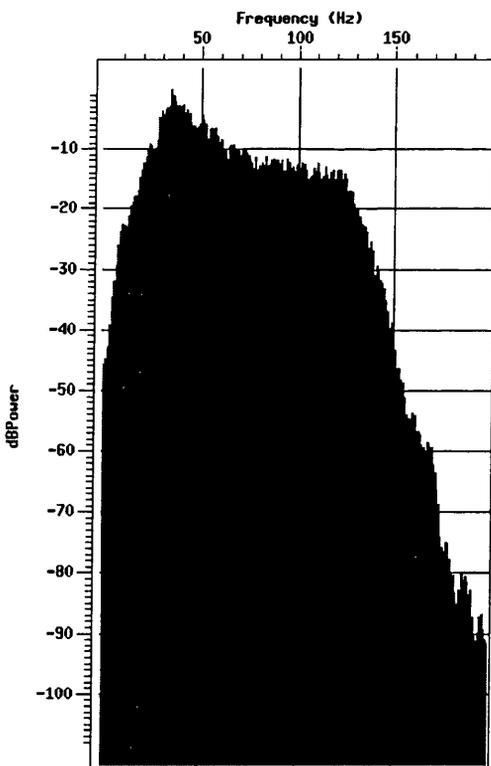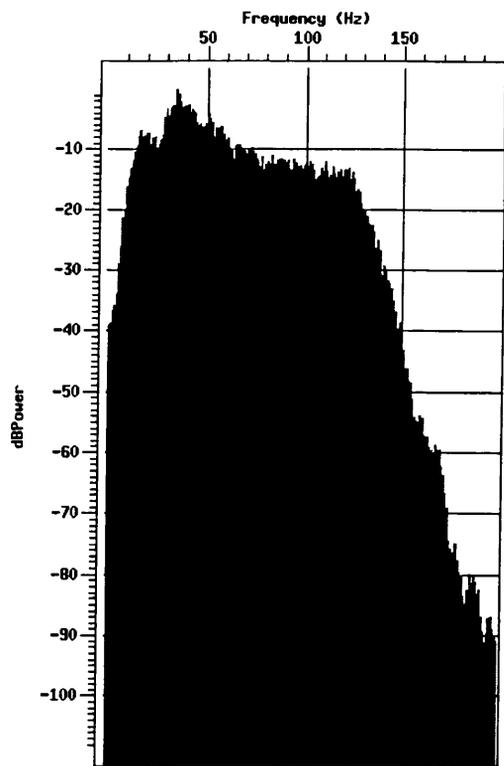
Figure 3-46 Spectral analyses of the common shot records shown in (a) Figure 3-41 *(b)* Figure 3-44 *and (c)* **Figure 3-45**. *The suppression of low frequency ground roll signal is apparent in the wavelet transform filtered spectra. The remaining signal has a wider bandwidth for the weighted filtered spectra shown in (c).*
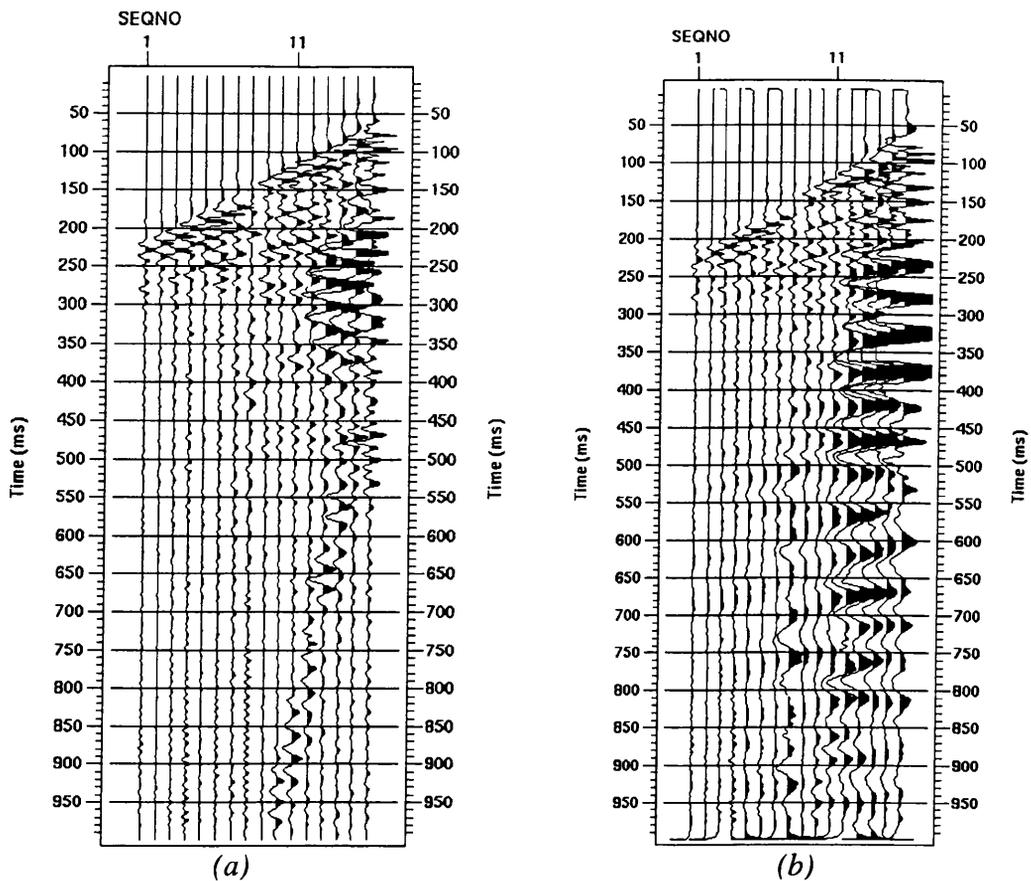
Figure 3-47 (a) Bandpass and (b) f-k *filtered versions of the common shot record shown in Figure 3-41. The bandpass filter works well yet suppresses frequencies from the entire trace as can be seen from the broadening of the first break arrival, and the* f-k *filtered section has a wormy appearance due to low fold and spatial aliasing.*

The next stage in testing the effectiveness of the wavelet transform filter was to see if any improvement is passed onto a brute stack. After the respective wavelet transform/bandpass filter, the processing stream consisted of trace kills, CMP sorting, NMO and then stack. No statics were applied to the data. The maximum fold of coverage was 16, which was reduced after trace killing for the band-pass filtered data set. The brute stacks are shown in **Figure 3-48** to

Figure 3-51 , whose quality are degraded by the rapid fall off of the fold of coverage. The improvement in the stack by simple bandpass filtering is considerable, again especially in the areas that were previously affected by ground-roll. The continuity and strength of reflections has been improved. Comparing the bandpass stack to those filtered by wavelet transform techniques, we can see that there is more energy present in the wavelet transform stacks (all stacks are plotted at the same scale). This can be attributed to the retention of energy outwith the ground roll area on the shot records which is removed by the global bandpass filtering process. This leads to reflector being more continuous across the stack, particularly at larger travel times.

3.9 Conclusions

The wavelet transform provides an efficient method of filtering seismic data in a domain similar to time-frequency, preserving frequency content of the seismic trace in areas which do not require filtering. The muting of coefficients in the wavelet domain has been shown to be effective in removing ground roll from seismic shot records whilst minimising the removal of any signal present. The use of a weighting procedure in the wavelet domain allows the preservation of frequencies within the area to be filtered whilst suppressing the ground roll. When compared to standard Fourier frequency filtering techniques the wavelet transform method gives effective results with no extra cost in computing time. Limitations associated with this method include the coarse nature of the time translation parameter and the fact that an octave may be too large a frequency range for filtering, especially at higher scales. To avoid coarse time translations other types of wavelet transforms could be used. The performance of the filters is dependant on the kernel wavelet used in the transform process, with the Battle-Lemarié kernel wavelets producing the best results.

The application of this method for the elimination of other unwanted signals, such as airblast, may be possible with variations in transform method such as the wavelet packet transform (Coifman *et al.*, 1989). We investigate this in the next chapter.
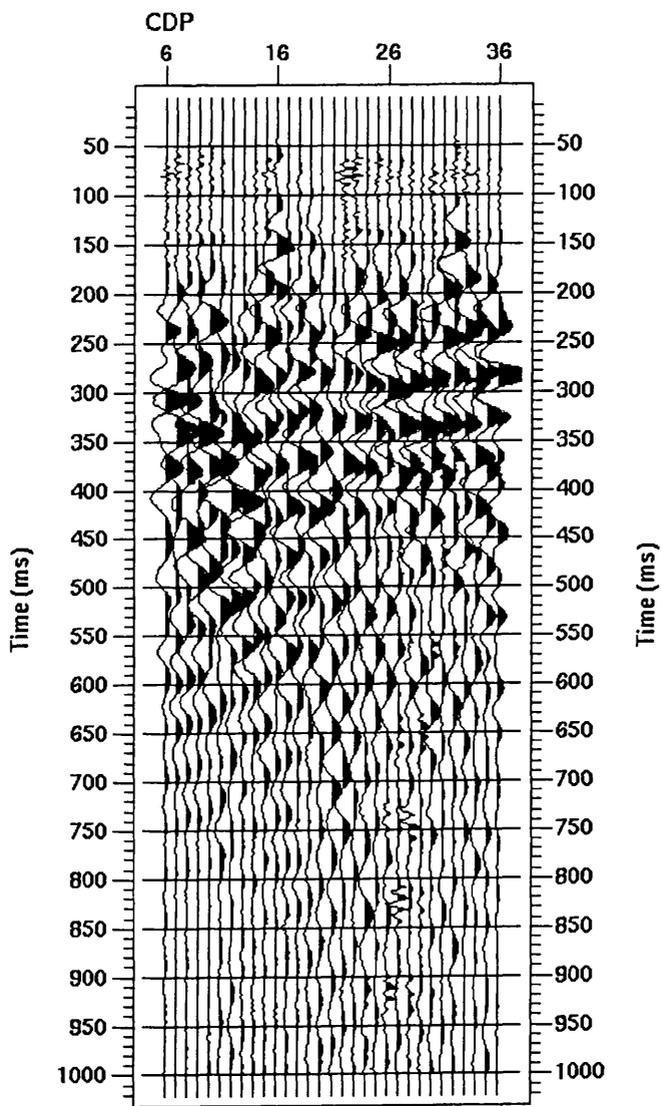
Figure 3-48 Brute stack of the Robroyston survey showing heavy contamination by low frequency signals.
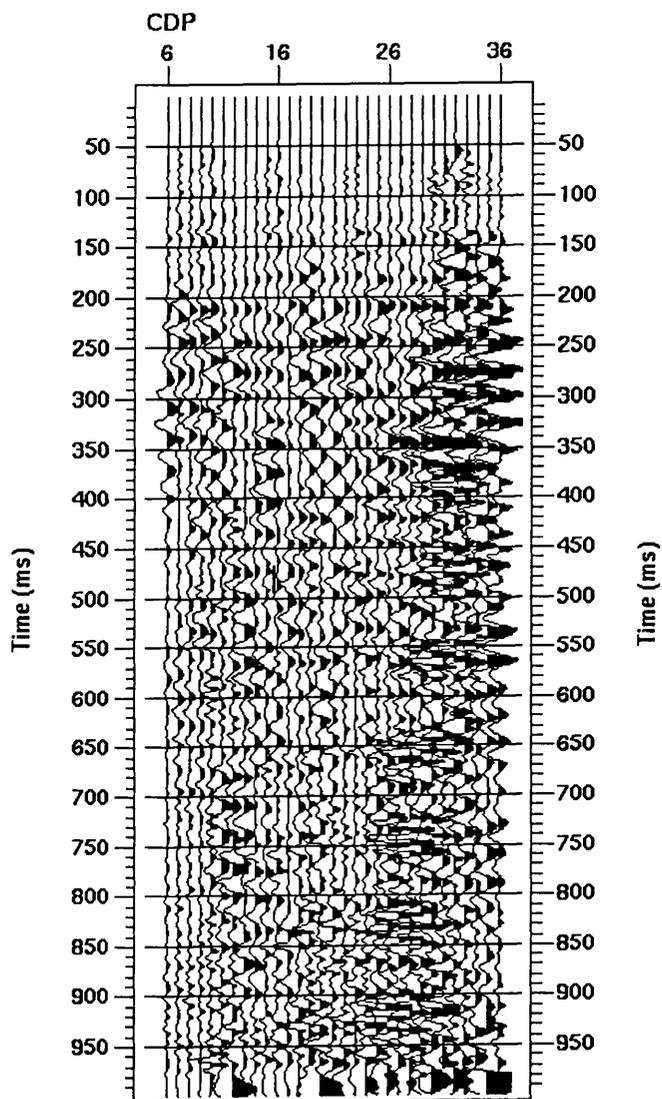
Figure 3-49 Robroyston stack from shot records filtered using the wavelet transform filtering technique (muting in wavelet domain). The stack has improved considerably.
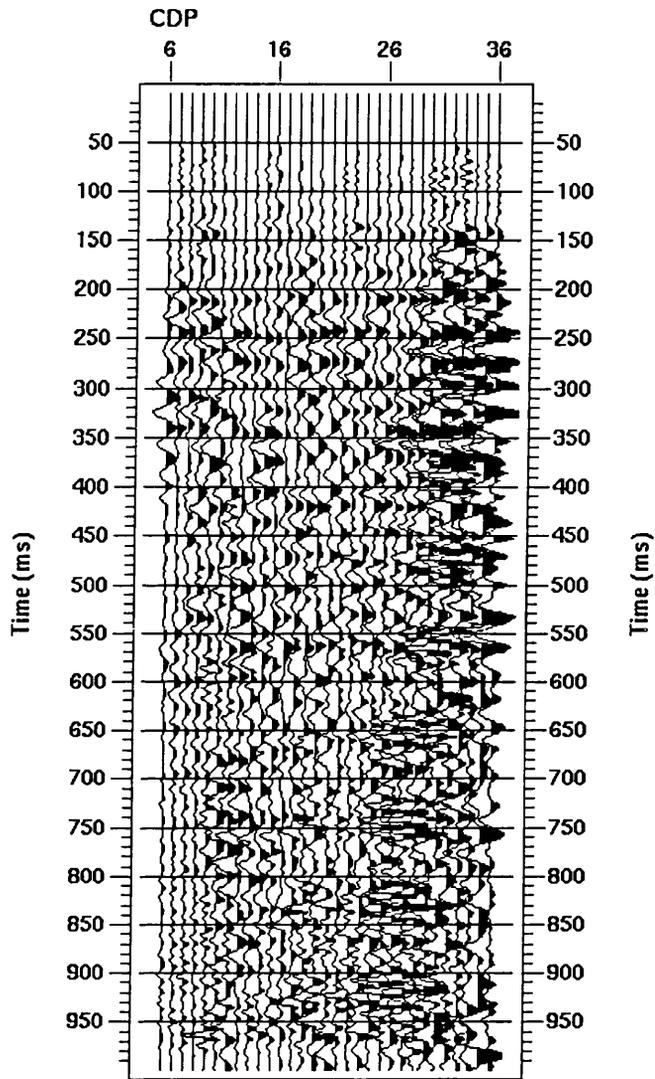
Figure 3-50 Robroyston stack from shot records filtered using the wavelet transform filtering technique (weighting in wavelet domain). The stack has also improved considerably.
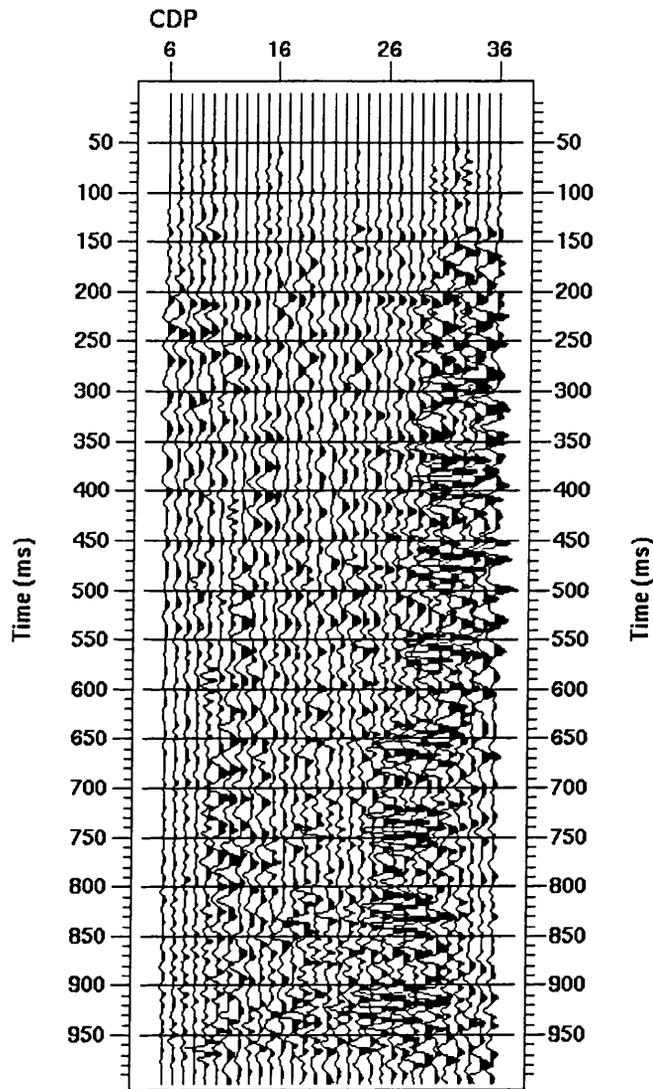
Figure 3-51 Robroyston stack from shot records which have been bandpass filtered. The scale is identical to the previous two figures. The stack is comparable to the previous two, yet has less energy present. This reduction in energy could be due to energy removed from the record by the filter outside the area contaminated by the ground roll.

4. Wavelet Packet Transform Filtering

4.1 Introduction

In the previous chapter we used the wavelet transform to filter seismic data in the translation-scale domain. We demonstrated that this filtering process has limitations in terms of time-frequency resolution at scales containing higher frequencies. Therefore this technique cannot be used effectively to filter noise centred around high frequencies or with high frequency content. To overcome this limitation we must substitute the wavelet transform with an alternative transform. As we discussed in chapter 2.5, the wavelet packet transform (Coifman *et al*, 1992) decomposes the octave scales of the wavelet transform into more adaptable time frequency cells and would appear to be the ideal tool for filtering the forms of noise that the wavelet transform cannot effectively filter. The wavelet packet transform is a natural extension of the wavelet transform and, in fact, the wavelet transform is a special case of the wavelet packet transform.

In this chapter we investigate the wavelet packet transform and as for the wavelet transform we develop it in the context of a pseudo time-frequency filter. Theoretically, the wavelet packet transform allows better frequency resolution, but we will show that this is at the expense of introducing additional side lobes into the basis wavelets frequency spectra that can lead to aliased noise in filtered signals. Due to the similarity of the wavelet and wavelet packet transforms, we can draw upon previous experience in terms of transform implementation and probable pitfalls. In this chapter we extend the investigation of wavelet transform implementation to wavelet packets and, as previously, we will look at the influence of basis wavelet on the filtering process, limitations of the technique and possible avenues of development of the process. We will demonstrate the filtering process by suppressing airblast from seismic shot records and evaluate the effectiveness of this technique compared to windowed Fourier based techniques and radon transform based techniques.

4.2 Wavelet Packet Representation and Basis Selection

With the wavelet transform, we used the scalogram to represent the wavelet transform coefficients. The wavelet packet transform is redundant in that for a N

sample signal it will return Nlog(N) wavelet packet coefficients from which we will select a wavelet packet basis to represent a basis for the original signal. Therefore, before the basis selection process we have multiple resolutions of the data in a overcomplete representation of the original data (**Figure 2-19**). Here we use resolution in the context of the shape of the Heisenberg cells. After one iteration of the transform process we have low frequency and high temporal resolution, whilst after several iterations we have higher frequency resolution, with correspondingly lower temporal resolution. We can represent a basis for the data with a single level of resolution, or by the combination of coefficients across levels of resolution, such as the wavelet transform representation shown in **Figure 2-19**. This process of selecting the basis across the levels of resolution is known as best basis selection. When the basis chosen consists of a single level of resolution, it is referred to as a best level.

For data compression, the selection of a best basis is fairly straightforward. We will select the basis from the wavelet packet coefficients that minimises the number of non-zero coefficients representing the data. That is we select the basis that produces the best compression ratio whilst minimising signal distortion (Bosman and Reiter, 1993). Correspondingly, when we filter a signal using the wavelet packet transform, we want to select the basis that leads to the best filtered record. However, what is the best basis to select for filtering a given signal? The basis selection techniques for data compression may not be well suited for suppressing coherent noise as they concentrate the maximum amount of energy into the least number of wavelet packet coefficients. Ideally we want to select the basis that leads to the optimum separation of signal and noise in the wavelet packet domain. We would then be able to filter noise from a signal, minimising distortion of the retained signal.

Rather than concentrating on optimising the best basis selection procedure for filtering, in this study we concentrate on evaluating the potential of the wavelet packet transform as a tool for filtering. When selecting the basis for filtering a signal, we will use a constant level of resolution. We will select the best level from the transform by visually inspecting the signal to be filtered at each resolution level of the transform and choose the level that best localises the noise. This can be considered as a human orientated best basis selection procedure.

At a constant level we have a regular grid of Heisenberg cells in frequency-time. We will refer to the rows of cells as 'packet scales', with the lowest scale containing the lowest frequency (**Figure 4-1**). This is not scale in the same sense as for the wavelet

transform where the basis functions were scaled by dilation. The wavelet packet transform modulates the basis functions using the associated quadrature mirror filters. The best level basis is the one that is most similar to the windowed Fourier transform as can be seen by comparing **Figure 2.4** to **Figure 2.19**. Therefore, the logical question to ask is why do we not use the windowed Fourier transform to filter signals? The answer to this is in several parts.

Time →

$$0 \quad \frac{t_{max}}{N} \quad \frac{2t_{max}}{N} \quad . \quad . \quad . \quad . \quad \frac{(N-1)t_{max}}{N} \quad t_{max}$$

M-1

M-2

Packet scale 2

1

0

0   1   2   3   .   .   .   N-1

Translation →

f_N

$\frac{(M-1)f_N}{M}$

Frequency

$2f_N/M$

f_N/M

0
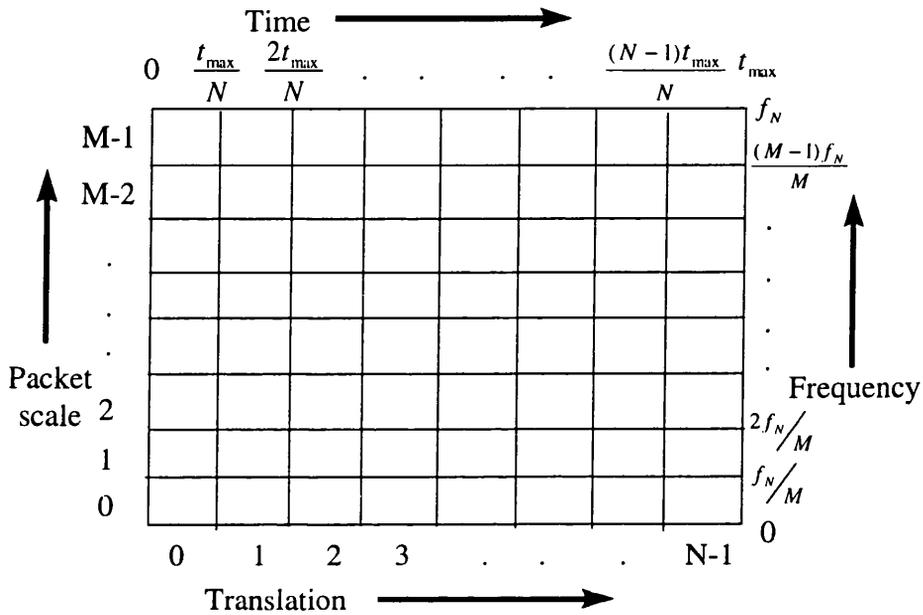
Figure 4-1 Schematic diagram showing the relationship between labelling of Heisenberg cells in the wavelet packet domain at a constant level of resolution, time and frequency. f_N corresponds to the Nyquist frequency and t_{max} the signal length in time.

First, the windowed Fourier transform uses a window of fixed length, inside of which is a basis of constant frequency. The wavelet packet basis at a constant level uses a window function that is almost fixed, but contains a range of frequencies, rather than a single frequency. Comparing the two processes at this stage is logical, but when we extend the selection of wavelet packet basis to across levels, we use a window function which changes in length adapting to the signal, which cannot be done with the windowed Fourier transform. We will be evaluating the wavelet packet transform, keeping in mind that it may be improved upon by a better basis selection procedure. This is not to say that the windowed Fourier transform is redundant, but only that it can be improved upon by the more adaptable wavelet packet transform.

Second, the windowed Fourier transform requires the selection of a window length. As we demonstrated in chapter 2.2, different window lengths are suitable for imaging different frequencies in the frequency-time plane. If we are filtering a signal that contains noise with a frequency range (as is the case with seismic signals), then the performance of the filter at each frequency will be dependant on window length. The wavelet packet transform automatically selects a range of window lengths (Heisenberg temporal cell widths) which are dependant on the sample rate of the original signal. For example **Figure 4-2** shows the wavelet packet transform of a signal at several resolutions. With each level, the temporal width of the Heisenberg cell increases and so can be considered to be similar to a series of windowed Fourier transforms with window widths incrementing by an integer number of sample intervals.

Third, the windowed Fourier transform samples the frequency-time plane more densely that the wavelet packet transform, with typical temporal increment of several sample intervals. With the wavelet packet transform, the sample density in the time-frequency plane is sparser, being dependant of the level of resolution in time-frequency space.

When we develop the wavelet packet transform as a tool for time-frequency filtering, we will compare and contrast the results with filtering the data with windowed Fourier techniques.
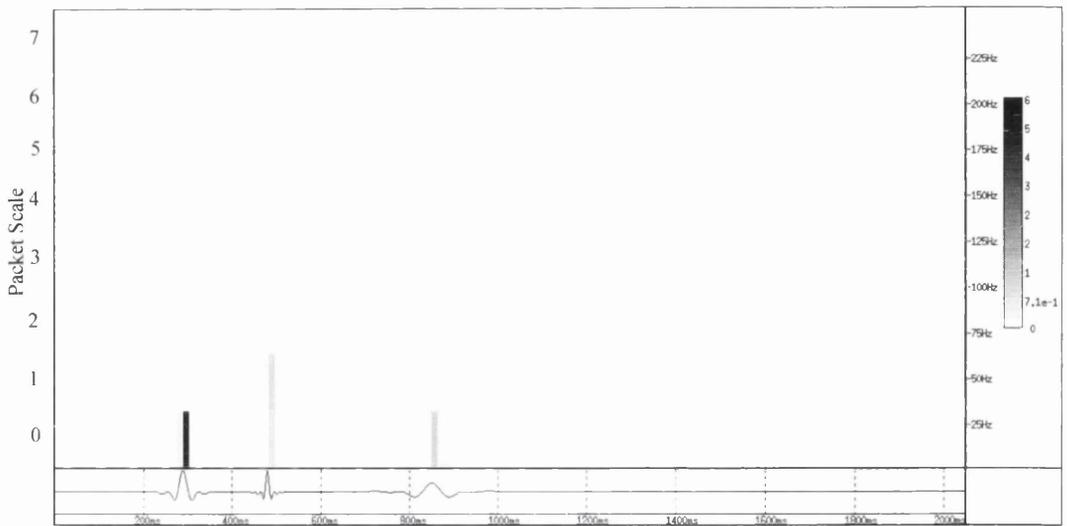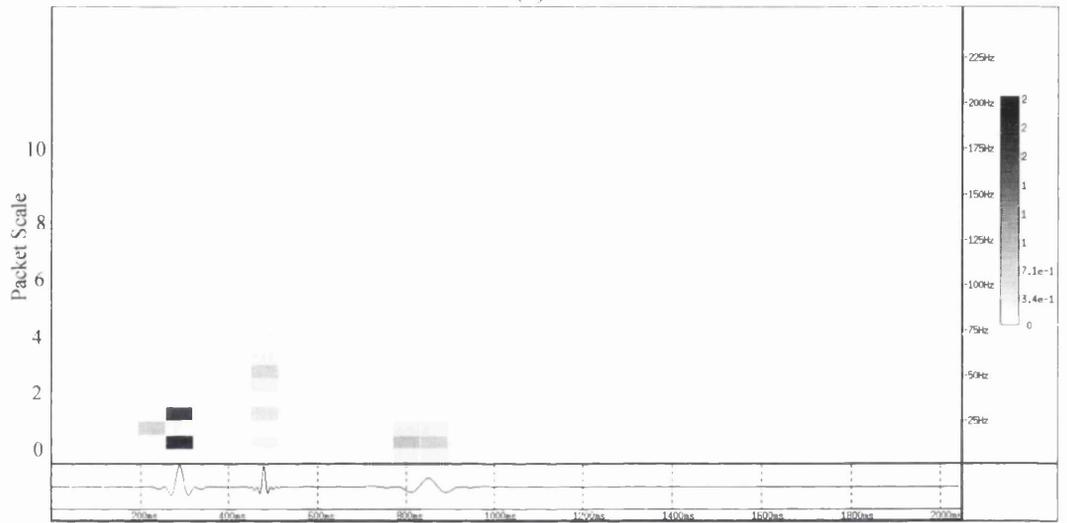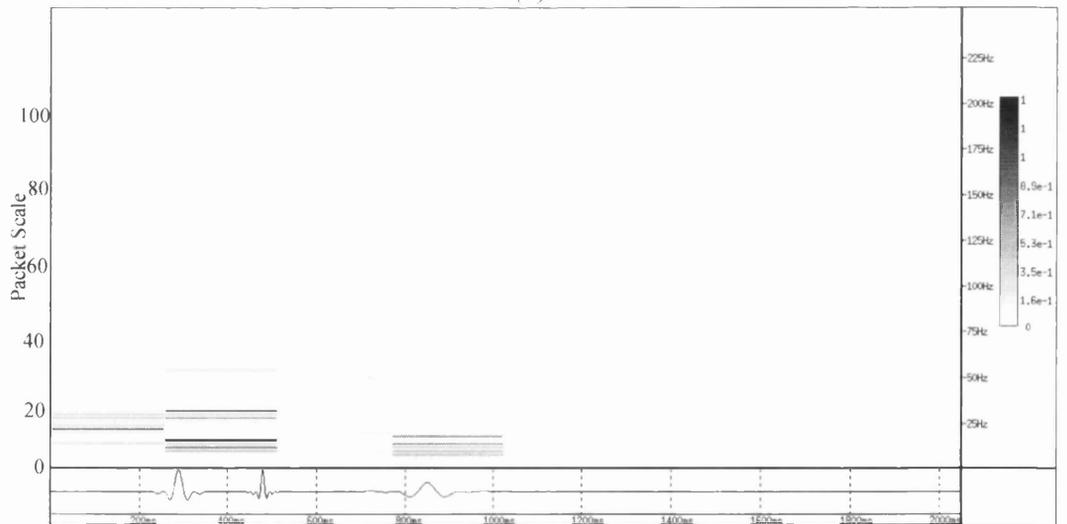
Figure 4-2 The wavelet packet transform of a signal at several resolutions after (a) four (b) five and (c) six iterations of the high/low pass transform operation. The signal is shown in the lower section of each diagram with translation being the horizontal axis and packet scale the vertical axis. The corresponding times and frequencies are shown for illustration.

4.3 Choice of Kernel Wavelet

From our investigations into the choice of kernel wavelet with wavelet transform filtering, we can state that the choice of kernel wavelet will also influence the performance of filters based on the wavelet packet transform. We base this statement on the fact that the wavelet transform and the wavelet packet transform are both implemented using quadrature mirror filters, as discussed in chapter 2.5. However, as the wavelet packet transform involves multiple combinations of the quadrature mirror filters (equations 2.39-2.41, chapter 2.5) we must investigate diminishes the influence of kernel wavelet choice when using wavelet packets.

Figure 4-3 shows the wavelet transform of a 10-50 Hz chirp signal using a cubic spline Battle-Lemarié basis wavelet. **Figure 4-4** shows the same signal at several levels of resolution using the wavelet packet transform. At the different levels of resolution in the wavelet packet transform the chirp signal is clearly identifiable as a diagonal line of coefficients (the clarity of which increase with the number of iterations). From this we can see that the wavelet packet transform allows more flexible resolution in time-frequency at higher frequencies apparently allowing better time-frequency resolution than could be achieved with the wavelet transform. However from **Figure 4-4c** (level 5) we can see that as well as localising the chirp signal in time-frequency, the wavelet packet transform has introduced some coefficients off the main diagonal of the chirp signal. These coefficients represent areas in frequency-time that were not present in the original signal and may be artefacts of the transform process.
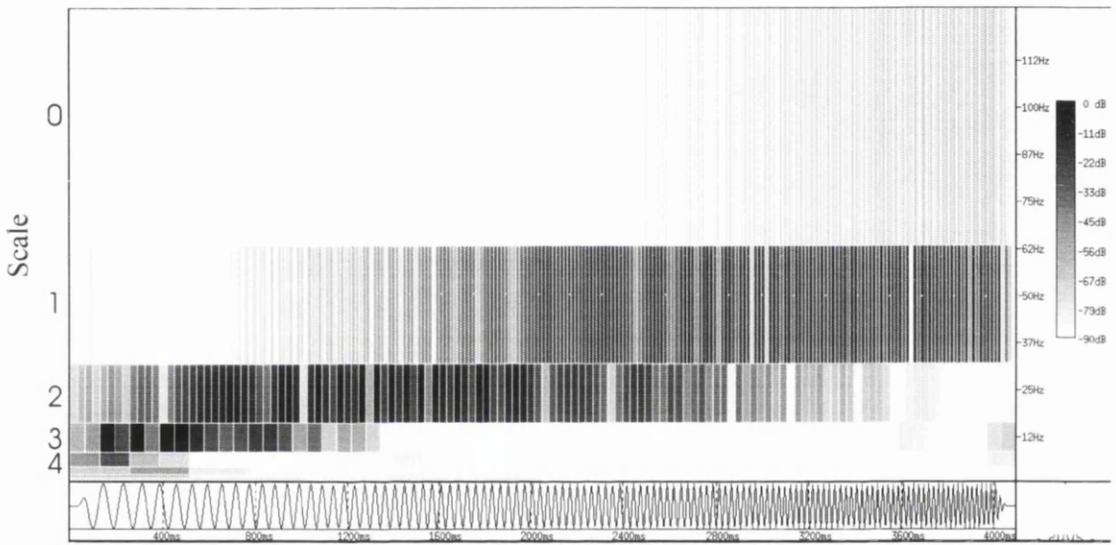
Figure 4-3 Scalogram representation (in decibels) of the wavelet transform of a 10-50 Hz chirp signal using the Battle-Lemarié cubic spline kernel wavelet.
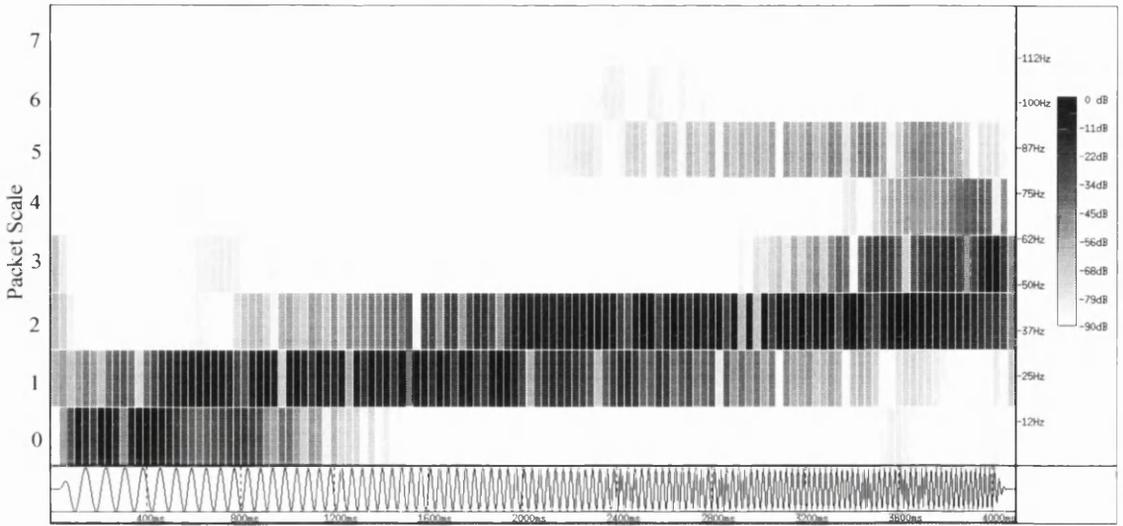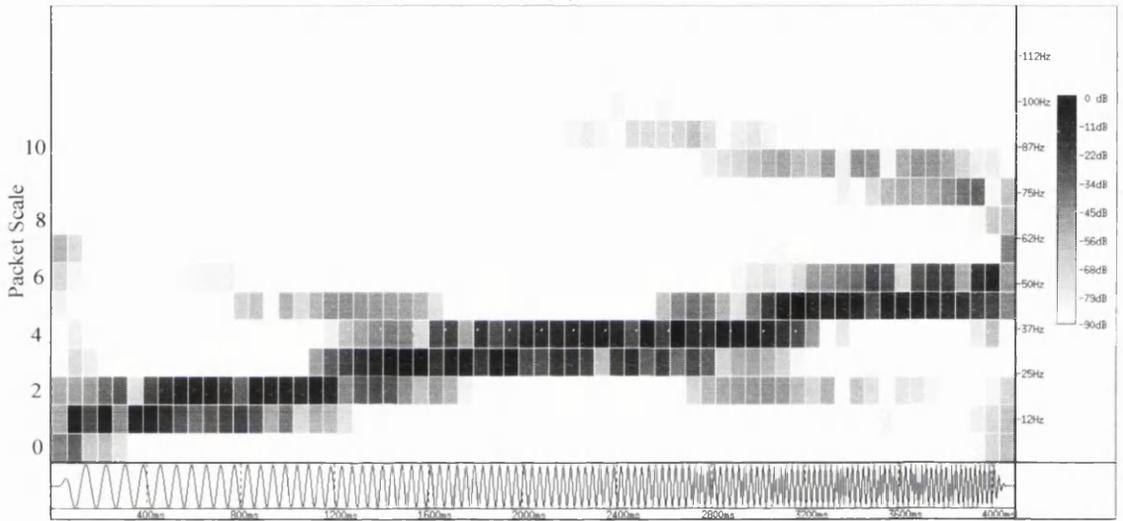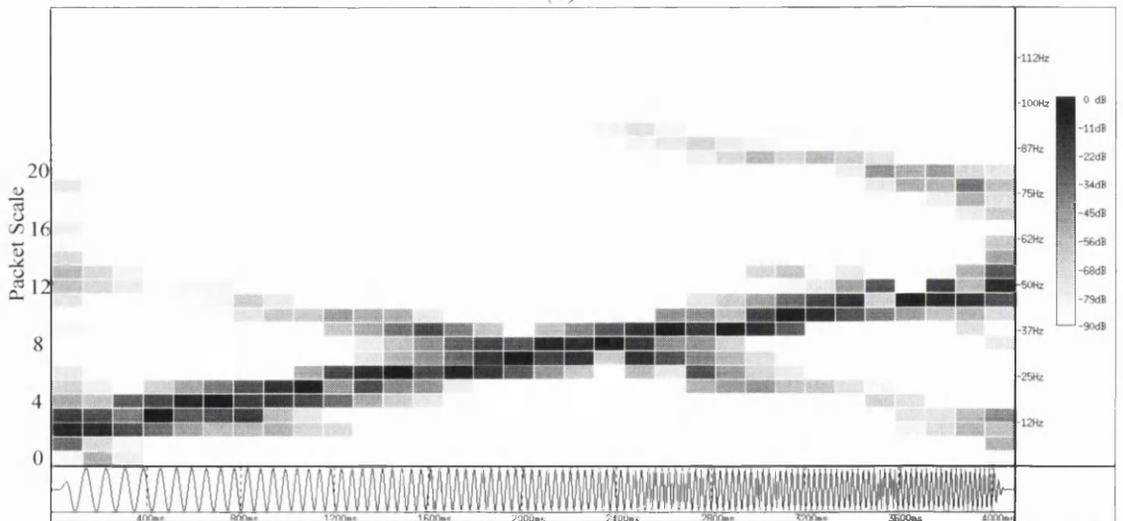
Figure 4-4 Wavelet packet transform a 10-50 Hz chirp signal at several resolutions using a Battle-Lemarié kernel wavelet. After (a) three (b) four and (c) five iterations of the transform process.

Analysis of the frequency spectra of the wavelet packet basis functions (**Figure 4-5**) at the resolution level shown in **Figure 4-4a** explains the source of these side lobes. The frequency spectra of the packet scales contain side lobes off the main frequency band centred around the Heisenberg cell limits (the boundaries marked on the diagram), which vary in amplitude depending on the packet scale of the basis function. During the decomposition process, the wavelet packet transform uses these basis functions with large side lobes to represent a part of the signal, and so the transform must later compensate for these side lobes by introducing a coefficient whose main frequency band matches the frequency of the side lobe. This leads to the off diagonal coefficients observed in **Figure 4-4**.

The large side lobes are a result of the splitting trick used in the implementation of the wavelet packet transform process. These side lobes are present in the wavelet basis as well as the wavelet packet basis as was seen in **Figure 3.2**, but not to the same extent. Amplification of these side lobes occurs when we extend the wavelet transform to wavelet packets. **Figure 4-6** shows the generation of the side lobes, by comparing the spectra of two wavelet packets at adjoining packet scales to the spectrum of the original wavelet from which they were split. In the transform process the quadrature mirror filters split the wavelet at a given scale into two wavelet packets, one consisting of the high frequency portion of the original spectrum, the other the low frequency content. We can see that the two wavelet packet spectra sum to give the wavelet spectra with the side lobe being an imperfection in the splitting process. This imperfection is a result of deviations of the quadrature mirror filters from the ideal sinc function high/low pass filters. As for the wavelet transform, the energy in these side lobes and other energy outwith the Heisenberg cell become aliased, a product of the downsampling process in the transform. The wavelet packet transform accounts for the aliased energy perfectly in the forward-inverse transform process. However if we filter the coefficients in between, the inverse transform will not account fully for the aliased energy and we may introduce aliased noise centred around the frequency of the filtered wavelet packet side lobes. Implementing other seismic processing techniques, such as migration, using the wavelet packet transform also suffers from this problem as was identified by Foster and Mosher (1994). This aliased noise may manifest itself as residual signal (of the signal that was being filtered) in the filtered area. One way of minimising the introduction of this aliased noise is to minimise any side lobe energy.
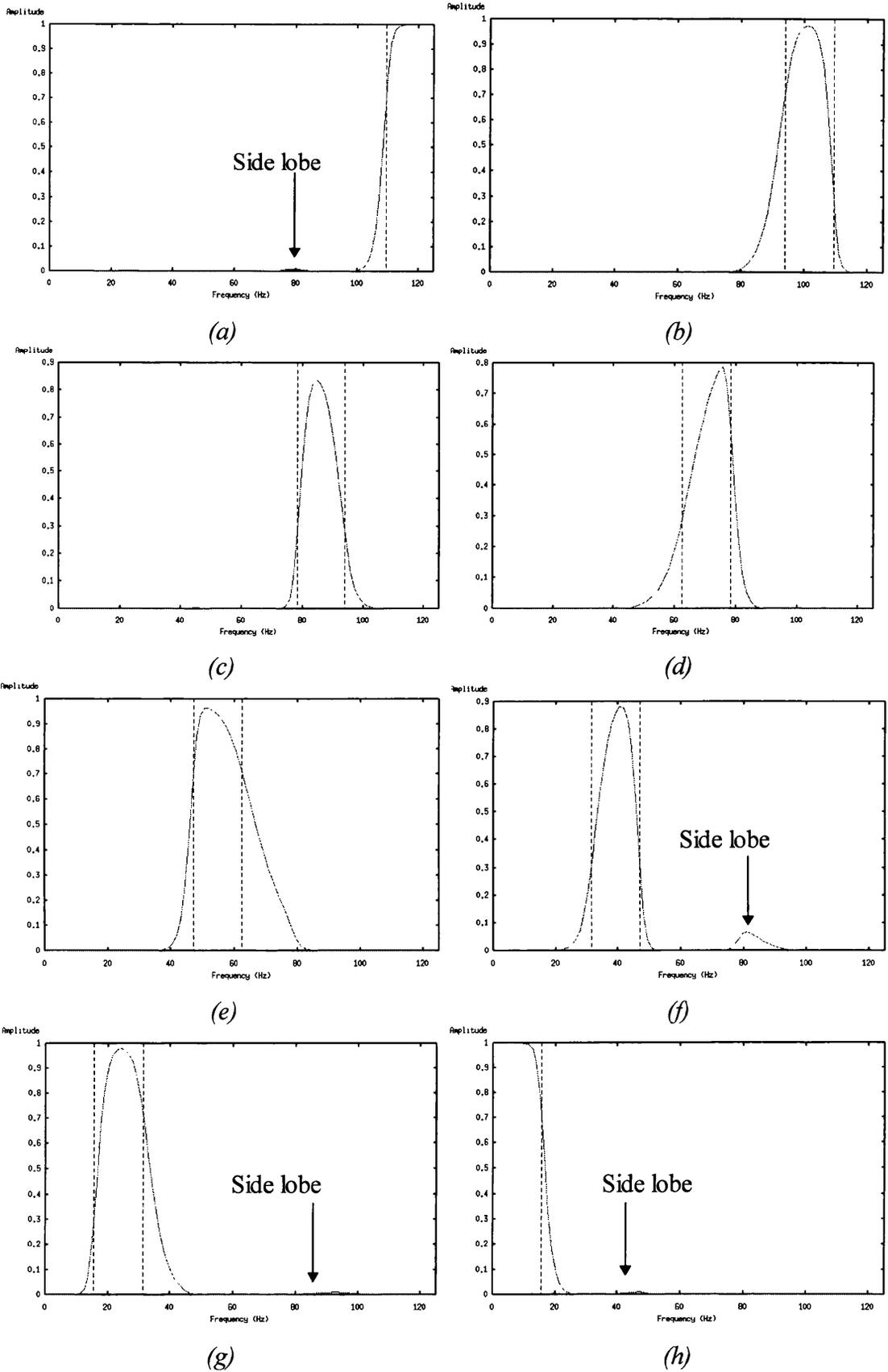
Figure 4-5 Frequency spectra of the wavelet packet basis functions for the cubic spline Battle-Lemarié kernel wavelet after three iterations of the transform step. Side-lobes to the main frequency band are apparent in the spectra leading to deviations from a true frequency-time representation.
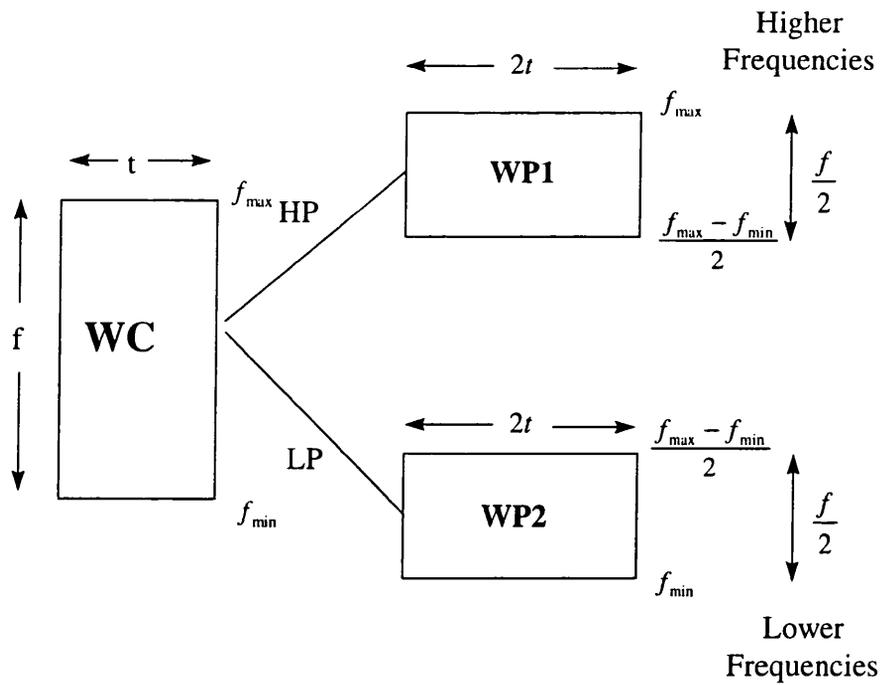
(a)
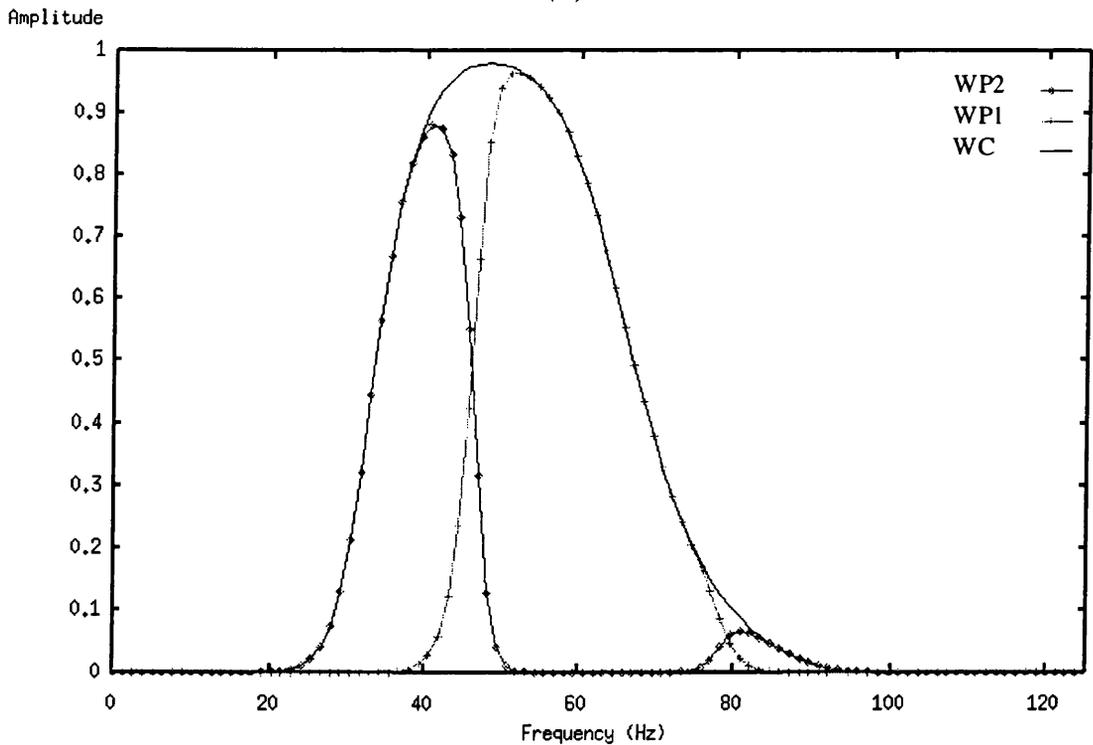
(b)

Figure 4-6 Splitting of the wavelet space, WC, by the high band high pass (HP) and low pass (LP) quadrature mirror filters into the two wavelet packet spaces WP1 and WP2 (a) shown schematically as the splitting of a Heisenberg cell and (b) shown by the splitting of the wavelet frequency into the two wavelet packet spectra.

From the analysis we performed on wavelet transform kernel wavelets, we can predict that kernel wavelets with the least amount of side lobe energy will minimise the amount of wavelet packet side lobe energy. **Figure 4-7** demonstrates this, showing the frequency spectra for a wavelet packet basis function for the 20 coefficient least asymmetric Daubechies wavelet and the quintic spline Battle-Lemarié kernel wavelet. From this it is apparent that the side lobe energy in the quintic spline kernel wavelet is considerably less than that contained in the Daubechies wavelet. An extension to this premise is that the smoother the wavelet, the smaller the side lobe energy in the frequency domain. From this figure we can also note that as for the wavelet transform, the quintic spline Battle-Lemarié kernel wavelet has a more desirable (flatter) response for filtering over the frequency range of the wavelet packet and a quicker amplitude decay at the Heisenberg cell edges.
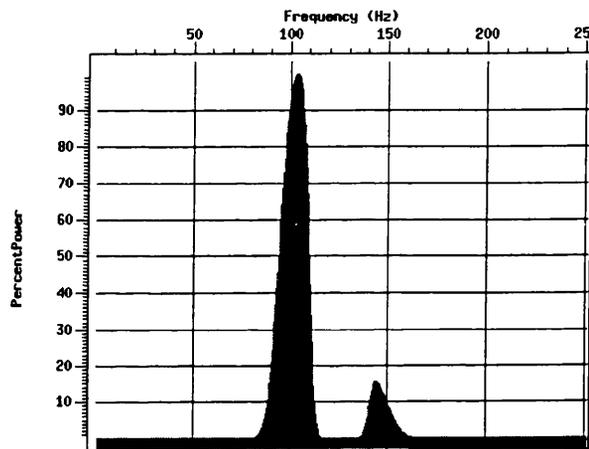
(a)
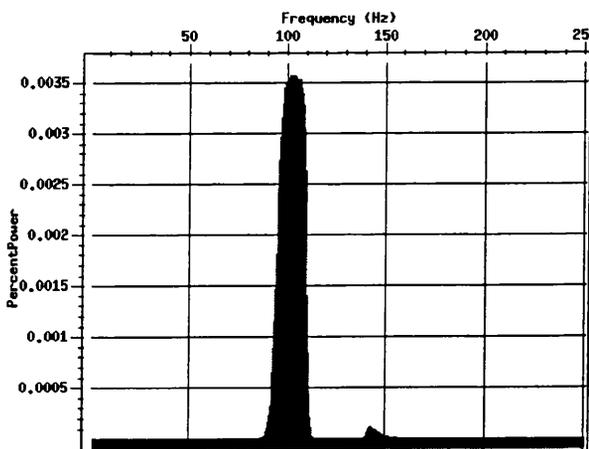
(b)

Figure 4-7 Frequency spectra of wavelet packets derived from (a) the 20 coefficient least asymmetric Daubechies and (b) Battle-Lemarié quintic spline kernel wavelet showing that the side lobe energy diminishes with smoother kernel wavelets.

To examine the influence of the side lobe energy on the performance of filters we will apply a crude wavelet packet filter to the synthetic signal shown in **Figure 4-8**. As for the wavelet transform we will zero all the coefficients in the range of packet scales corresponding to the signal, examining the performance of the filter without using the temporal properties of the transform. This will allow us to estimate the effect of aliased noise on any filtered signal. **Figure 4-8** shows the wavelet packet transform at a single level before and after filtering, illustrating the packet scales removed. **Figure 4-9** shows the result of applying this filtering procedure to the signal with several different kernel wavelets. The varying degrees of residual signal show the sensitivity of this process to the choice of kernel wavelet. From the figure we can see that the amount of residual signal due to the aliased energy is minimised by the Battle-Lemarié and 20 coefficient least asymmetric Daubechies wavelet. As for the wavelet transform, the kernel wavelets with fewest coefficients in the associated filters give the poorest filtering results.

Figure 4-10 shows the frequency spectra of the synthetic trace before and after filtering from which we can see the main frequency band that has been suppressed. We can also see that the quintic spline Battle-Lemarié wavelet minimises the peak energy left after filtering. The 20-coefficient Daubechies wavelet also performs well but introduces notches into the frequency spectrum which may be problematic if the signal is to be processed further.

The influence of the kernel wavelets in the wavelet packet transform and the wavelet transform are similar, which is not surprising due to the relationship between the transforms. We can apply the criteria we developed for selecting a kernel wavelet for use with wavelet transform filtering for filtering with the wavelet packet transform. The linear phase and flat frequency response properties of the kernel wavelet are desirable for wavelet packet filtering, and the minimisation of energy outwith the band represented by the Heisenberg cell is a tighter constraint with wavelet packet filtering than for wavelet transform filtering due to potentially large side lobe energy. The temporal extent of the basis wavelet packets outside the bounds of the Heisenberg cell directly corresponds to the smoothness of the wavelets as for the wavelet transform as shown in **Figure 4-11**. Therefore we can state that the criteria for choosing a suitable wavelet for filtering with the wavelet packet transform is the same as for selecting a kernel wavelet for wavelet transform filtering with the extra constraint on the frequency bounds of the signal.
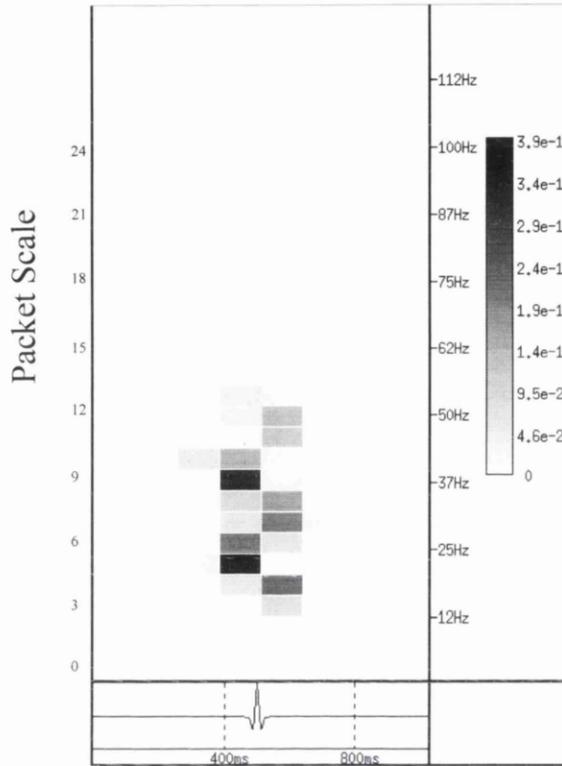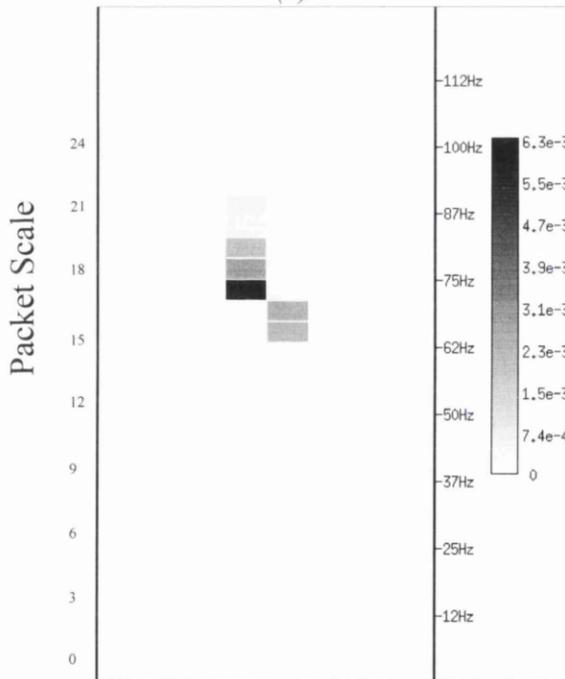
Figure 4-8 (a) A synthetic signal and the associated wavelet packet transform using a Battle-Lemarié cubic spline kernel wavelet (b)The transform after filtering showing the packet scales removed by the filtering process to evaluate the effect of aliased energy.
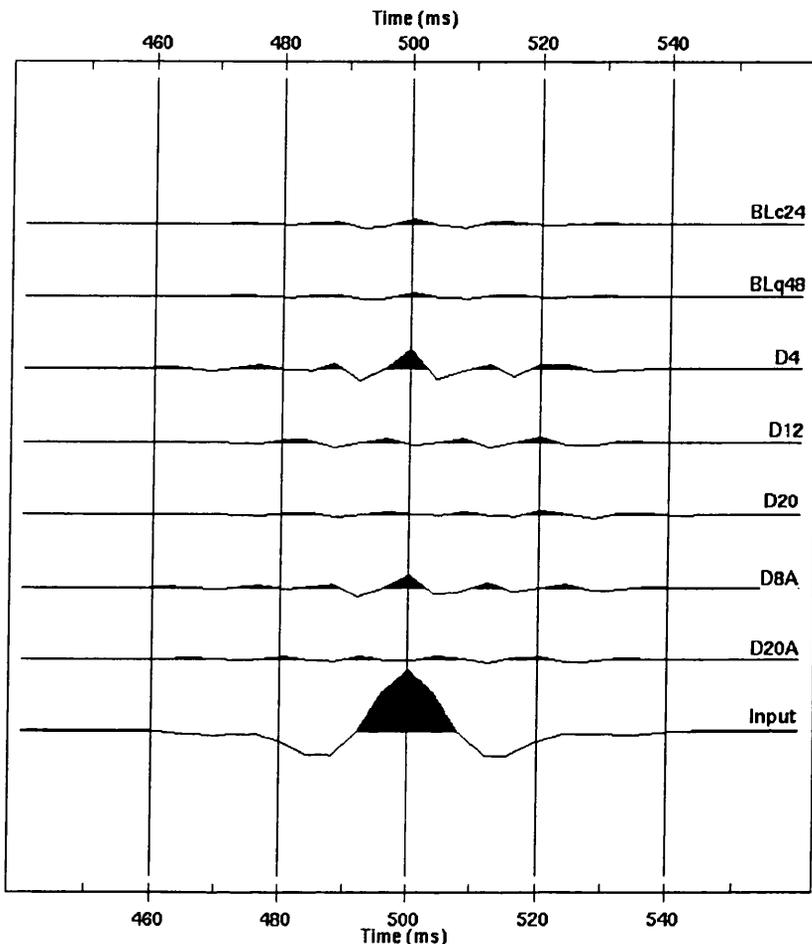
Figure 4-9 Filtering of the synthetic signal using various kernel wavelets showing the sensitivity of the filtering process to kernel wavelet choice and so to the kernel wavelet side-lobe energy. The following kernel wavelets were used: BLc24- Battle-Lemarié cubic spline wavelet with 24 coefficients, BLq48- Battle-Lemarié quintic spline wavelet with 48 coefficients, D4- 4 coefficient Daubechies wavelet, D12- 12 coefficient Daubechies wavelet, D20- 20 coefficient Daubechies wavelet, D8A- 8 coefficient least asymmetric Daubechies wavelet, D20A- 20 coefficient least asymmetric Daubechies wavelet.
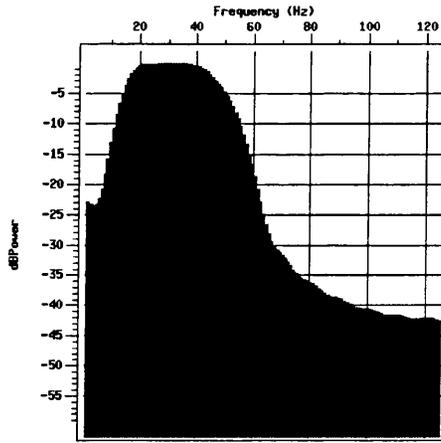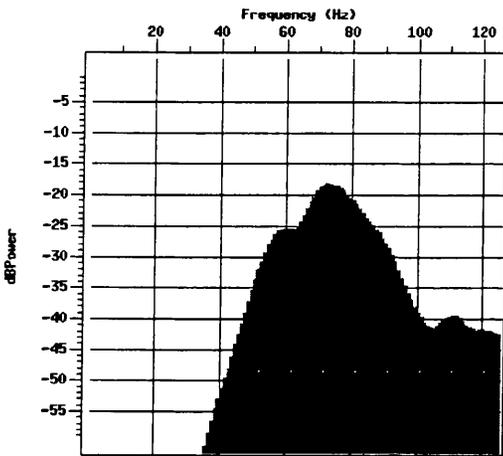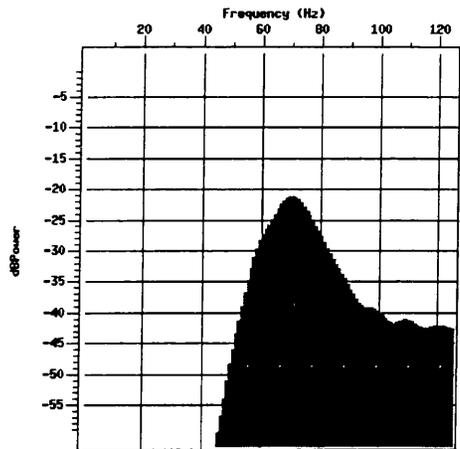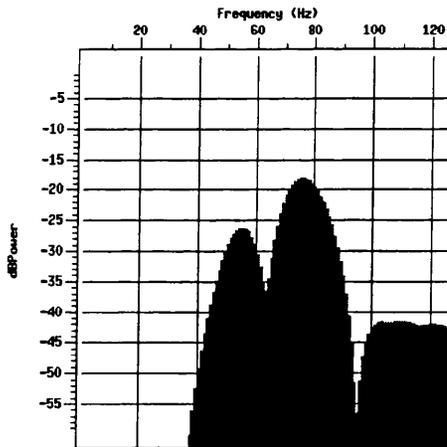
Figure 4-10 Frequency spectra of selected traces from Figure 4-9 showing the varying degrees of frequency suppression by different kernel wavelets (a) Input signal, output signal from the filtering process using the (b) Battle-Lemarié cubic spline wavelet (c) the Battle-Lemarié quintic spline wavelet and (d) the 20 coefficient least asymmetric Daubechies wavelet.
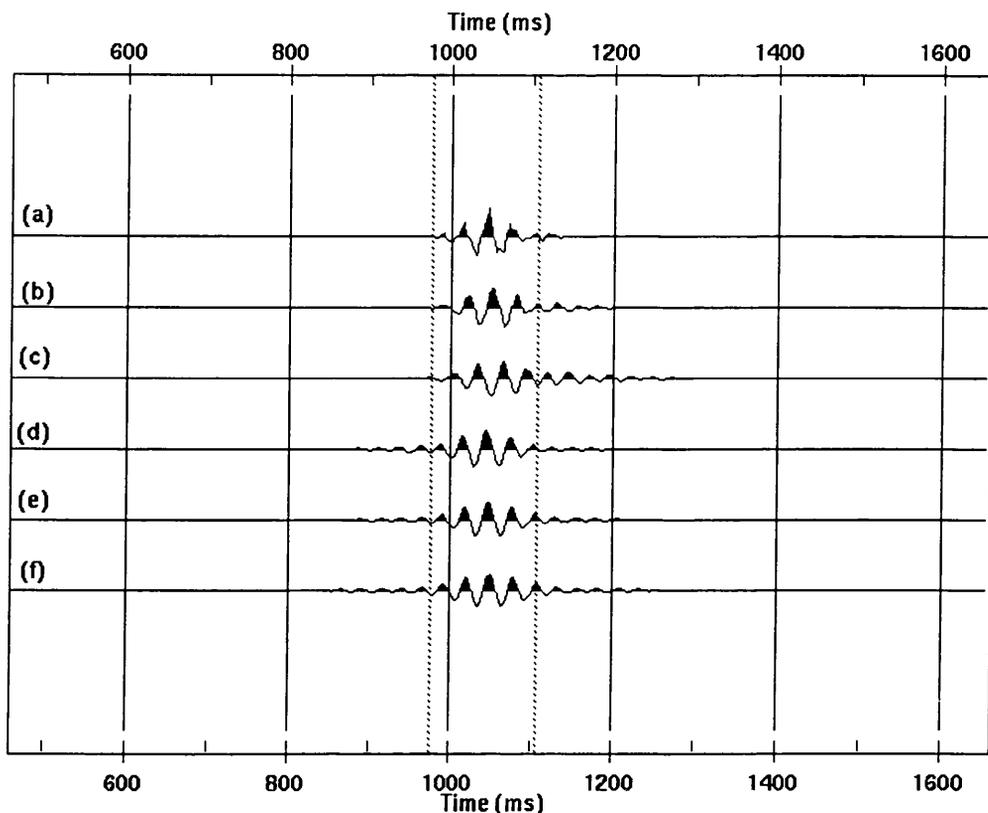
Figure 4-11 Temporal support of wavelet packets at a single level derived from selected kernel wavelets showing the variation in temporal support with kernel wavelet. The smoother the kernel wavelet in time the greater the extent of the wavelet packet outside the boundaries of the Heisenberg cell (marked by dashed lines). The wavelet packets are derived from (a) the 4 coefficient Daubechies wavelet (b) the 8 coefficient Daubechies wavelet (c) the 16 coefficient Daubechies wavelet (d) the 20 coefficient least asymmetric Daubechies wavelet (e) the Battle-Lemarié cubic spline wavelet and (f) the Battle-Lemarié quintic spline kernel wavelet.

Again, as for the wavelet transform, we must make a trade off between speed of the transform and the properties of the kernel wavelet. The smoother the kernel wavelet, the more coefficients in the corresponding quadrature mirror filters. This in turn leads to extra iterations in the transform process.

4.4 Transform Implementation

As the discrete wavelet packet transform is implemented in the same way as the discrete wavelet transform, using quadrature mirror filters as was discussed in chapter 2.5. The restrictions and rules governing the discrete wavelet transform also apply to the discrete wavelet packet transform. That is, the signal length must be an integer power of two, and the transform assumes that the data are cyclical. Therefore, the techniques we used for implementing the discrete wavelet transform are used for

implementing the discrete wavelet packet transform. Padding of the signal can be performed using zero padding, cosine tapering or signal mirroring. The edge handling procedure is the same in that the signal can be treated as if it were periodic, giving exact reconstruction, or aperiodic which can introduce an associated error. In the filtering examples in this chapter we will use the periodic boundary condition combined whilst mirroring the samples about the last sample when signal padding is required.

The discrete wavelet packet transform, like the discrete wavelet transform is translationally variant. If we translate the input signal by one sample, the discrete wavelet packet transform changes. This is an obvious point as the downsampling which causes the translational variance is included in both transforms. Therefore, the points raised in the previous chapter regarding translational variance apply equally to the discrete wavelet packet transform.

4.5 Filtering Methodology

We can use the wavelet packet transform as a time varying filter in the same way that we used the wavelet transform. However, if we ignore the side lobe problem (assuming we have chosen a kernel wavelet that minimises this), the wavelet packet transform allows better time-frequency resolution than the wavelet transform. This will allow us to filter noise localised in higher frequency areas. To filter a given signal we must specify several parameters that will control the application of this process. Firstly we must define the resolution level to decompose the data to in the wavelet packet domain. This choice depends on the form of noise that we want to remove and its distribution in the wavelet packet domain. We must make the choice of decomposition level by inspecting the data in the transform domain and choosing the level that best localises the noise in the wavelet packet domain which will subsequently lead to the minimum distortion to the retained signal.

After choosing the level of decomposition, we must define the area in the wavelet domain that we wish to filter. Again, we determine this by inspection, using time-gates and packet scale parameters to filter the data.

As for the wavelet transform, to filter the data in the wavelet packet domain we can mute or weight the coefficients. Muting coefficients leads to a form of time-frequency mute, and we subsequently lose any signal information contained in the muted coefficients. In the weighting procedure, again as for the wavelet transform, we

apply a weight based on the Q-value of the signal, as discussed in the previous chapter. With the wavelet packet transform at a constant level of resolution we have a regular grid of coefficients in a domain similar to frequency-time. The Q-surface for a signal can be estimated and used to weight wavelet packet coefficients that have energies larger than that expected form the Q-surface in frequency-time. Again, this weighting technique is simply one technique that could be used in an attempt to preserve signal and suppress noise and is primarily used to demonstrate the potential of the weighting process. The technique is based on Fourier techniques and therefore does not use the full potential of the wavelet packet transform, but simply illustrates an alternative to muting the wavelet packet coefficients.

Figure 4-12 shows a signal containing a localised noise burst between 2800 ms and 3200 ms. Figure 4-13 shows a three dimensional view of the absolute amplitudes of the corresponding wavelet packet transform at a constant level. The estimated best-fit Q-value surface for the signal is shown in Figure 4-14. Figure 4-15 shows the absolute amplitudes of the wavelet packet transform that are greater than the Q-value surface. From this we can see that at the time of the noise burst, at frequencies of 15 to 70 Hz the wavelet packet coefficients lie above the estimated Q-surface for the data. We can also see at earlier times that the wavelet packet coefficients peak above the surface. This is due to a combination of the side lobe problem discussed earlier where we observe artefacts of higher frequencies that do not represent true time-frequency components of the signal and the quality of the best-fit Q-value surface. These will not be filtered as the zone to be filtered is specified by a combination of packet scales and temporal translations.
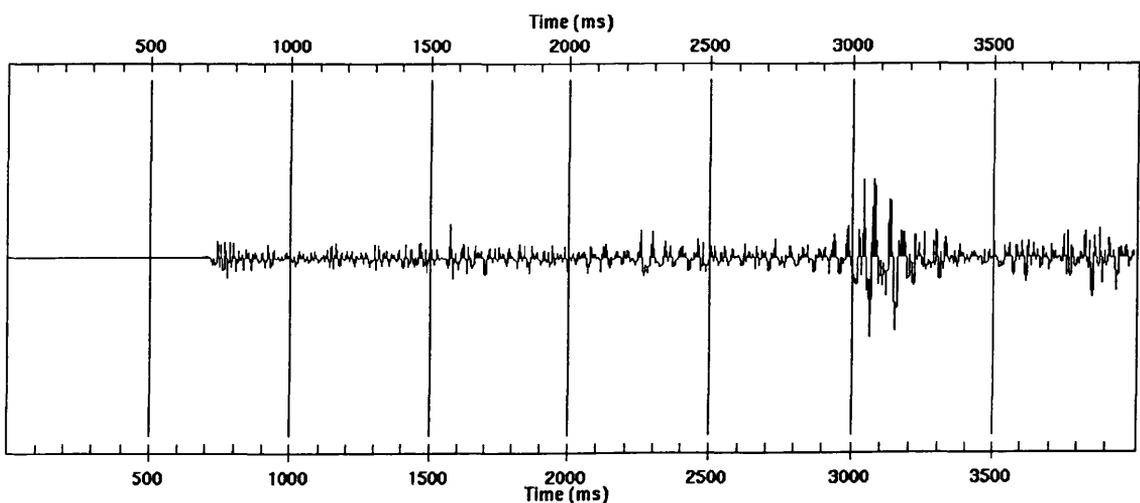
Figure 4-12 A seismic signal containing noise burst between 2800 ms and 3200 ms.
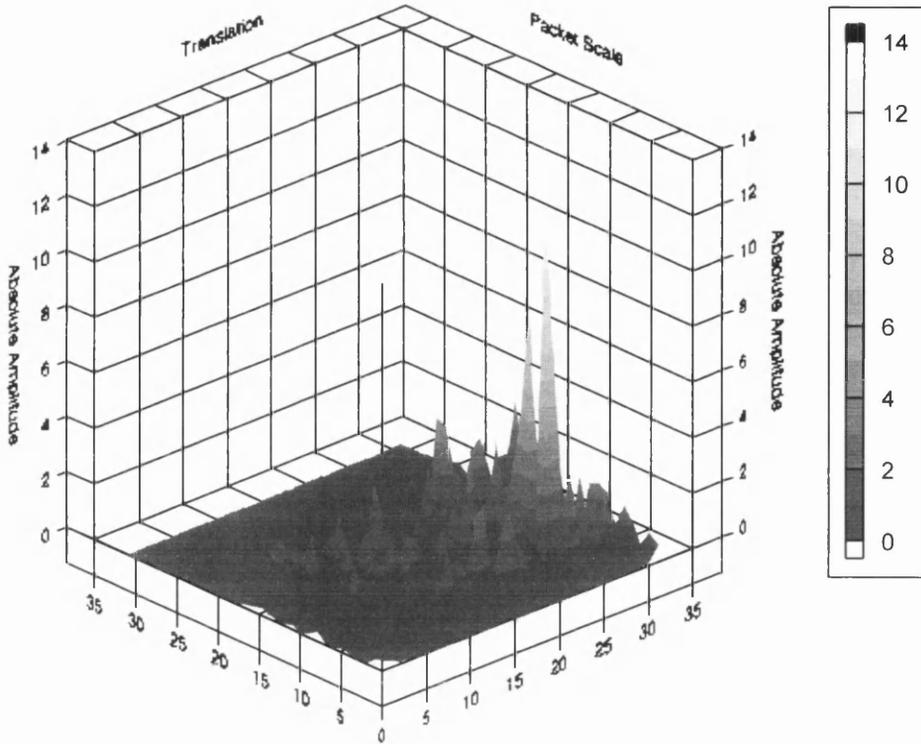
Figure 4-13 A 3-dimensional representation of the wavelet packet transform of the signal shown in Figure 4-12. The absolute amplitude of the coefficients is plotted against packet scale and translation.
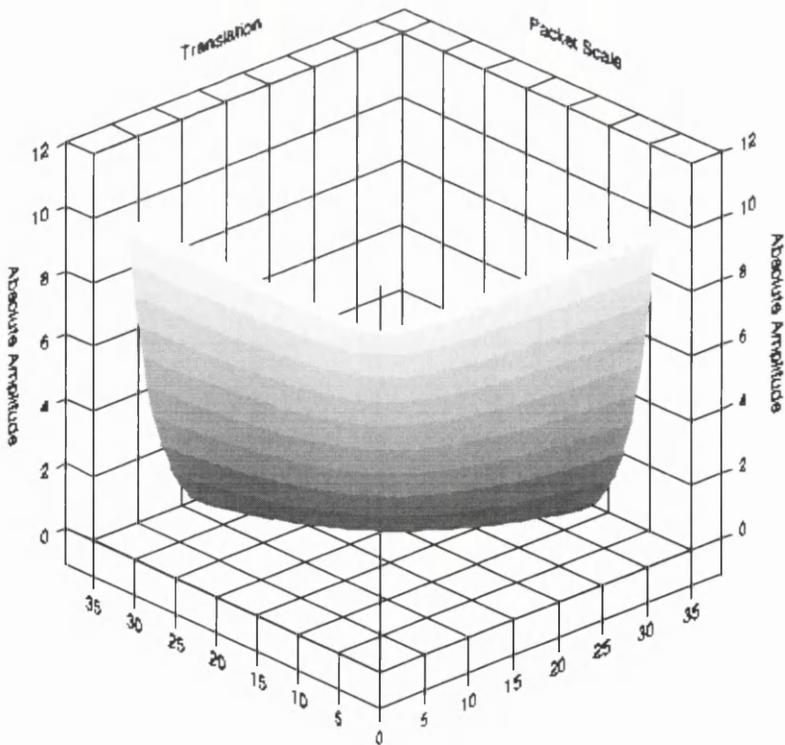
Figure 4-14 The best fit Q-value surface for the signal shown in Figure 4-12 which we will use to weight the wavelet packet coefficients.
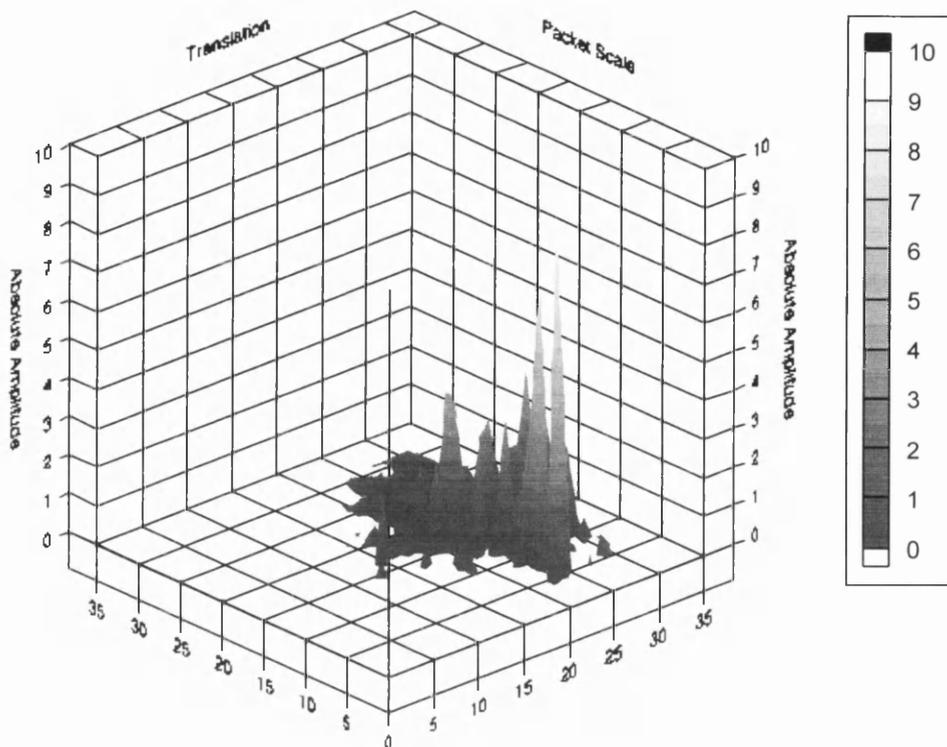
Figure 4-15 The absolute amplitude of the component of the wavelet packet transform shown in Figure 4-13 which lies above the Q-value surface shown in Figure 4-14.

In the zone we want to filter, we multiply the wavelet packet coefficients by a weight, suppressing them such that the weighted coefficients lie underneath the Q-value surface. The weight is the value that, when multiplied by the maximum wavelet packet coefficient of the zone being filtered, gives the value of the Q-surface at the centre point of the Heisenberg cell in frequency-time. We then weight all the coefficients by multiplying all the coefficients in the filter area by the weight. This effectively lowers the wavelet coefficients so that they lie below the Q-value surface. The result of this process on the transform is shown in **Figure 4-16** and the resultant filtered signal in **Figure 4-17**. From this we can see that the noise burst has been suppressed successfully.
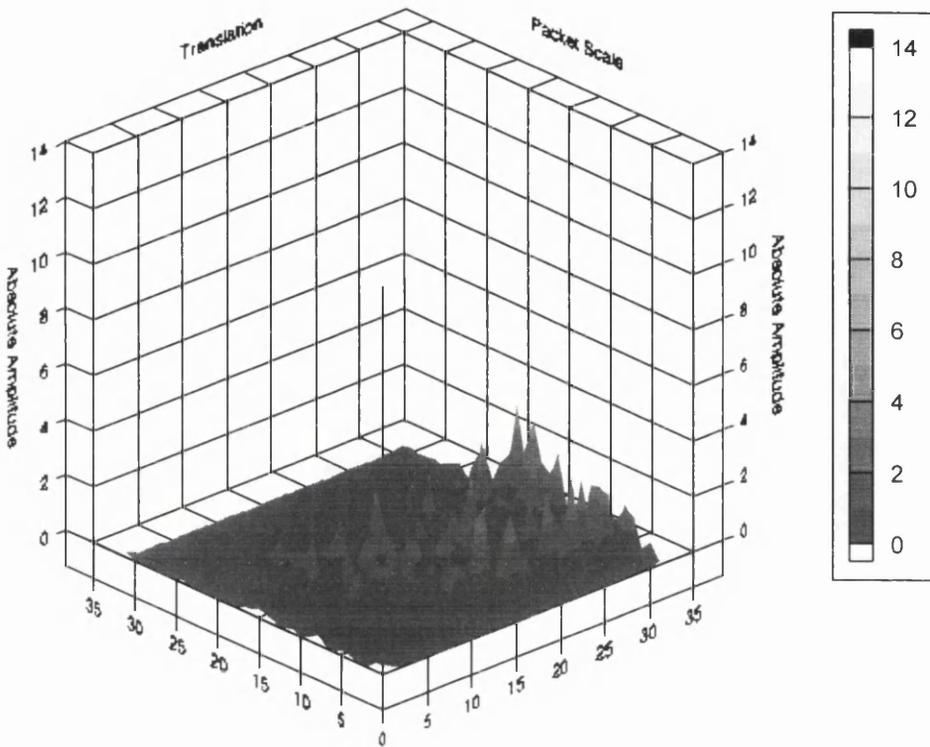
Figure 4-16 The wavelet packet representation shown in Figure 4-13 after filtering using the weighting procedure.
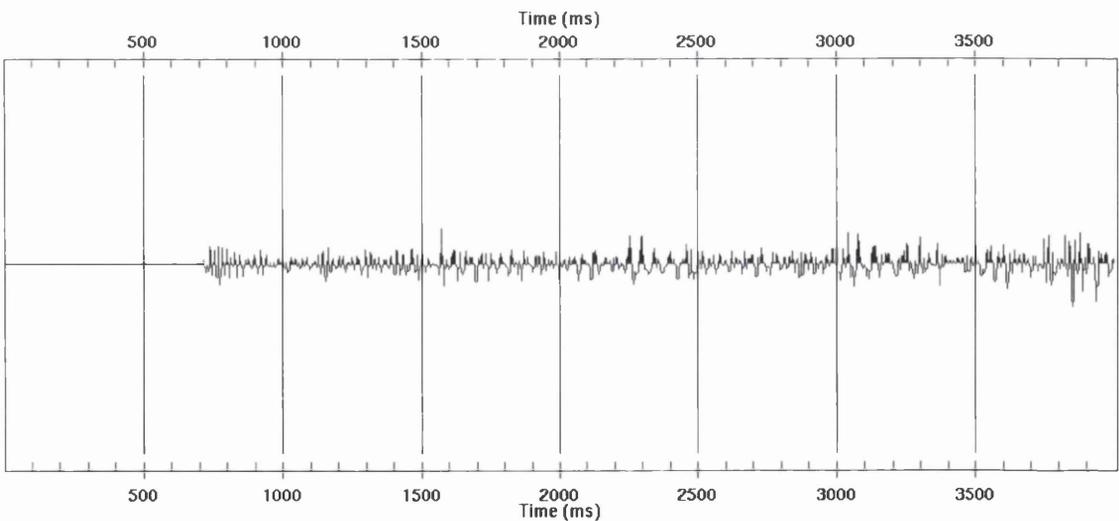
Figure 4-17 The signal shown in Figure 4-12 after filtering using the weighting procedure in the wavelet packet domain. The noise burst has been successfully suppressed whilst preserving the bandwidth of the signal.

4.6 Filtering Examples

In the last section we developed the wavelet packet transform as a form of time-frequency filter. In this section we apply this filtering technique to the suppression of airblast from seismic records.

Airblast, an air-coupled Rayleigh wave, is a result of sound energy coupling with the Earth's surface and is a common feature of land-based seismic records, particularly when shooting with surface sources. Evison (1956) showed that when the phase velocity of a surface wave is approximately equal to the velocity of airborne sound travelling over the surface, then energy transfers across the surface in spite of the great disparity in acoustic impedance. The airblast energy transfers to the ground and so insulation or burial of geophones is an ineffective method of removing airblast from the seismic record (Knapp, 1986). The airblast manifests itself on shot records as an arrival with linear moveout equal to the velocity of sound in air. **Figure 4-18** shows a common receiver gather from a land-based seismic survey with acquired using an explosive source. We can see from this figure that the receiver gather contains noise associated with airblast. The low velocity of the airblast leads to spatial aliasing of the signal which precludes the use of *f-k* filters for airblast suppression and limits radon transform based filters (Yilmaz, 1987).

The use of radon techniques to filter data that is highly spatially aliased can lead to end effects which generate linear streaks on radon filtered gathers. **Figure 4-19** shows a synthetic common shot gather containing signal similar to airblast and the corresponding linear radon transform. From this we can see that the airblast event is not localised in the radon domain. Therefore, for the suppression of airblast, radon techniques may lead to the introduction of artefacts into the data (Yilmaz, 1987). **Figure 4-20a** shows the linear radon transform of the record shown in **Figure 4-18**. In the radon domain we can see the linear streaks corresponding to the airblast that we saw with the synthetic gather. Filtering the record in the radon domain involved muting the data to the right of the line marked on **Figure 4-20a** and inverting the transform. The filtered record is shown in **Figure 4-20b**. From the filtered record we can see that the spatially aliased airblast energy remains and the filtered record has a wormy character.

In addition to these limitations, radon transform based techniques are two dimensional, filtering seismic gathers, and so lead to the additional problem of dealing with dead traces, and are inherently more time consuming to apply. Other airblast
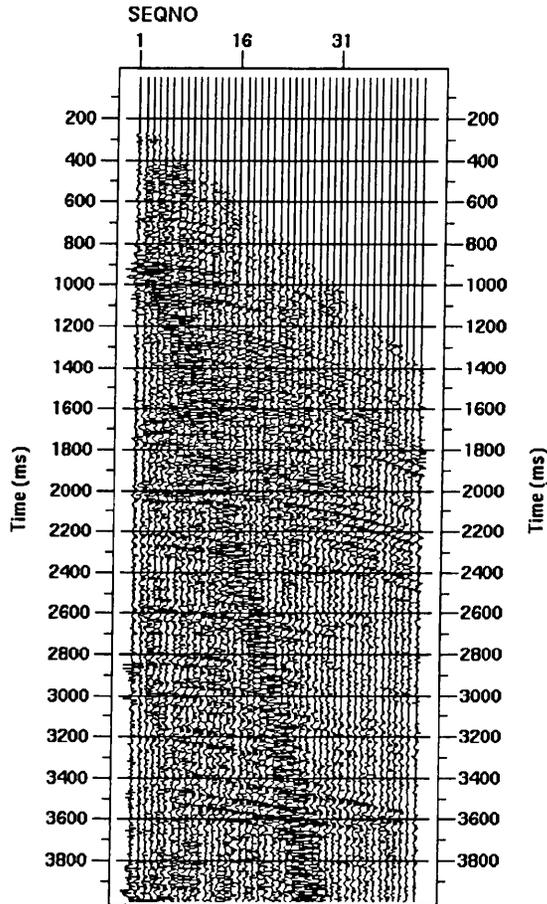
Figure 4-18 A common receiver gather contaminated by steeply dipping airblast.
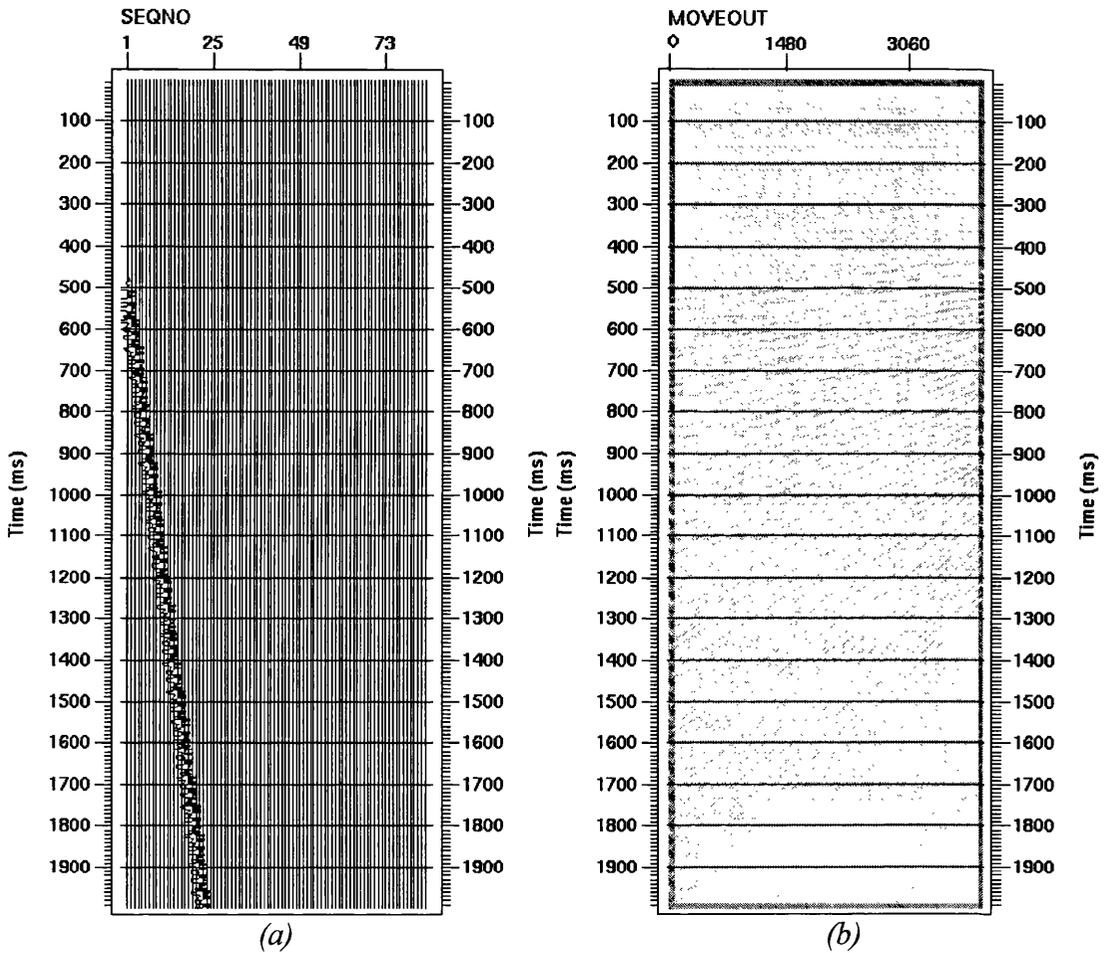
Figure 4-19 (a) A synthetic shot record modelling airblast and (b) the corresponding linear radon transform showing the poor localisation of this form of noise in the radon domain. This is due to the spatial aliasing associated with this form of noise.
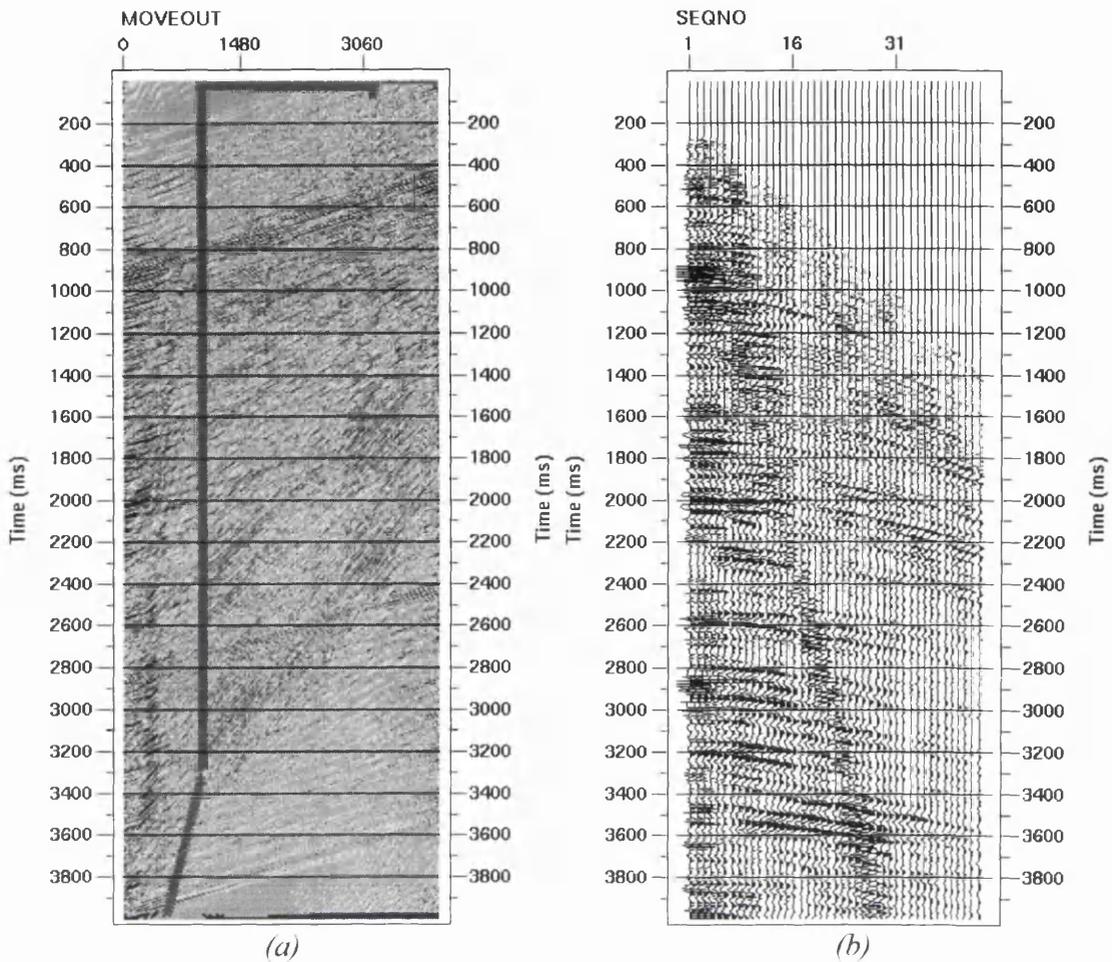
Figure 4-20 (a) The linear radon transform of the gather shown in Figure 4-18 where the linear signals similar to Figure 4-19 are due to the airblast. (b) The common receiver gather after filtering using the radon transform where all the signal to the right of the line in (a) has been muted. The airblast is still present on the record due to poor signal noise separation in the radon domain.

attenuation techniques include surgical muting, which is a crude but effective method of removing airblast from seismic records but also leads to the removal of any associated signal.

Figure 4-21 shows a common receiver gather containing 84 traces with a 30 m offset increment between traces contaminated by airblast. The data were collected using an explosive source in a transition land-marine area. **Figure 4-22** shows the twenty-second trace from this common receiver gather at several resolutions in the wavelet packet domain. From this we can see that five iterations of the transform process gives acceptable localisation of the airblast noise in the wavelet packet domain. Time gates were selected on the common receiver gather to define the translation areas in the wavelet packet domain to filter. Similarly, the packet scales to be filtered were selected by inspection of the wavelet packet transform best level basis. The 5th to 18th packet scales were selected for filtering after five iterations of the wavelet packet transform

process. **Figure 4-23a** shows the result of applying the wavelet packet filter using the muting technique in the wavelet packet domain. We can see that the airblast has been successfully suppressed, but there is degradation in the reflected signal crossing the airblast contaminated area. **Figure 4-23b** shows the result of filtering after weighting in the wavelet packet domain. The Q-value for the signal was estimated from examining the wavelet packet transform and comparing it to several Q-value surfaces. We can see that the airblast has, again, been successfully suppressed without the degree of signal degradation observed with the muting technique. Finally, for comparison purposes, **Figure 4-23c** shows the result of windowed Fourier filtering of the common receiver gather to suppress the airblast. The same time gate used for calculating the wavelet packet translations to be filtered was used for windowed Fourier filtering. The frequency parameters for the filtering process were derived from the frequency limits of the Heisenberg cells of the packet scales filtered, allowing for filter tapering. We can see that this produces a similar result to the wavelet packet muting technique, with the signal degradation being slightly more pronounced. If we were more precise in our frequency definition and taper specification for the windowed Fourier filter we could match the performance of the wavelet packet muting technique. However, we would be unable to match the wavelet packet weighted filter result.
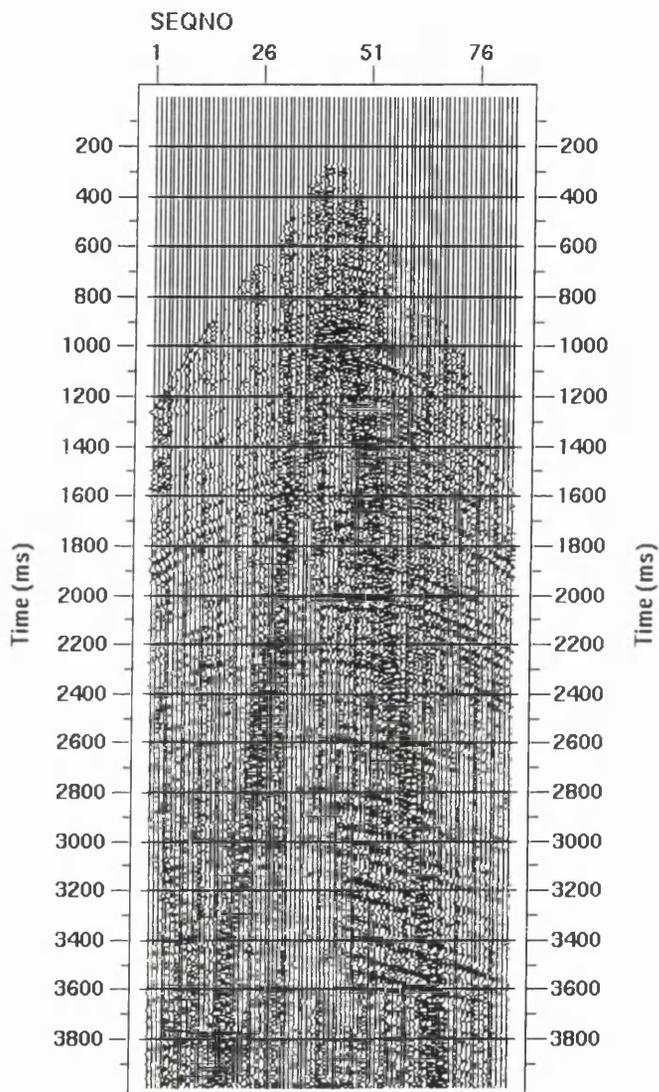
Figure 4-21 A common receiver gather contaminated by airblast. The offset increment between traces is 30 m.
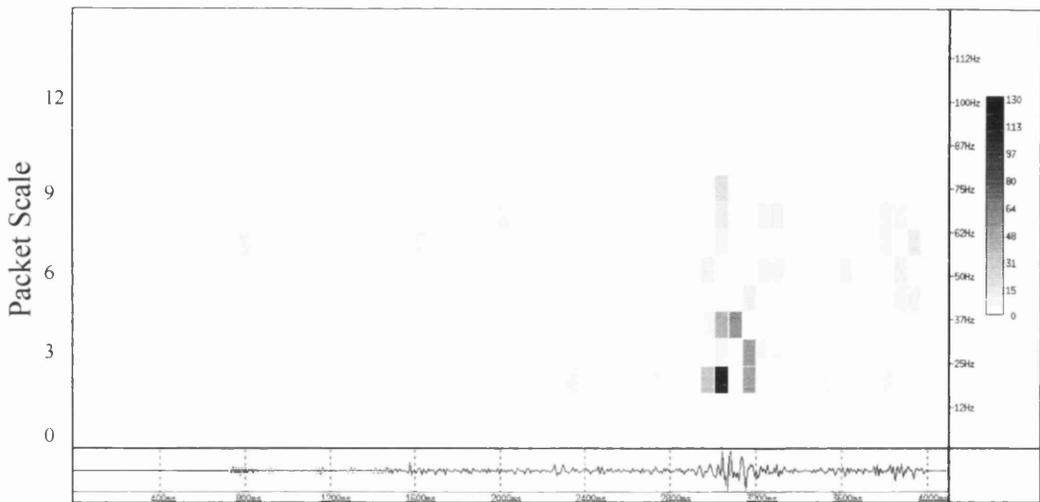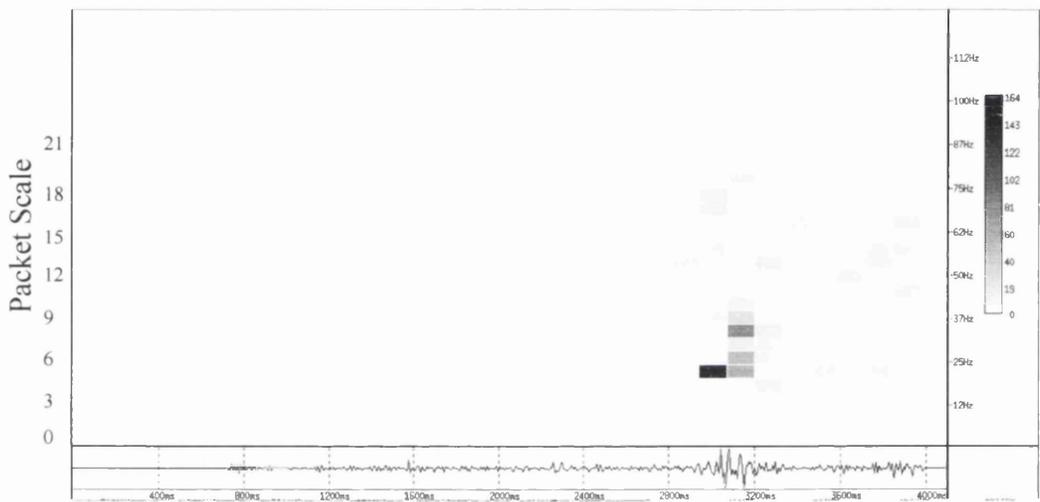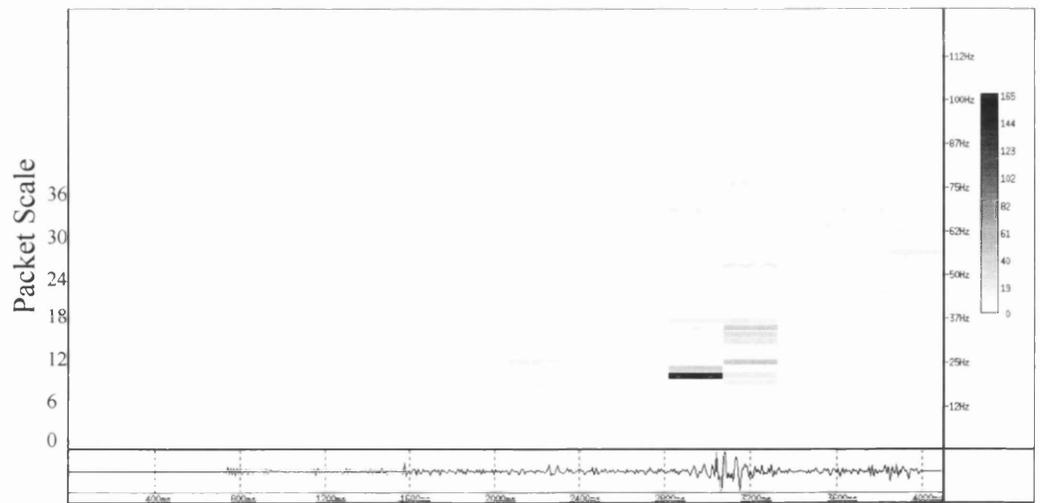
Figure 4-22 The wavelet packet transform of trace 22 from Figure 4-21 at several levels of resolution (a) after four, (b) five and (c) six iterations of the transform process. The airblast noise burst seems most localised in (b).
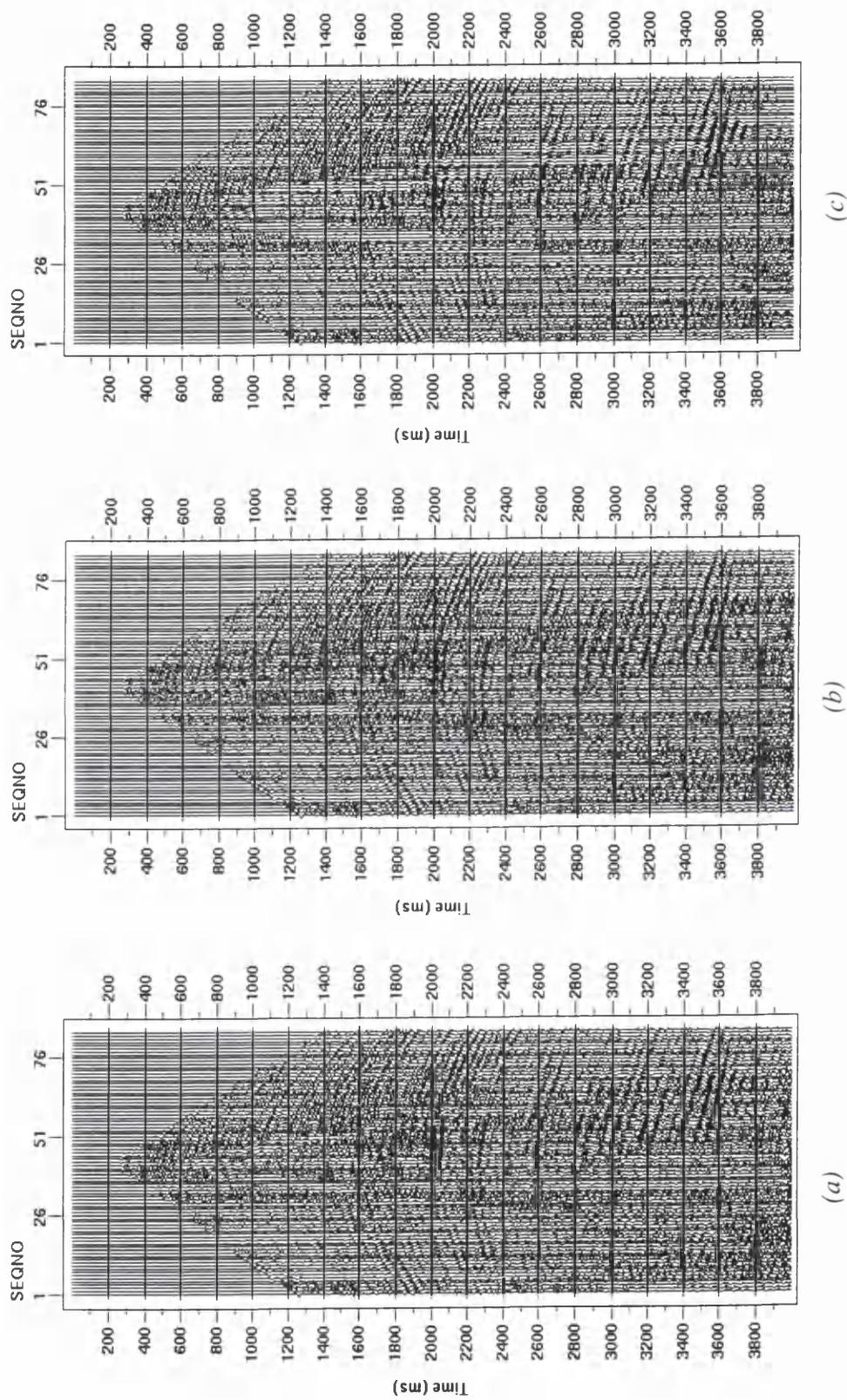
Figure 4-23 The common receiver gather shown in Figure 4-21 after filtering by (a) muting in the wavelet packet domain (b)weighting in the wavelet packet domain and (c) windowed Fourier filtering.

4.7 Conclusions

The wavelet packet transform allows a more adaptable tiling of the time frequency plane by the modulation of the standard wavelet basis by the associated quadrature mirror filters. We have demonstrated that this adaptability is at the expense of introducing side lobes into the basis function frequency spectra which could degrade the quality of any subsequent filtering process. The size of side lobe energy is related to the smoothness of the original kernel wavelet, and through careful choice of kernel wavelet we can minimise this effect. We have shown that this adaptability allows the wavelet packet transform to be used to filter frequency-time space in the packet scale-translation sense allowing the suppression of noise that is localised in frequency time. Although perfect separation of signal and noise is not possible using this technique, the filtering process minimises any signal losses by localising the filtering process. This technique allows filtering of areas in the frequency-time plane that could not be filtered effectively using the wavelet transform.

We have shown that through a simple wavelet packet basis selection process, satisfactory filtering results can be achieved and so, through possible future development of the basis selection procedure, we could improve upon these filtering results. Filtering signals using best level bases are effective, but further improvements in performance could be achieved if we adapt the basis selection procedure to adapt to the form of noise that is to be suppressed in the wavelet packet domain.

The filtering process has been shown to be effective by both muting and weighting the wavelet packet coefficients of the signal. The weighting technique based on the estimated Q-value for the signal is not the best technique as it is based on Fourier rather than wavelet theory. However, it effectively demonstrates that the filtering process can be improved by using a weighting rather than a crude muting technique. Future work could develop better weighting techniques based on the wavelet packet coefficients themselves and relationships between changes of coefficients between packet scales and the noise to be suppressed.

In this and the previous chapter, we have looked at filtering seismic data, using the one-dimensional wavelet and wavelet packet transforms. These filters are one-dimensional in terms of the transforms they use and that they are trace by trace processes, but two-dimensional in the way that they filter in a two-dimensional transform domain. In the next chapter we will look at the two-dimensional forms of

these transforms and develop these as two dimensional wavelet and wavelet packet transform filters which filter a four-dimensional transform space.

5. Two Dimensional Filtering

5.1 Introduction

A natural extension of filtering seismic signals using one dimensional discrete wavelet and wavelet packet transforms is to filter using two dimensional versions of the transforms. In the one dimensional case we decompose the temporal axis into frequency scales and time translation, and in extending this to two dimensions we additionally decompose the offset axis into wavenumber scale and offset translation. Therefore, the two dimensional transform potentially allows us to filter seismic data in a frequency-time-wavenumber-offset sense. We will refer to this form of filtering as two dimensional as it utilises a two dimensional transform event though we will be filtering a four dimensional transform space.

Two dimensional filters, such as those based on the two dimensional Fourier transform (f-k filters) and the radon transform (τ-p filters), are commonly used in seismic processing. The f-k filter, used for the suppression of ground roll and multiples (Yimaz, 1989), utilises the two dimensional Fourier transform which decomposes a signal using basis functions that have infinite extent in space and time. Therefore, as for one dimensional filters, to implement these filters in a time varying fashion, we must use some form of windowed processing. Filters based on radon transforms are also commonly used in seismic processing for the suppression of ground roll (Yilmaz, 1989) and multiples (Taner, 1980, Verschuur and Berkhout, 1996). The radon transform is not perfectly invertible, errors are introduced in the transform-inverse transform process, but good reconstruction is possible with a least squares inversion. This is not a major limitation and can be considered similar to using a Battle-Lemarié kernel wavelet with the discrete wavelet transform, in that there is an associated error in reconstruction.

To investigate the possible uses of two dimensional wavelet transforms for filtering seismic data we must investigate how they decompose f-k and x-t space. Again, we investigate the influence the choice of basis wavelet on any filtering process.

5.2 The Two Dimensional Wavelet Transform

In this section we will discuss the two-dimensional discrete wavelet transform from the perspective of multiresolution analysis. This is effectively an extension of the analysis performed in chapter 2.4 to two dimensions. The two dimensional transform can be constructed from two one-dimensional multiresolution analyses. In one-dimensional multiresolution analysis we divide the space of square integrable functions (finite energy signals), $L^2(R)$, into closed subspaces V_j. We can use this method to extend the concept to two-dimensions. If we define a two-dimensional subspace $\mathbf{V}_0$ as the tensor product of two one-dimensional subspaces (Daubechies, 1992),

$$\mathbf{V}_0 = V_0 \otimes V_0 \tag{5.1}$$

which is spanned by a two-dimensional function in offset-time, $F(x,t)$,

$$F(x,t) = f(x)g(t); \quad f,g \in V_0 \tag{5.2}$$

and as for the one-dimensional case the subspaces are related by a scaling law such that

$$F \in \mathbf{V}_j \Leftrightarrow F(2^j x, 2^j t) \in \mathbf{V}_0 \tag{5.3}$$

then we have a two-dimensional multiresolution ladder in $L^2(R)$ such that

$$... \mathbf{V}_2 \subset \mathbf{V}_1 \subset \mathbf{V}_0 \subset \mathbf{V}_{-1} \subset \mathbf{V}_{-2} ... \tag{5.4}$$

where the subspaces do not intersect and the union of all the subspaces gives $L^2(R)$ (Daubechies, 1992). This is the two-dimensional equivalent of equation 2-26. Since the one-dimensional function $\phi(t-n), n \in Z$, constitutes an orthonormal basis for V_0, then the product functions

$$\Phi_{0:n_1 n_2}(x,t) = \phi(x-n_1)\phi(t-n_2), \quad n_1, n_2 \in Z \tag{5.5}$$

constitute an orthonormal basis for $\mathbf{V}_0$. Therefore, generalising we find that

$$\Phi_{j:n_1,n_2}(x,t) = \phi_{j,n_1}(x)\phi_{j,n_2}(t)$$

$$= 2^{-j}\Phi(2^{-j}x - n_1, 2^{-j}t - n_2), \quad n_1, n_2 \in Z \tag{5.6}$$

constitute an orthonormal basis for the two dimensional subspace $\mathbf{V}_j$, where j is the scaling index and n_1, n_2 are the translation parameters for the two dimensions. As for the one-dimensional case we define, for each $j \in Z$, the complement space $\mathbf{W}_j$ to be the orthogonal complement in $\mathbf{V}_{j-1}$ of $\mathbf{V}_j$. Therefore we have

$$\mathbf{V}_{j-1} = V_{j-1} \otimes V_{j-1} = (V_j \oplus W_j) \otimes (V_j \oplus W_j)$$

$$= V_j \otimes V_j \oplus [(W_j \otimes V_j) \oplus (V_j \otimes W_j) \oplus (W_j \otimes W_j)]$$

$$= \mathbf{V}_j \oplus \mathbf{W}_j . \qquad (5.7)$$

using equation 2-33. From this we can see that $\mathbf{W}_j$ consists of three components, with the orthonormal bases given by $\psi_{j,n_1}(x)\phi_{j,n_2}(t)$ for $W_j \otimes V_j$, $\phi_{j,n_1}(x)\psi_{j,n_2}(t)$ for $V_j \otimes W_j$, and $\psi_{j,n_1}(x)\psi_{j,n_2}(t)$ for $W_j \otimes W_j$. This leads to the definition of three wavelets for each scale j,

$$\Psi^h(x,t) = \phi(x)\psi(t)$$
$$\Psi^v(x,t) = \psi(x)\phi(t) \qquad (5.8)$$
$$\Psi^d(x,t) = \psi(x)\psi(t)$$

where h, v, and d stand for horizontal, vertical and diagonal respectively. This is related to the form of the wavelets in x-t space and will be explained further when we analyse the decomposition of f-k space by the transform. We will now explain this formation more clearly in terms of the quadrature mirror filter implementation of the transform.

5.3 Transform Implementation

The transform is implemented by the half-band low pass and half-band high pass quadrature mirror filters used in the one-dimensional transform. One iteration of the transform (one pass of the high pass and low pass filters) applies the filters to one dimension, say the offset dimension, followed by one iteration of the transform in the second (say time) dimension as indicated schematically in **Figure 5-1**. This leads to an x-t and f-k representation shown in **Figure 5-2**. The wavelet coefficients are gathered into three sectors and the scaling coefficients into a fourth sector. Each of the three sectors of wavelet coefficients at the first scale (scale 0) $d^{0,v}$, $d^{0,d}$ and $d^{0,h}$, corresponds to one of the three basis wavelets Ψ^v, Ψ^d and Ψ^h derived from the original kernel wavelet. The final sector which contains the scaling coefficients is passed onto the next iteration of the transform. The transform splits f-k space into octave sectors (which are symmetrical about the frequency axis). Each of the three zones of wavelet coefficients in the transform representation contains $\frac{M}{2} x \frac{N}{2}$ wavelet coefficients, when the input data are of dimension MxN. These coefficients cover the input x-t space in a regular grid, so that in x-t space at the lowest scale the dimension of the Heisenberg cell is 2x2 samples. A wavelet coefficient from one sector has the f-k support of the corresponding sector in f-k space. From the relationship between the f-k zones and the corresponding basis wavelets at the first scale we can see that the Ψ^v

wavelet coefficients will contain more steeply dipping (vertical) energy in the x-t domain, that is energy with low wavenumber and low frequencies. The Ψ^d and Ψ^h wavelet coefficients represent diagonally and shallowly dipping (horizontal) energy respectively in the x-t domain, from their associated partitions in f-k space. The regions of support of the basis wavelets in x-t and f-k space (Heisenberg cells) are for illustration purposes, and as for the one-dimensional transforms, the boundaries are fuzzy depending on the kernel wavelet used in the decomposition. The Heisenberg cells in the f-k domain represent the sinc basis function, and in the x-t domain the Haar square wave function.

The full transform, as for the one-dimensional transform, is applied in a cascade fashion on the remaining scaling coefficients, leading to a partitioning f-k space shown in **Figure 5-3**. The arrangement of the coefficients in the output matrix is also shown. As the scale increases, the number of wavelet coefficients at each scale decreases, and correspondingly the support of the basis wavelets in x-t increases (doubles) and f-k support decreases (halves).
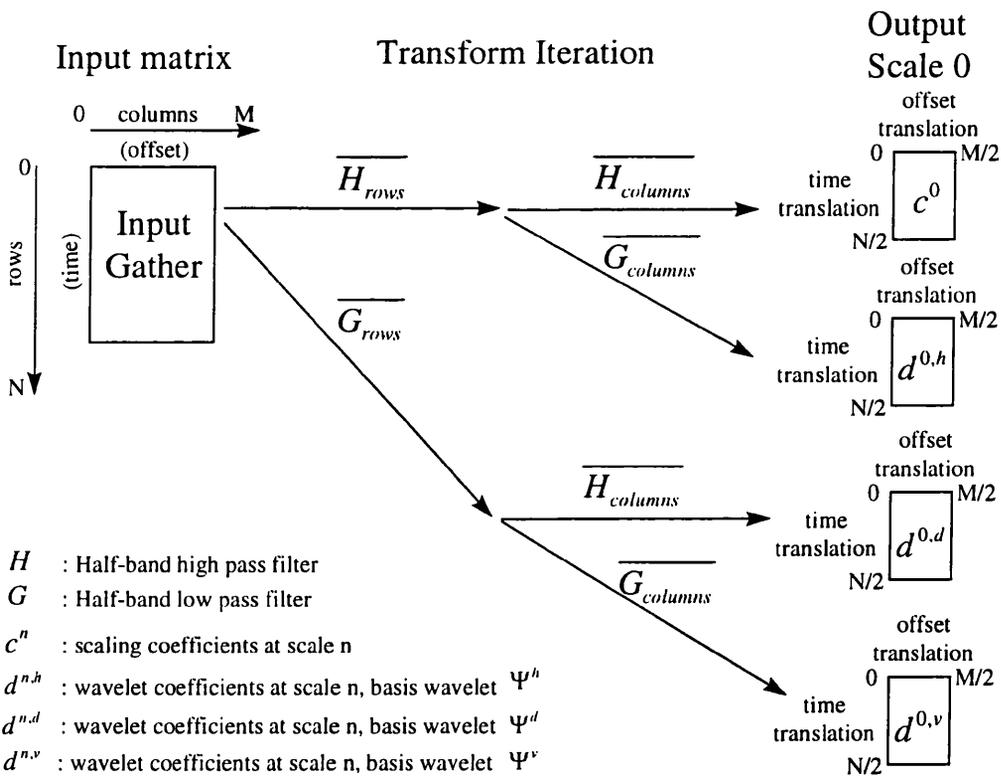
*Figure 5-1 Schematic diagrams showing how the quadrature mirror filters associated with the kernel wavelet are used to decompose a two-dimensional signal in **one iteration** of the wavelet transform.*
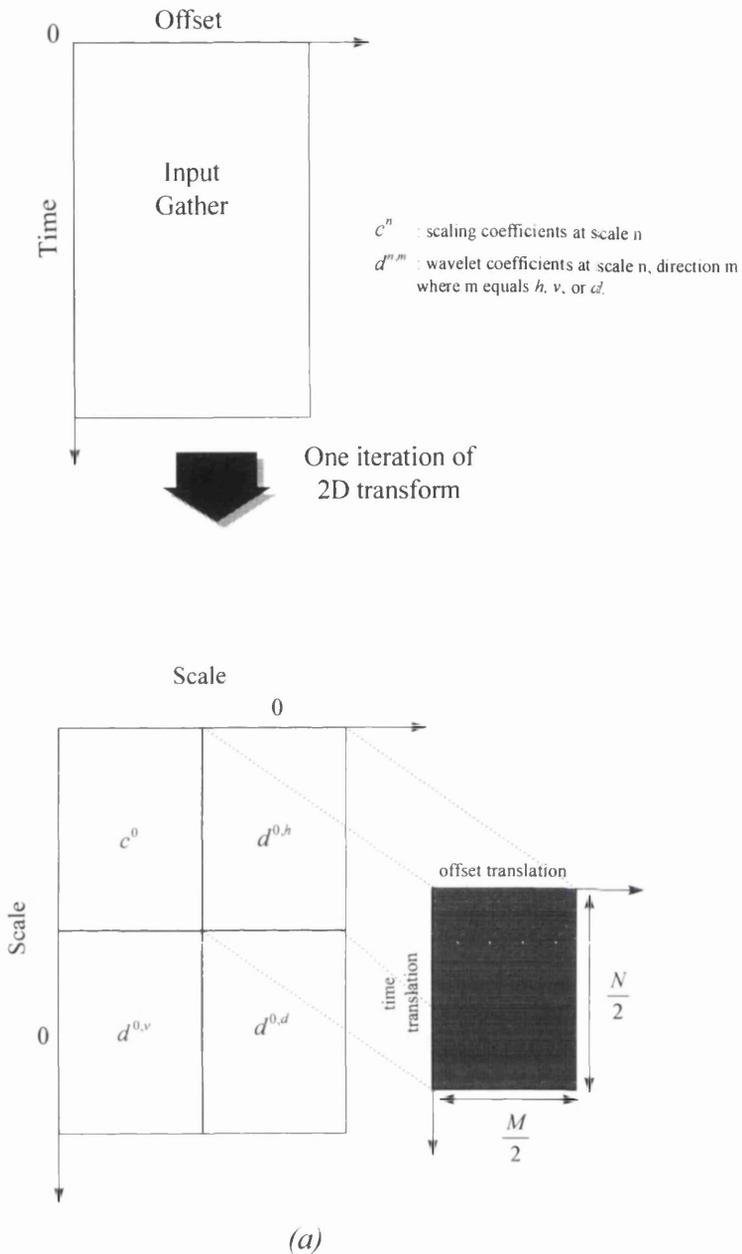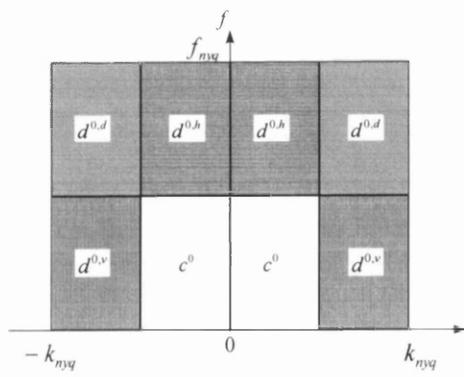
Figure 5-2 Partitioning of the (a) x-t and (b) f-k domain by one iteration of the 2-D discrete wavelet transform. The transform decomposes the x-t domain into three sets of wavelet coefficients corresponding to the three areas of f-k space shown and one set of scaling coefficients that are passed on to the next iteration. F-k space is decomposed into octave segments.
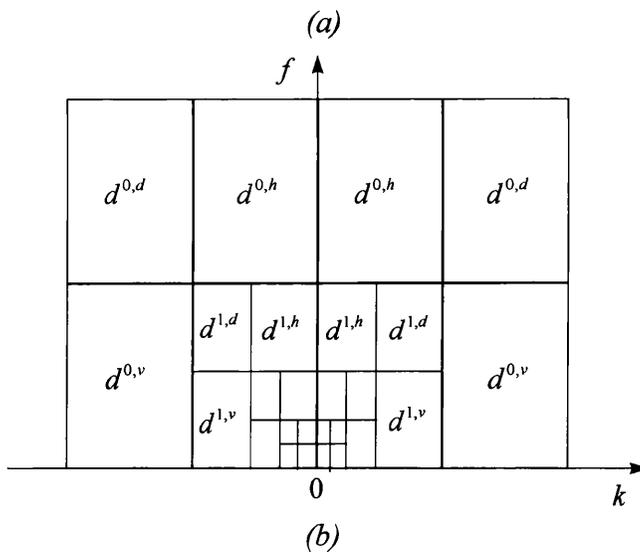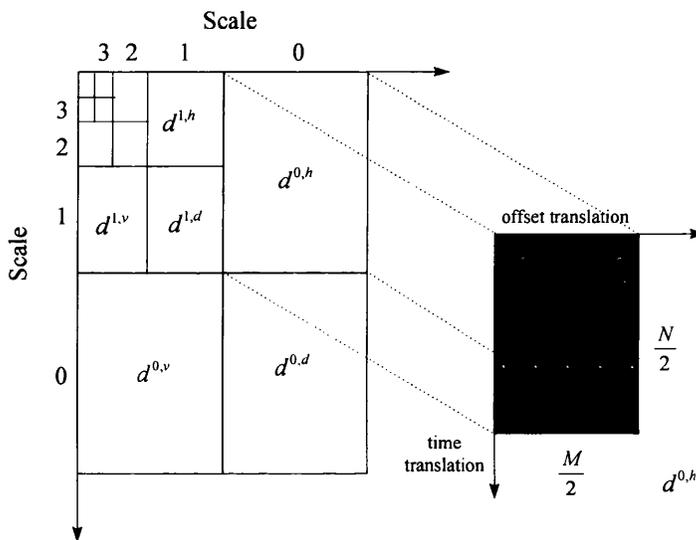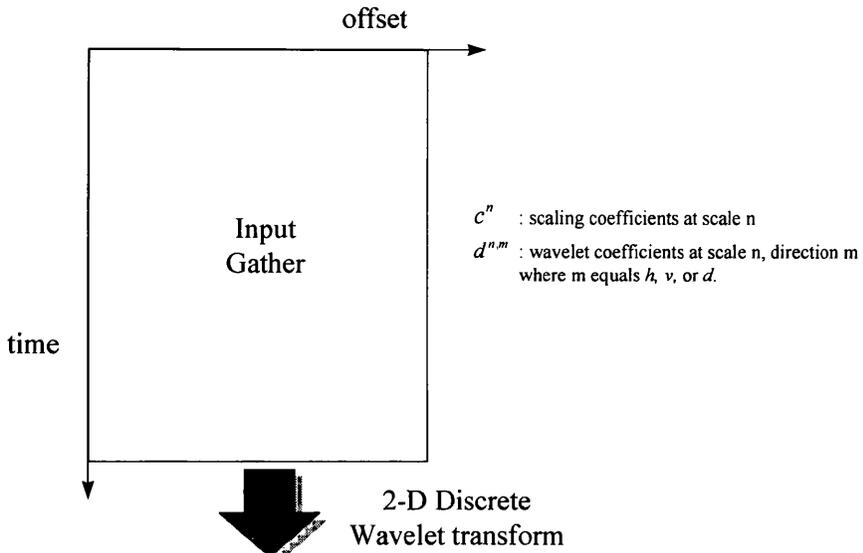
Figure 5-3 Partitioning of (a) x-t space and (b) f-k space by the two-dimensional wavelet transform showing the partitioning of the scaling coefficient areas from the previous figure by the cascaded application of the transform process.

We can see that the three wavelet functions defined in equation 4.9 are derived from combinations of the quadrature mirror filters used in the one-dimensional transform applied in the different dimensions. Examples of these basis functions for the cubic spline Battle-Lemarié kernel quadrature mirror filters are shown in **Figure 5-4**. **Figure 5-5** shows a seismic shot record and the corresponding two dimensional wavelet transform using these basis wavelets. From this figure and **Figure 5-2** to **Figure 5-4** several interesting points arise.

(a)   (b)   (c)
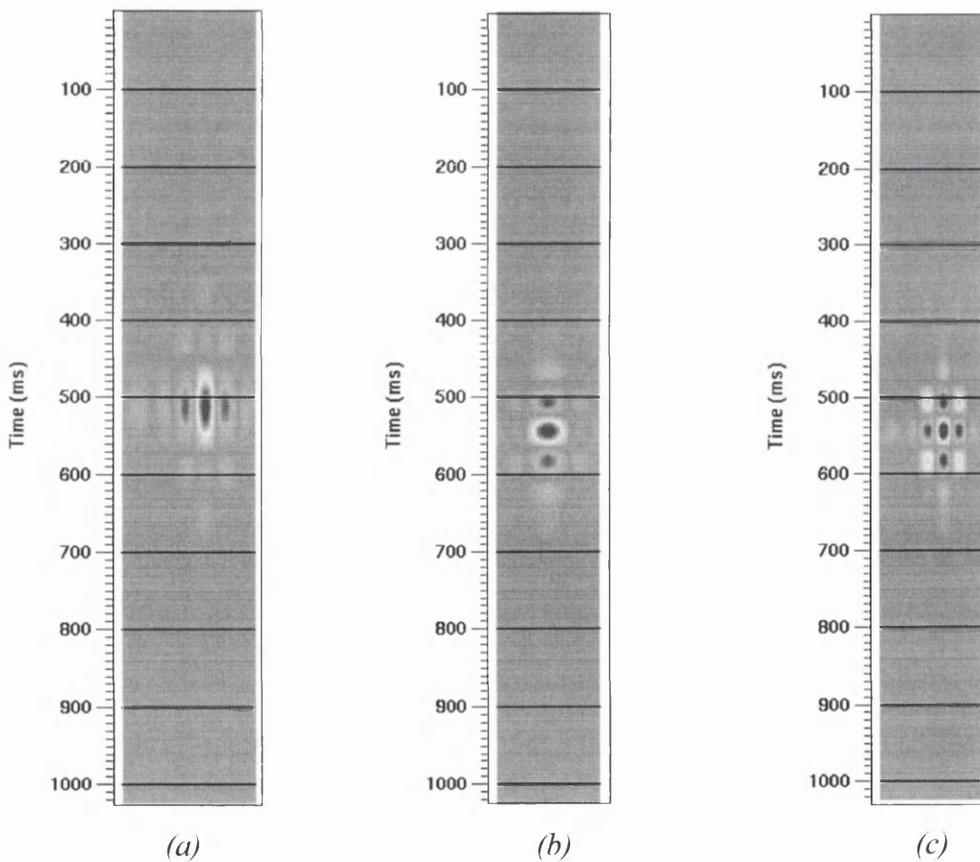
Figure 5-4 Two-dimensional wavelet basis functions built from cubic spline Battle-Lemarié kernel wavelets. The three wavelets are the basis for one scale (n) orientated in the (a) vertical direction, $d^{n,v}$, (b) horizontal direction, $d^{n,h}$, and (c) diagonal direction, $d^{n,d}$.
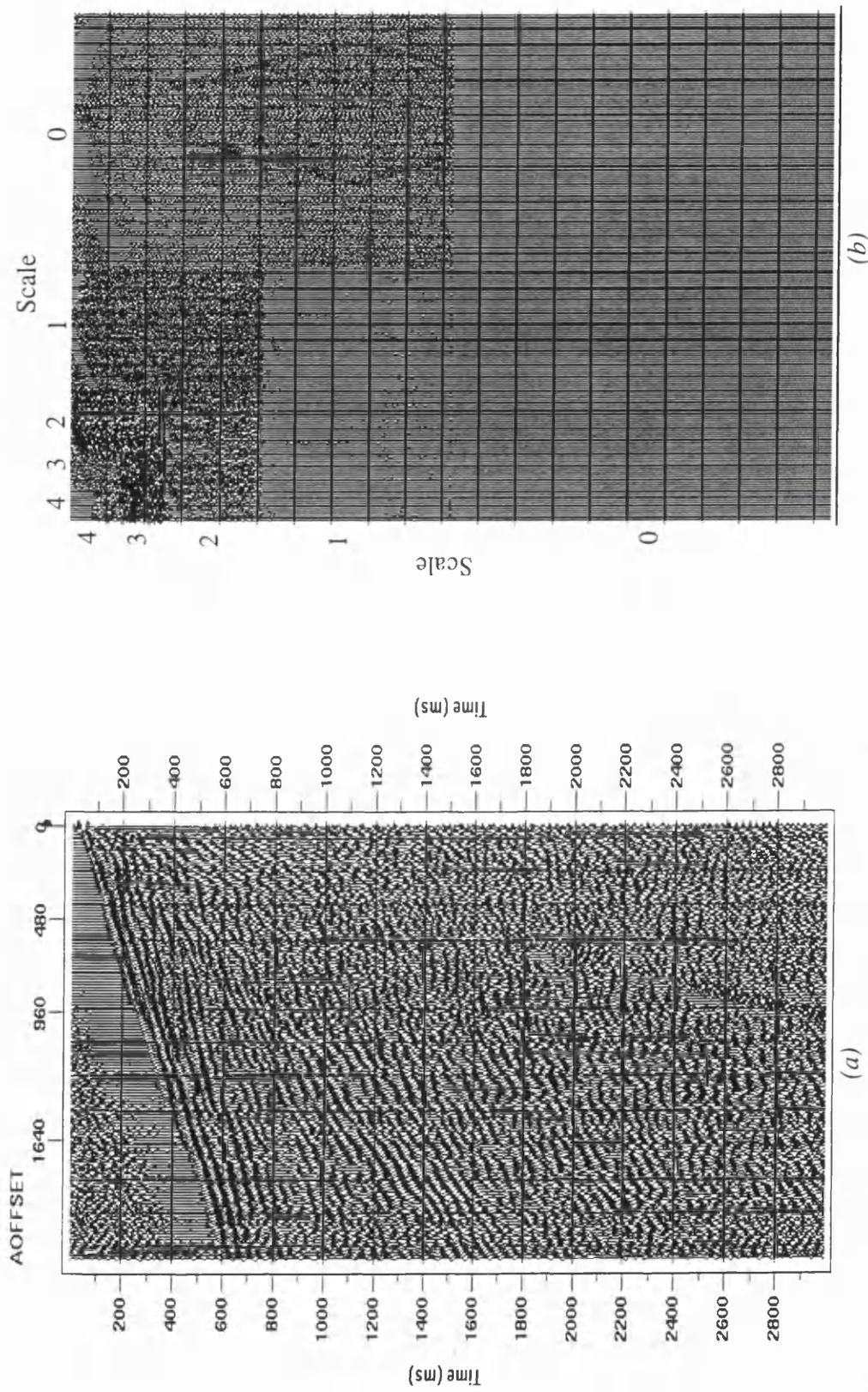
Figure 5-5 (a) A common shot record and (b) the corresponding two dimensional wavelet transform, where the wavelet coefficients are stored as depicted in Figure 5-3.

Chapter 5: Two Dimensional Filtering

Firstly, as we mentioned previously, the partitioning of f-k space by the basis wavelets is symmetrical about the frequency axis. It is also symmetrical about the wavenumber axis, but this is not typically shown using seismic f-k plots. This may present problems when attempting to filter spatially aliased data where the spatially aliased signal lies in one sector in f-k space, with the mirror sector (from symmetry) containing signal we wish to preserve.

Secondly, seismic data tends not to be equidimensional, that is, the number of samples in the time domain does not equal the number of samples in the offset domain. In the time domain, there are usually thousands of samples, whilst in the offset domain there tends to be tens or hundreds of traces. In seismic acquisition, the sampling of the seismic wave field is denser in time that in space. The two dimensional discrete wavelet transform assumes that the sample interval in time is the same as in space. Therefore, when we display the two-dimensional wavelet transform, as in **Figure 5-5**, we are actually showing a distorted view as the offset dimension is undersampled compared to the time. This explains why there is very little energy apparent in the $d^{0,h}$ and $d^{0,d}$ tiles of the transform. If we looked at the input shot record with the perspective of the transform, that is with equal spacing of horizontal and vertical samples, the shot record would look as if it contained energy dipping almost vertically.

Thirdly, as for the one-dimensional transform, the two dimensional transform array lengths need to be an integer power of two in length. Therefore we have to pad out the signal in offset and time in the same manner as for the one-dimensional case. When implementing the transform process we must define how the edges of the signal are handled and how the signal is padded, if this is required. The techniques used for the two dimensional transform are identical to those discussed in chapter 3. In **Figure 5-5** the signal has been padded out to the next power of two number traces/samples by mirroring the data about the last trace/sample. The transform boundary condition is periodic.

Fourthly, the input gather (a matrix of samples) is not equidimensional. Therefore, the number of scales possible when decomposing one dimension of the gather may not be the same as for the other dimension. Therefore, when applying the full transform, we cascade the process until the lowest possible scale is reached in the dimension with the lowest number of possible scales.

Finally, the representation of the two-dimensional transform shown in **Figure 5-5** is particularly cumbersome. For the one-dimensional transform we had the scalogram

which gave an efficient representation of the coefficients and their relationships in frequency and time. However with the two dimensional wavelet transform, there are four variables, offset scale, temporal scale, temporal translation and offset translation, which makes visualisation of the transform difficult. In this work we will use a combination of the x-t, *f-k* and two dimensional wavelet representations to display, analyse and design filters for the data. The coefficients will be displayed as samples in two-dimensional wavelet space as shown in **Figure 5-3** with each sample for a given scale and orientation (horizontal, vertical or diagonal) represents a given area in *f-k* space. In this representation there are two sets of Heisenberg cells, one associated with the wavelet coefficient in the x-t domain and another in the *f-k* domain. As mentioned previously, these boundaries are used to indicate the support of the basis wavelets in these two domains and the true supports vary depending on choice of kernel wavelet.

5.4 Kernel Wavelet Selection

Figure 5-6 shows examples of basis wavelets at one scale with different orientations and the corresponding *f-k* spectra. Superimposed on this diagram are the corresponding Heisenberg cells. From this we can see that the boundaries represented by the Heisenberg cells are not exact, as for the one-dimensional case. The criteria we developed in chapter 3 for selecting a suitable basis wavelet for filtering seismic data using the one dimensional transform are equally applicable for the two dimensional version. The flat frequency response over the main *f-k* support of the basis function and fast fall off at scale boundaries are desirable as is symmetry of the kernel wavelet. However, with the two dimensional transform and seismic signals there is an added constraint when we extend the filtering process to two dimensions.

Seismic data typically has wide dynamic range, that is the amplitude range of signals in a common shot record can be particularly large, especially from trace to trace. When we selected the Battle-Lemarié kernel wavelets as the most suitable kernel wavelet for one dimensional filtering, we assumed that the error in reconstruction associated with these kernel wavelets (which depended on the number of coefficients used) was of minor consequence compared to the benefits gained from the other properties of this wavelet. However, when using these kernel wavelets in the two-dimensional transform, where there can be large trace to trace amplitude variations, this error in reconstruction is more significant.

Wavelet transforms have been developed mainly in the image processing community which deal with digital images whose samples normally have a range of values from 0 to 255. Therefore this problem of reconstruction error does not create a problem when transforming one or two dimensional images. For seismic data, the one-dimensional case, we typically used twenty four Battle Lemarié cubic spline coefficients and forty eight quintic spline coefficients to apply the transform. In two dimensions this would not be acceptable for a typical shot record, unless some form of gain and/or trace equalisation had been previously applied to the shot record. For a shot record, with no gain applied and large trace to trace amplitude variations, one-hundred and twenty four coefficients are typically required for the quintic spline wavelet to produce an acceptable reconstruction with minimal error. When comparing this to the Daubechies wavelets which will exactly reconstruct without any need to expand the number of kernel wavelet coefficients, we can observe that using Battle Lemarié wavelets will increase any processing time considerably. In terms of processing cost, the 20 coefficient Daubechies wavelet will apply the transform process in one-sixth of the time taken by the quintic spline Battle-Lemarié wavelet with 124 coefficients.

Figure 5-7 shows the 2-D reconstruction error manifesting itself on a common shot record using the 24 coefficient Battle-Lemarié cubic spline wavelet and how this is minimised by using a greater number of coefficients. Rather than increase the number of coefficients we use, we can apply a gain function to the traces before transformation, which can later be removed after reconstruction. This gain can be in the form of an AGC. The application and removal of the gain to the data is far less time consuming than the use of the extra coefficients and so, this leads to a reduction in processing time. When filtering signals use two-dimensional wavelet transforms, we will use this technique rather than increasing the number of coefficients used. This means that when comparing the basis wavelets, the extra cost in using the Battle-Lemarié wavelets is the two additional gain application/removal steps. This gain technique is also commonly used in seismic processing streams when filtering data using *f-k* filters where the data contains large amplitude variations.
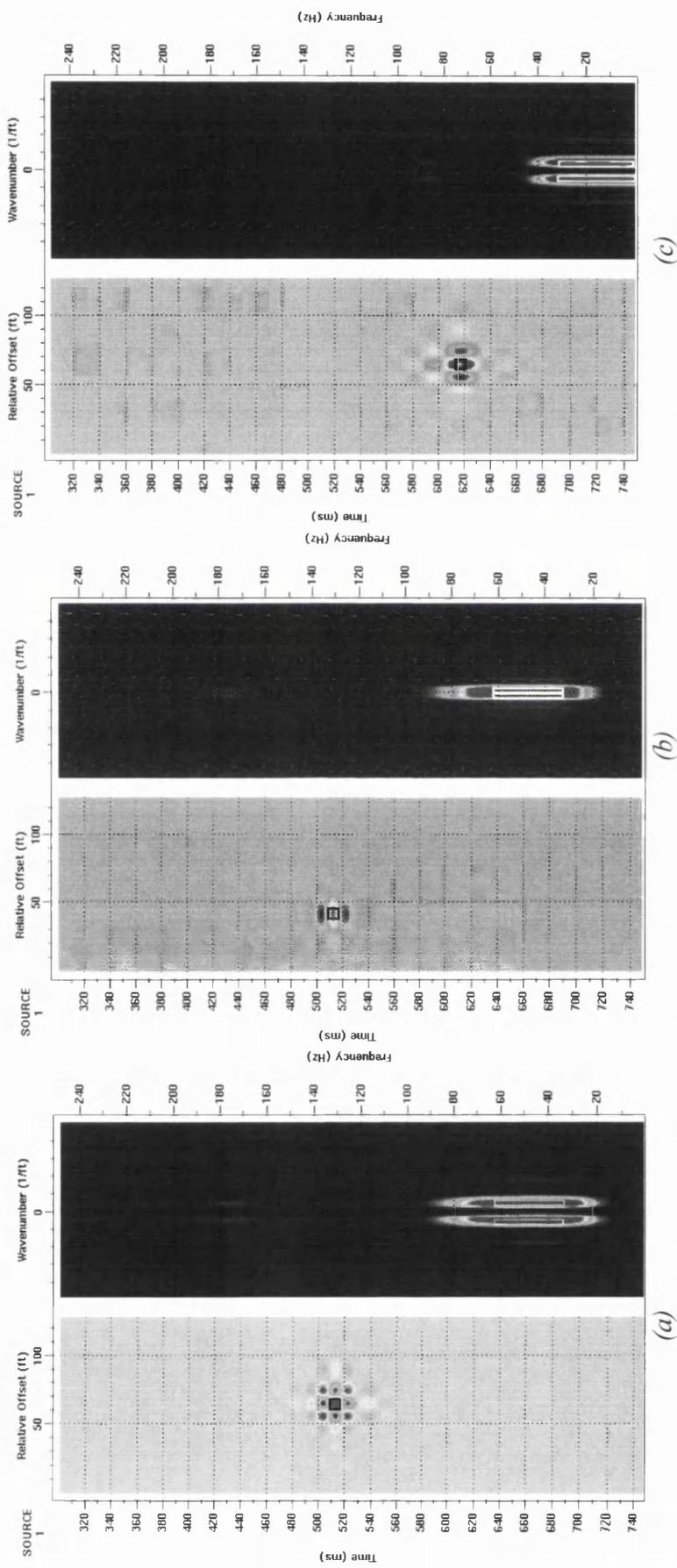
Figure 5-6 Two-dimensional basis wavelets at one scale (n) for the 20 coefficient least asymmetric Daubechies kernel wavelet and their corresponding f-k representations and x-t (black) and f-k (white) Heisenberg cells (a) $d^{n,d}$, (b) $d^{n,h}$ and (c) $d^{n,v}$.
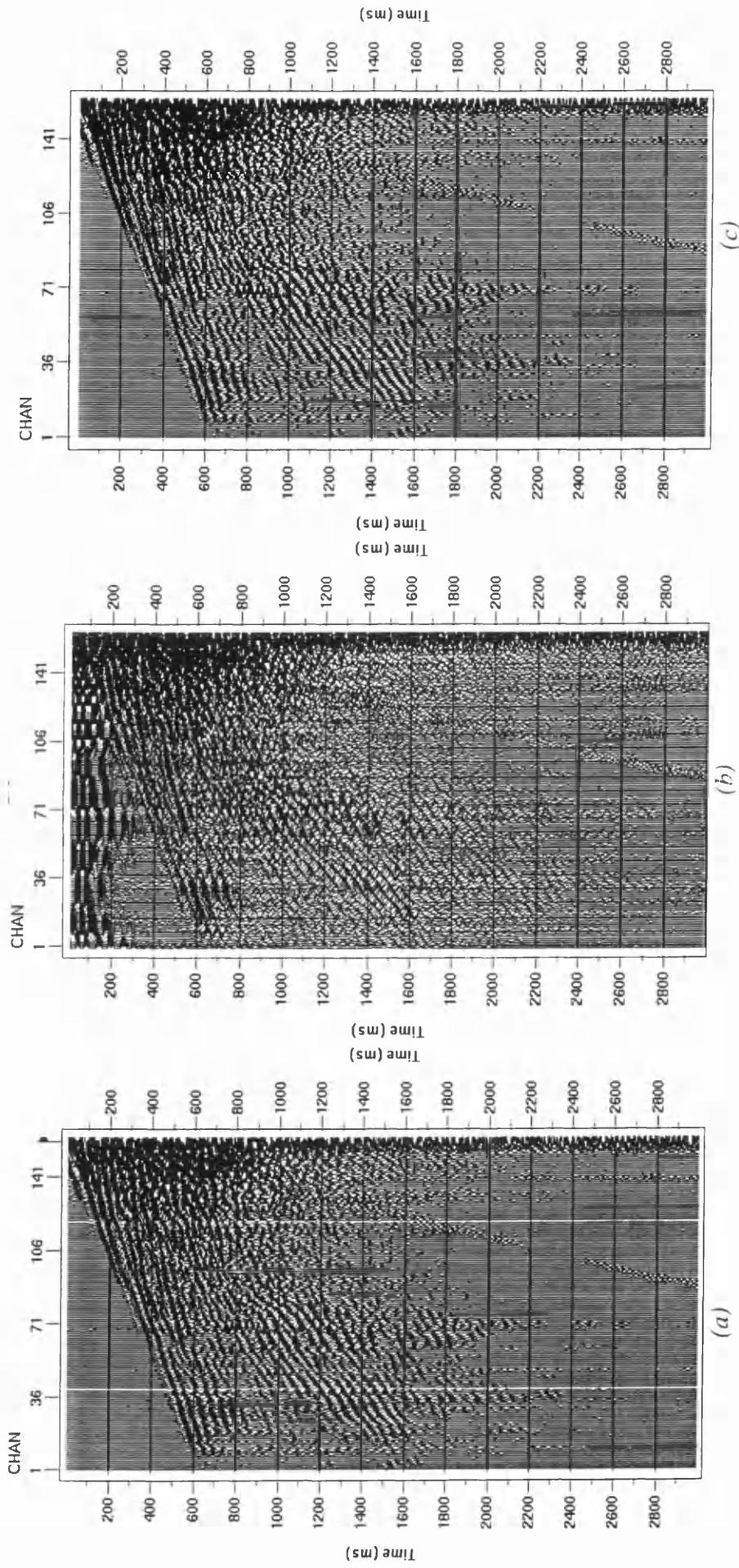
Figure 5-7 (a) A common shot record, (b) after application of the forward and inverse two dimensional wavelet transform using 24 Battle-Lemarié cubic spline kernel wavelet coefficients in the transform processes and (c) using 124 Battle-Lemarié cubic spline kernel wavelet coefficients.

Chapter 5: Two Dimensional Filtering

5.5 Filtering with the Two-Dimensional Wavelet Transform

The two-dimensional wavelet transform potentially allows the filtering of seismic data in the four dimensions of the transform space. Analysing a given signal in the two dimensional wavelet domain, we can select coefficients that can be muted out and the inverse transform applied. The poor representation of the two-dimensional transform means that it is difficult to visualise the coefficients to be filtered for a given signal at higher scales. However, with the added knowledge of the range of the corresponding Heisenberg cells in x-t and *f-k* space we can specify the coefficients to be removed. This process is shown in **Figure 5-8** to **Figure 5-10**. **Figure 5-8** shows a simple synthetic signal containing four reflected events, a refracted event, and three low velocity, low frequency events, described in **Table 5-1**. **Figure 5-9** shows the two dimensional wavelet transform of the gather using the cubic spline Battle-Lemarié kernel wavelet. We can see that the majority of the low velocity signal is contained in the vertical basis wavelet sector of scales 0-2, whilst reflected and refracted signals are contained in these and other scales. This is as expected when we look at the corresponding *f-k* regions for these sectors. On the transform display, within the wavelet coefficient sectors that contain the low velocity signal, we have selected areas (between the crosses on each trace) to mute the signal in an attempt to suppress the low velocity signals. We do not mute the entire sector, only the portion of the wavelet coefficients that contain low velocity signal. This localises the filter in time and offset. **Figure 5-10** shows the gather after filtering in the wavelet domain and the application of the inverse transform. We can see that apart from some edge effects associated with the filtering process, the low velocity signals have been successfully suppressed.

Event	Velocity (m/s)	Wavelet
Refractor	4500	50 Hz Ricker wavelet
Reflector 1	3500	30 Hz Ricker wavelet
Reflector 2	3700	30 Hz Ricker wavelet
Reflector 3	4500	30 Hz Ricker wavelet
Reflector 4	5500	30 Hz Ricker wavelet
Low velocity 1	1600	4-6-15-35 Hz corner frequencies
Low velocity 2	1000	4-6-15-20 Hz corner frequencies
Low velocity 3	600	4-6-15-20 Hz corner frequencies

Table 5-1 Velocities and seismic wavelet frequiencies used to form the synthetic shot gather shown in Figure 5-8.
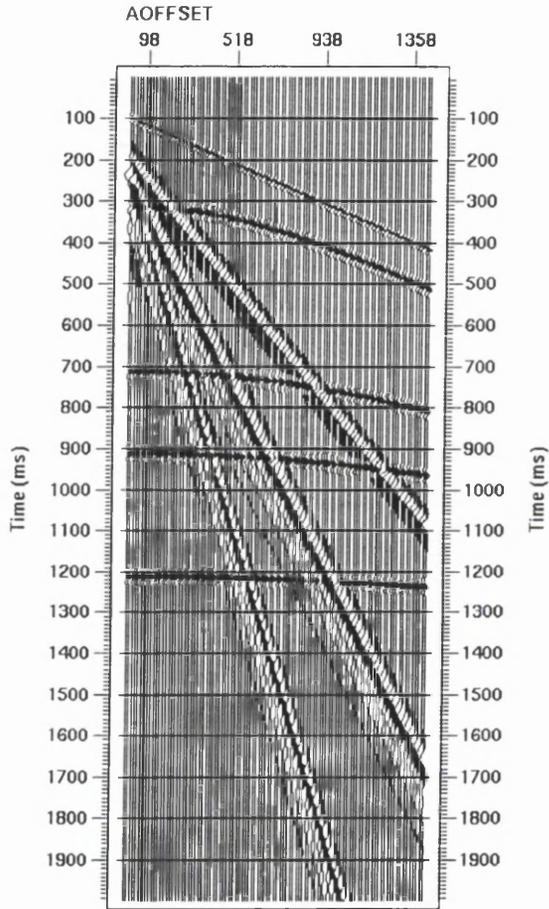
Figure 5-8 A synthetic shot gather containing four reflected events, one refracted event and three low velocity, low frequency events which we will attempt to suppress using the two dimensional wavelet transform.
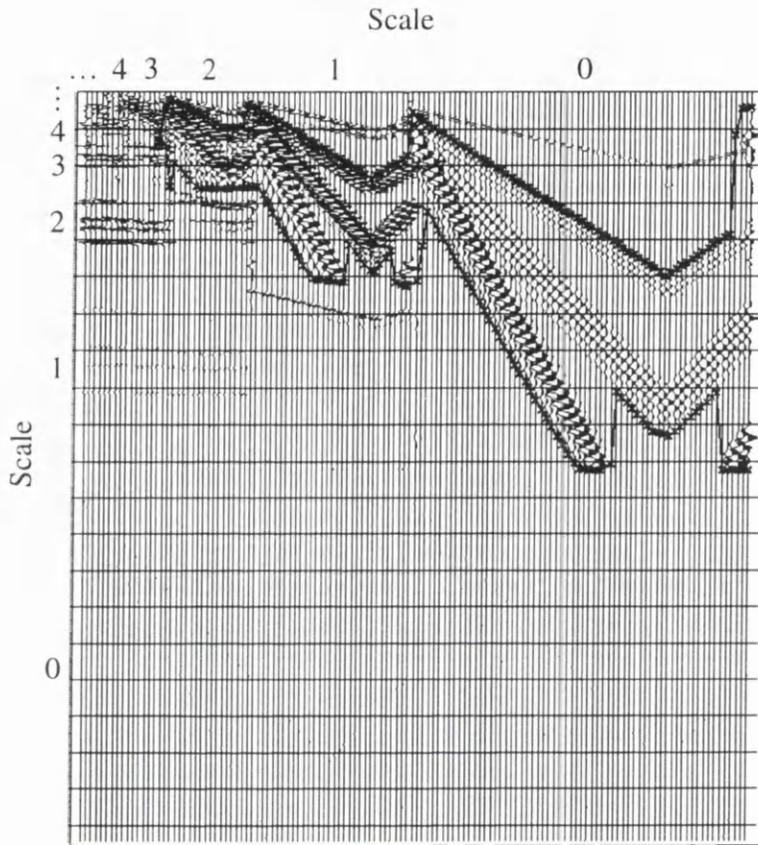
Figure 5-9 The two dimensional discrete wavelet transform of Figure 5-9 using the cubic spline Battle Lemarié kernel wavelet. The record has been padded by mirroring the signal about the last sample in the offset and time domains. Large wavelet coefficients due to wrap around effects can be seen clearly at the right hand side of the transform display.
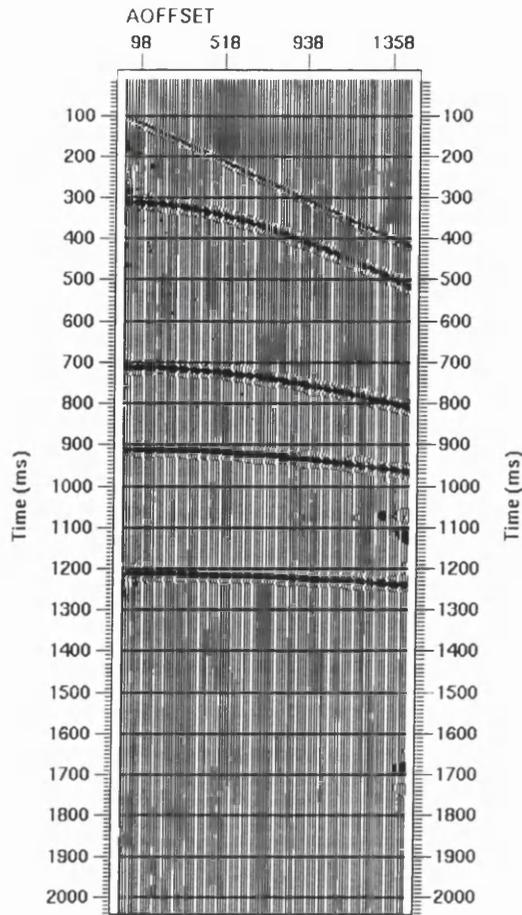
Figure 5-10 The shot record shown in Figure 5-8 after filtering using the two dimensional discrete wavelet transform. The low velocity signal has been successfully suppressed apart from the residual edge effects associated with the filtering process.

In seismic processing, two-dimensional filtering techniques are principally used for the suppression of low velocity noise from seismic common shot records. The appearance of low velocity noise in the f-k domain is shown schematically in **Figure 5-11** where we have also superimposed the partitioning of the domain by the two-dimensional wavelet transform. In the previous synthetic example we had good separation of the reflections and refractions from the low velocity signal in f-k space. In field data, this degree of signal and noise separation is unlikely and so the octave splitting of f-k space by the two-dimensional discrete wavelet transform may be too coarse for suppression of low velocity noise without seriously deteriorating the signal we wish to preserve, particularly at lower frequencies. This problem is demonstrated in **Figure 5-12** to **Figure 5-14**. **Figure 5-12** shows a common shot record contaminated by ground roll and the corresponding f-k spectrum after the application of an AGC to enhance weak signals. The record was acquired using an explosive source and clusters of geophones at each recording station with a 3 metre receiver group interval. **Figure 5-13** shows the two-dimensional wavelet transform of the record using the cubic spline

Battle-Lemarié kernel wavelet, from which we can see the areas to be filtered in the wavelet domain in an attempt to suppress the ground roll. The resultant filtered record is shown in **Figure 5-14**, from which we can see that some ground roll remains. In addition, the areas that have been filtered have a wormy appearance, a result of signal as well as noise being removed, leaving distorted signal behind. Using alternative kernel wavelets does not improve on this result. The poor filtering performance is a direct result of too coarse a sampling in f-k space by the two dimensional wavelet transform. To circumvent this coarse sampling problem we must apply a shift to the data, moving the signal away from the noise in f-k space. To do this we will apply a normal moveout to the signal before filtering, which we will subsequently remove after filtering.

The use of a normal moveout (NMO) corresponding to reflected energy before filtering rotates reflected energy towards the frequency axis. Ground roll energy is also rotated, but not to the same extent. The purpose of the moveout is to rotate the reflected energy and ground roll into separate Heisenberg cells in f-k space so that when we have energy overlap in the wavelet domain, any filtering process minimises distortion of the primary reflectors. This result of this process is shown in **Figure 5-15**. A constant velocity NMO has been applied to the signal in the common shot domain before the wavelet transformation process and removed after the inverse transform. As before, the filter was designed by the selection of filter zones in the wavelet domain. The result of this filtering process is a great improvement from the previous filtered record, with more noise removed, more reflected signal apparent and less worminess on the filtered record.

Rather than apply a normal moveout, we would like to use a technique that decomposes f-k space into smaller Heisenberg cells in the area we wish to filter. Although this would lead to correspondingly larger wavelets in the x-t domain, we would like a decomposition that allows more flexibility in terms of f-k and x-t support of the basis wavelets. Therefore, we will extend the use of the two-dimensional filtering process to the wavelet packet transform in two dimensions. As for the one-dimensional transform the wavelet packet transform theoretically allows higher resolution in f-k space in areas where the wavelet transform has poor f-k resolution and in general allows a more flexible representation in the frequency-wavenumber-offset-time domain.
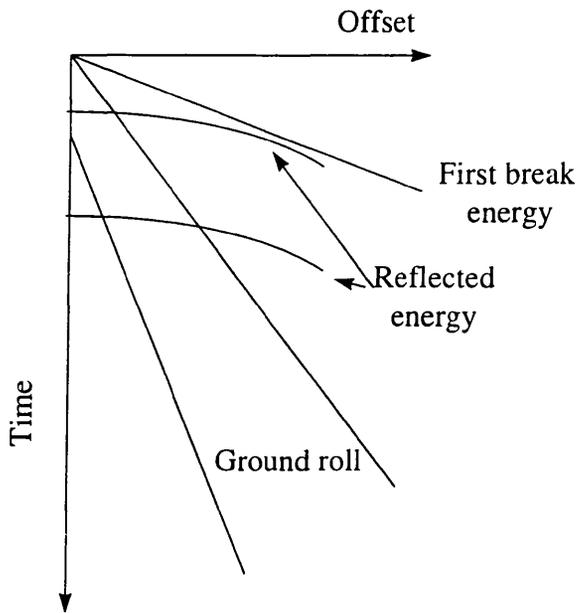
(a)
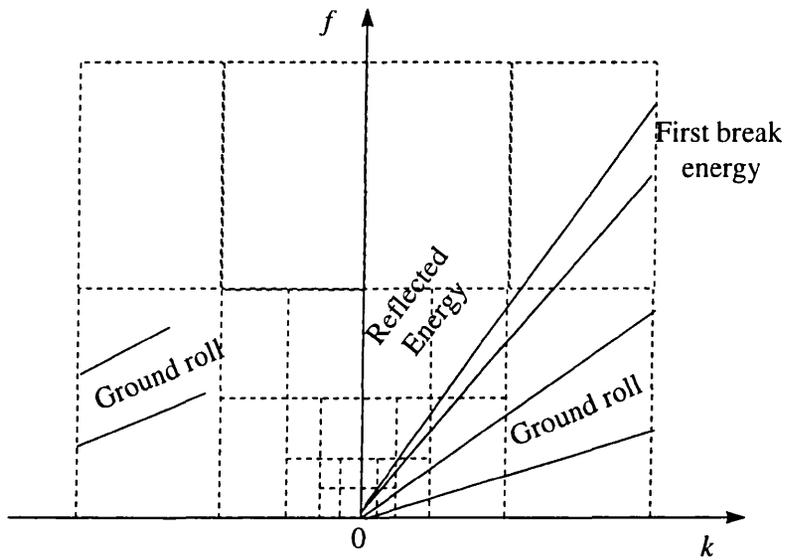
(b)

Figure 5-11 The typical positioning of signal and noise in (a) the x-t domain and (b) the f-k domain and comparing the partition of the f-k domain by the 2-D wavelet transform. The octave band splitting may be too coarse a partitioning of the f-k domain for filtering at lower frequencies and wavenumbers.
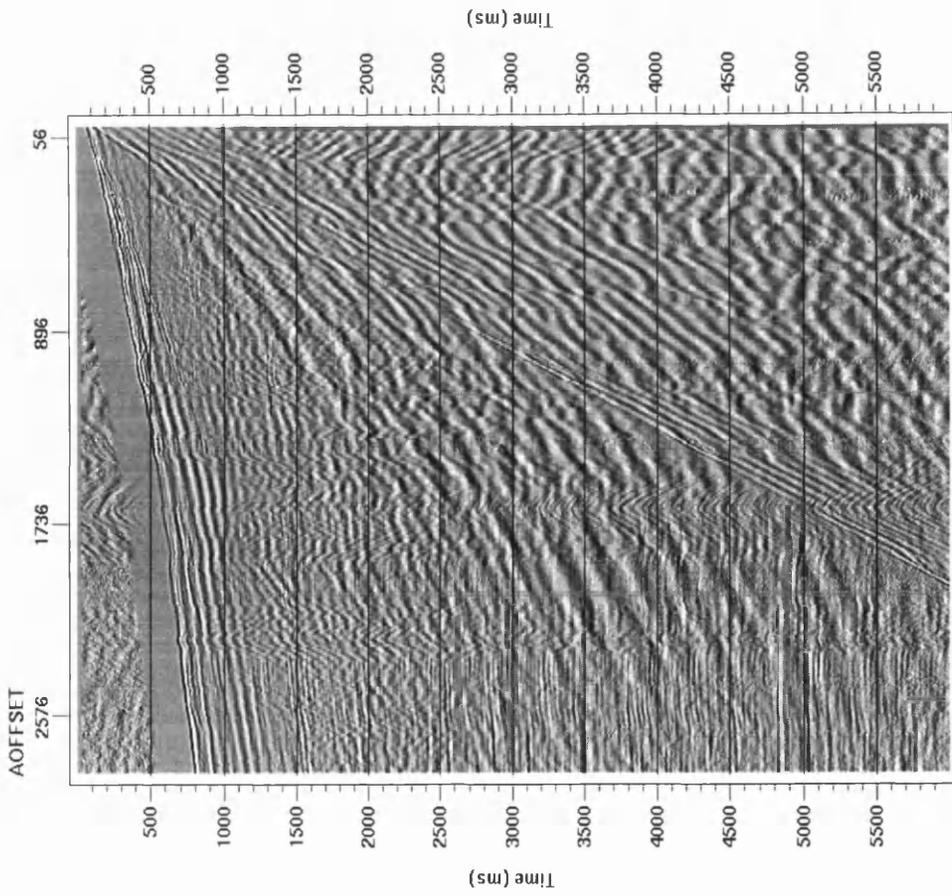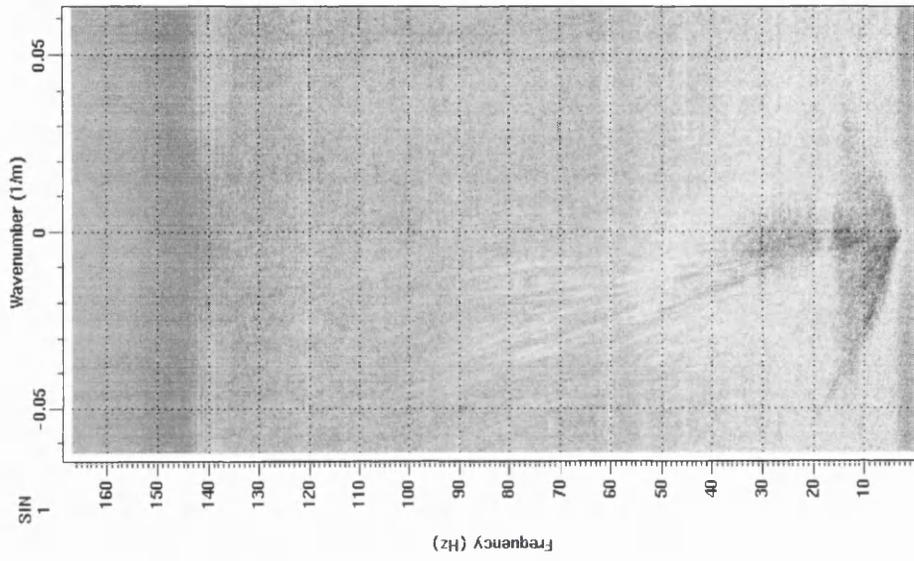
Figure 5-12 (a) A common shot record from a land based seismic survey contaminated by low velocity ground roll and (b) the corresponding f-k spectrum.

Chapter 5: Two Dimensional Filtering
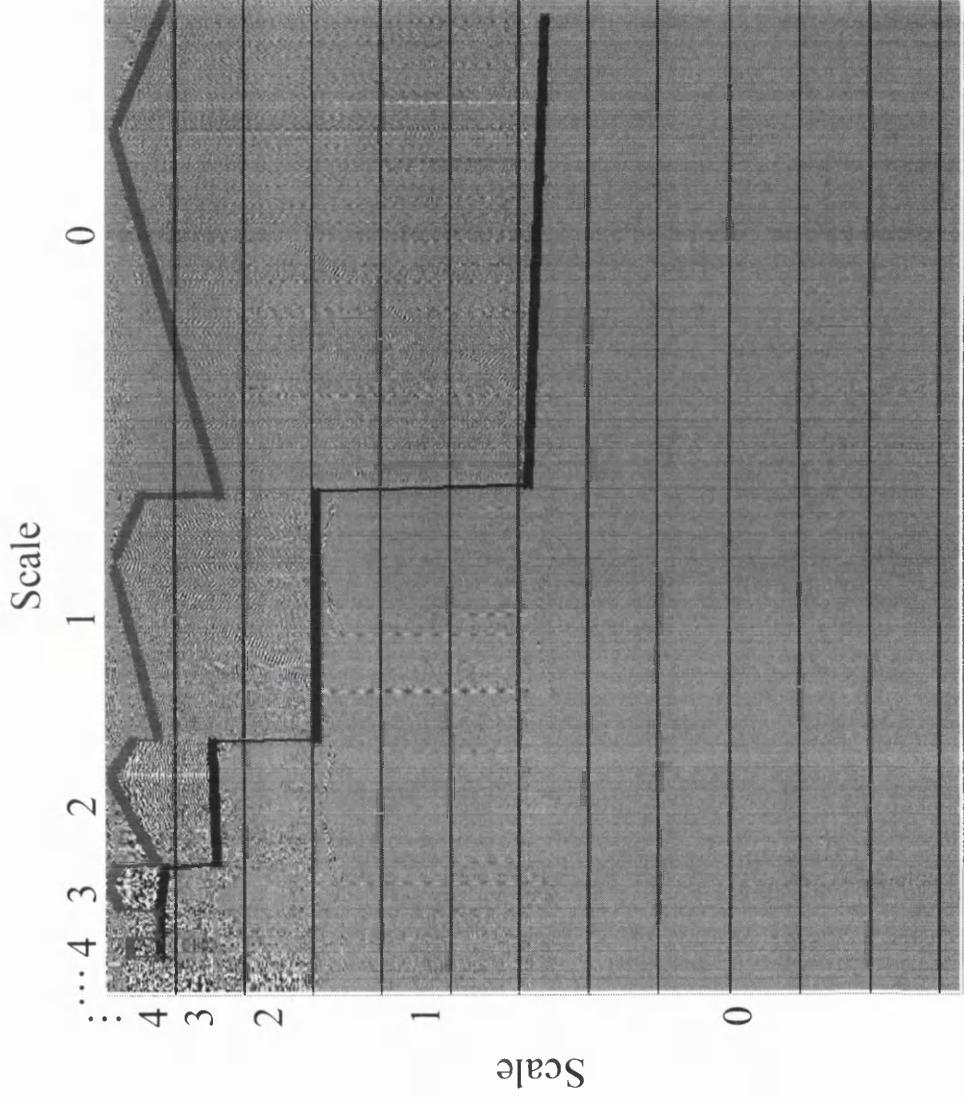
Figure 5-13 The two dimensional wavelet transform of the shot record in Figure 5-12(a) using the quintic spline Battle-Lemarié kernel wavelet. The filtered zones are bounded by the two thick lines marked on the figure.
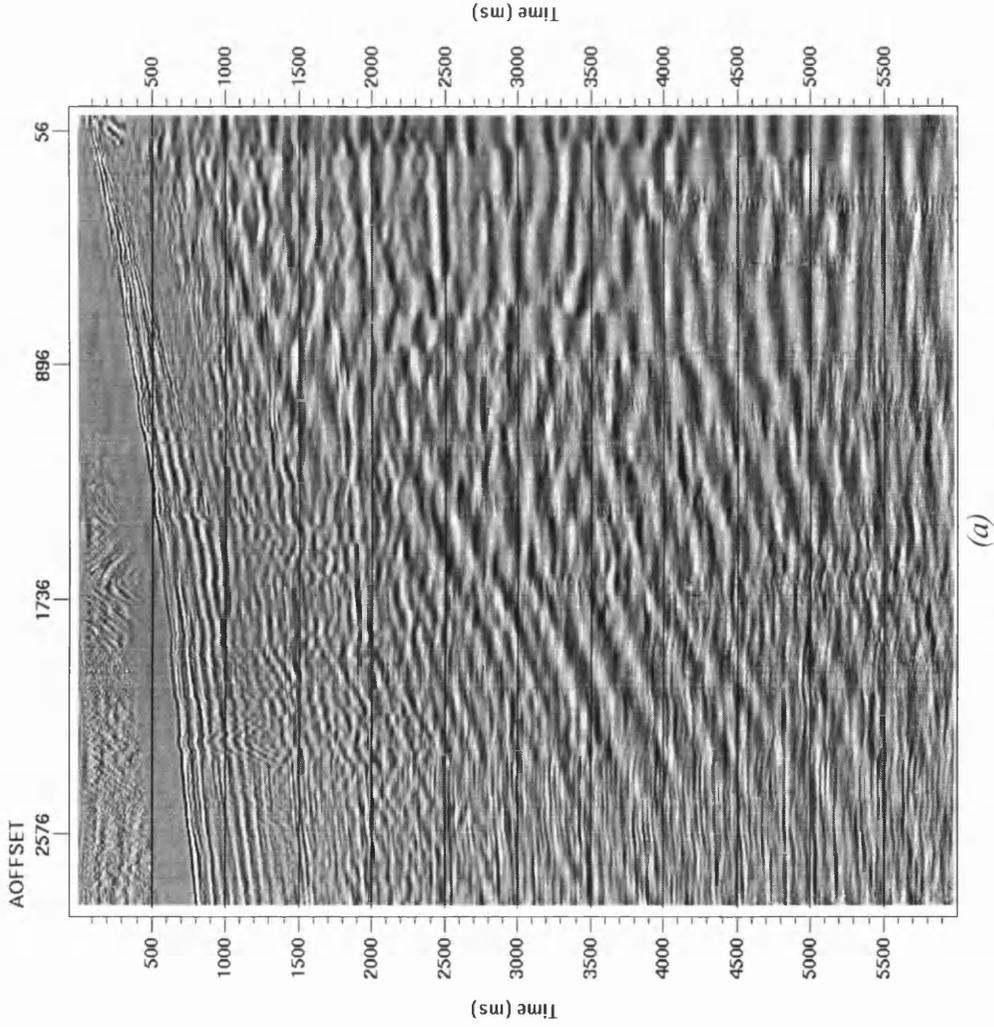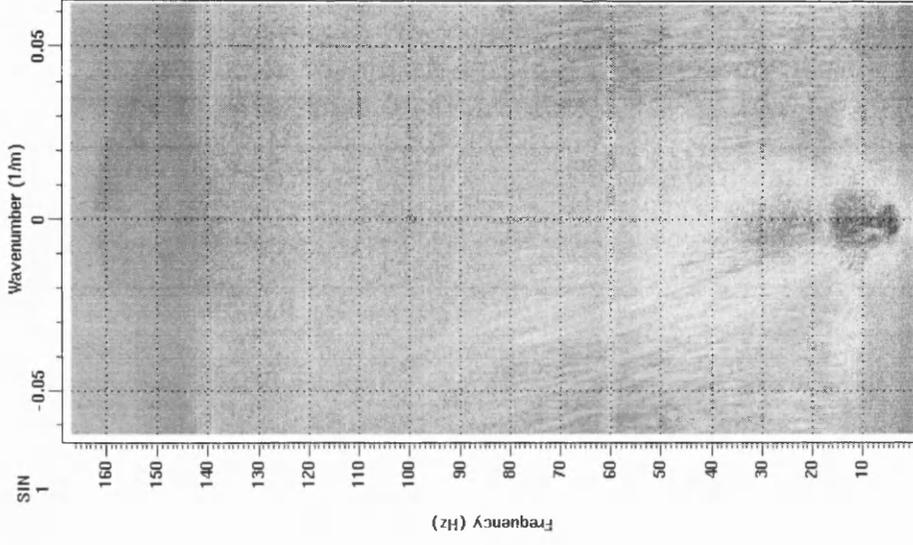
Figure 5-14 (a) The common shot record in Figure 5-12(a) after filtering using a two-dimensional wavelet transform and (b) the associated f-k spectrum.

Chapter 5: Two Dimensional Filtering
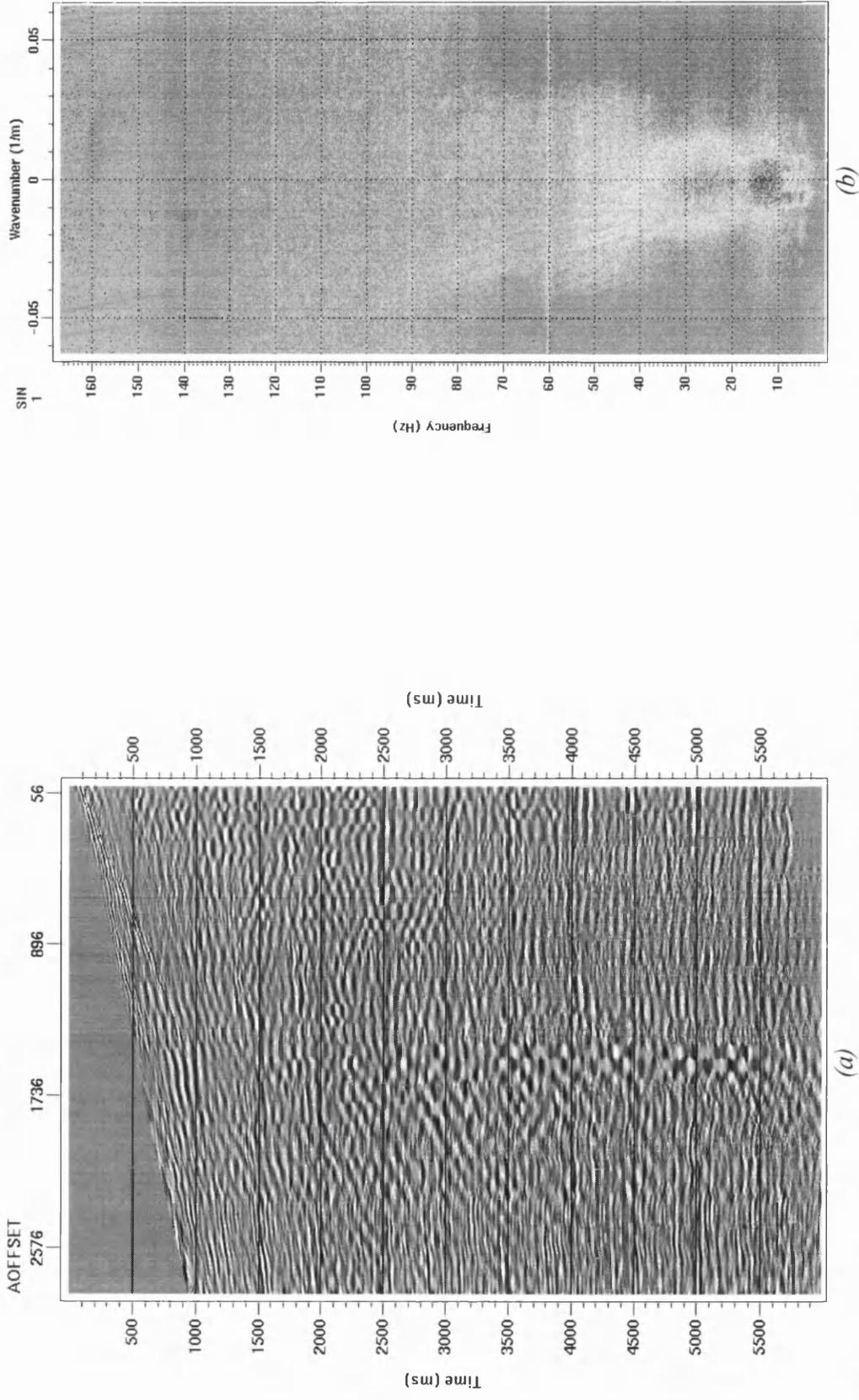
Figure 5-15 (a)The common shot record shown in Figure 5-12(a) after filtering with the two dimensional wavelet transform with the application of a normal moveout before filtering, and removed after. (b) The corresponding f-k spectrum.

5.6 Wavelet Packets in Two-Dimensions

As for the wavelet transform, the wavelet packet transform in two dimensions is a direct extension of the one-dimensional application. In one-dimension, the transform applies the quadrature mirror filters to the wavelet coefficients as well as the scaling coefficients at each iteration of the transform. In two dimensions, the same process is applied, with the exception that it is applied in two dimensions. One iteration of the wavelet packet transform in two dimensions leads to the same partitioning of x-t and *f-k* space shown in **Figure 5-2**. The second iteration applies the quadrature mirror filters to the sectors containing both the wavelet coefficients and the scaling coefficients in both dimensions leading to the partitioning of x-t and *f-k* space as shown schematically in **Figure 5-16**. Each successive iteration splits a sector of coefficients into four smaller sectors.

Therefore, after one iteration we have four basis wavelet packets (identical to the three basis wavelets plus the scaling function) which in the z-domain are

$$
\begin{aligned}
W_0^1(z_1,z_2) &= G(z_1)G(z_2) \\
W_1^1(z_1,z_2) &= G(z_1)H(z_2) \\
W_2^1(z_1,z_2) &= H(z_1)G(z_2) \\
W_3^1(z_1,z_2) &= H(z_1)H(z_2)
\end{aligned}
\tag{5.9}
$$

where we use the quadrature mirror filters $H(z)$ and $G(z)$, and where z_1 and z_2 represent the z-domain for offset and time respectively. After two iterations we have sixteen basis wavelet packets

$$
\begin{aligned}
W_0^2(z_1,z_2) &= G(z_1)G(z_2)G(z_1^2)G(z_2^2) & W_4^2(z_1,z_2) &= G(z_1)H(z_2)G(z_1^2)G(z_2^2) \\
W_1^2(z_1,z_2) &= G(z_1)G(z_2)G(z_1^2)H(z_2^2) & W_5^2(z_1,z_2) &= G(z_1)H(z_2)G(z_1^2)H(z_2^2) \\
W_2^2(z_1,z_2) &= G(z_1)G(z_2)H(z_1^2)G(z_2^2) & W_6^2(z_1,z_2) &= G(z_1)H(z_2)H(z_1^2)G(z_2^2) \\
W_3^2(z_1,z_2) &= G(z_1)G(z_2)H(z_1^2)H(z_2^2) & W_7^2(z_1,z_2) &= G(z_1)H(z_2)H(z_1^2)H(z_2^2) \\
\\
W_8^2(z_1,z_2) &= H(z_1)G(z_2)G(z_1^2)G(z_2^2) & W_{12}^2(z_1,z_2) &= H(z_1)H(z_2)G(z_1^2)G(z_2^2) \\
W_9^2(z_1,z_2) &= H(z_1)G(z_2)G(z_1^2)H(z_2^2) & W_{13}^2(z_1,z_2) &= H(z_1)H(z_2)G(z_1^2)H(z_2^2) \\
W_{10}^2(z_1,z_2) &= H(z_1)G(z_2)H(z_1^2)G(z_2^2) & W_{14}^2(z_1,z_2) &= H(z_1)H(z_2)H(z_1^2)G(z_2^2) \\
W_{11}^2(z_1,z_2) &= H(z_1)G(z_2)H(z_1^2)H(z_2^2) & W_{15}^2(z_1,z_2) &= H(z_1)H(z_2)H(z_1^2)H(z_2^2)
\end{aligned}
\tag{5.10}
$$

These basis wavelet packets correspond to the areas of x-t and *f-k* space shown in **Figure 5-16**. The packet scales are indexed on both axis, with the sum of the axes labels giving the index of the wavelet packet scale. Again as for the wavelet transform,
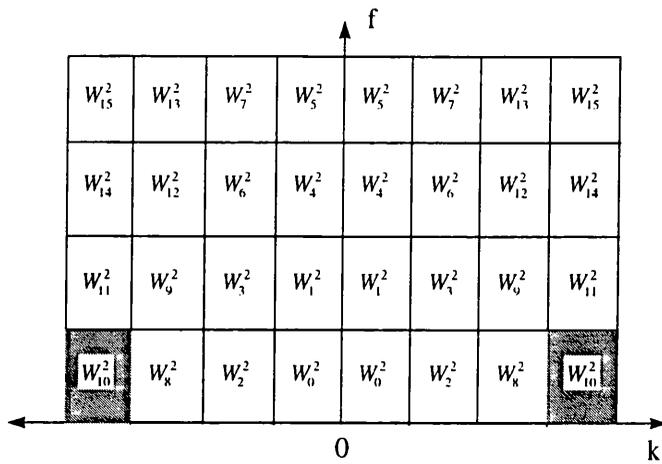
(a)
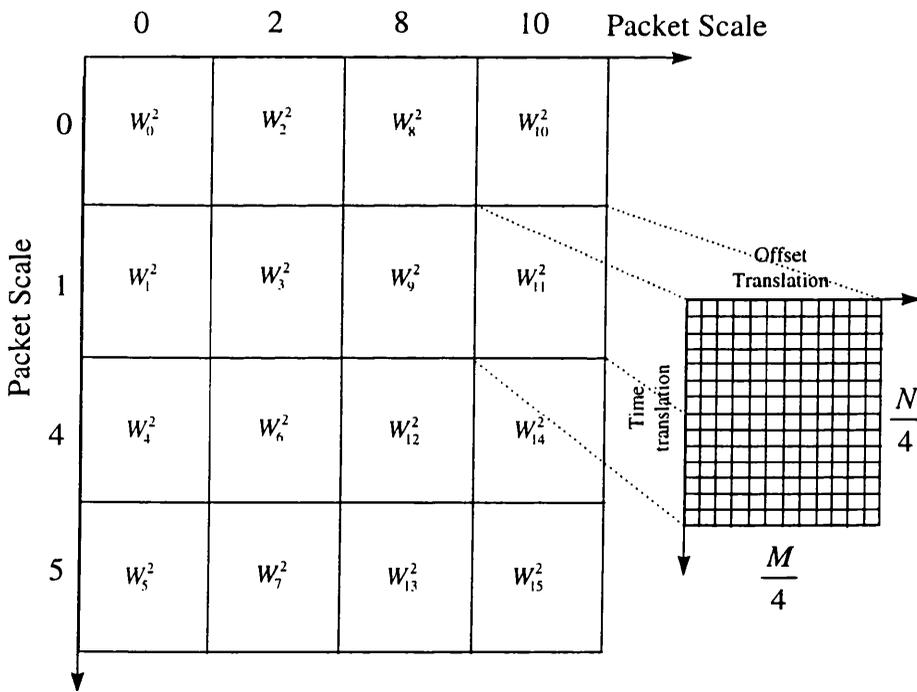
(b)

Figure 5-16 Partitioning of the (a) f-k domain after two iterations of the wavelet packet transform process and (b) the wavelet packet coefficients are organised into sectors as for the wavelet transform with each sector of wavelet packet coefficients tiling the input x-t space such that at the resolution shown, the Heisenberg cells in x-t space will be 4x4 samples in size. Each sector of coefficients is associated with a corresponding mirror pair of sectors in f-k space.

the sectors in *f-k* space are symmetrical about the frequency axis such that the two shaded areas in **Figure 5-16** are represented by one wavelet packet. After each iteration the number of basis functions is squared and the two-dimensional wavelet packet decomposition is in the form of a linear weighted sum of these basis functions,

$$f(x,t) = \sum_f \sum_s \sum_{p=-\infty}^{\infty} \sum_{q=-\infty}^{\infty} A^s_{f,p,q} W^s_{f,p,q}(x,t) \tag{5.11}$$

where $A^s_{f,p,q}$ is the wavelet packet coefficient, that is the double inner product of the signal *f(x,t)* with the two dimensional wavelet packet $W^s_{f,p,q}(t)$, having selected values of the level index *f*, the packet scale index *s* (number of iterations) and the offset and time translation parameters *p* and *q*. After each iteration we have a constant level of resolution each of which, as for the one-dimensional transform, represent a basis for the original signal. Again, as for the one-dimensional case, bases can be selected across levels of resolution. This is performed by a best basis selection procedure.

So as for the one-dimensional wavelet packet transform we have a more adaptable tiling of the *f-k* and x-t domain, and in filtering we will select the best **level** that localises the noise in the wavelet packet domain. When we investigated possible kernel wavelets to use in the wavelet packet decomposition in one-dimension, we stated that we must minimise any aliased noise introduced by a given kernel wavelet by minimising the side lobe energy in the associated frequency spectrum. **Figure 5-17** shows the x-t and *f-k* wavelet packet impulse response for three kernel wavelets, the four coefficient Daubechies wavelet, the least asymmetric 20 coefficient Daubechies wavelet and the quintic spline Battle-Lemarié wavelet. From this we can see that the side lobe energy off the main *f-k* support is present in two dimensions and as for the one-dimensional case, is minimised by the spline wavelet. This, however, leads to a corresponding increase in the x-t support of the wavelet. From section 5.4, we saw that to achieve a good reconstruction of the input signal using the Battle-Lemarié quintic spline wavelets (that minimise the side-lobe energy) we need to substantially increase the number of coefficients in the associated quadrature mirror filters. So again we have a quandary between optimising the speed of the filter or optimising the performance of the filter. Again, we can use the application of a gain function to minimise this problem. However, even with the gain function the quintic spline wavelet still has greater than twice the coefficients of the 20 coefficient Daubechies wavelet and so will take twice as long in the transform process. Therefore when examining the filtering
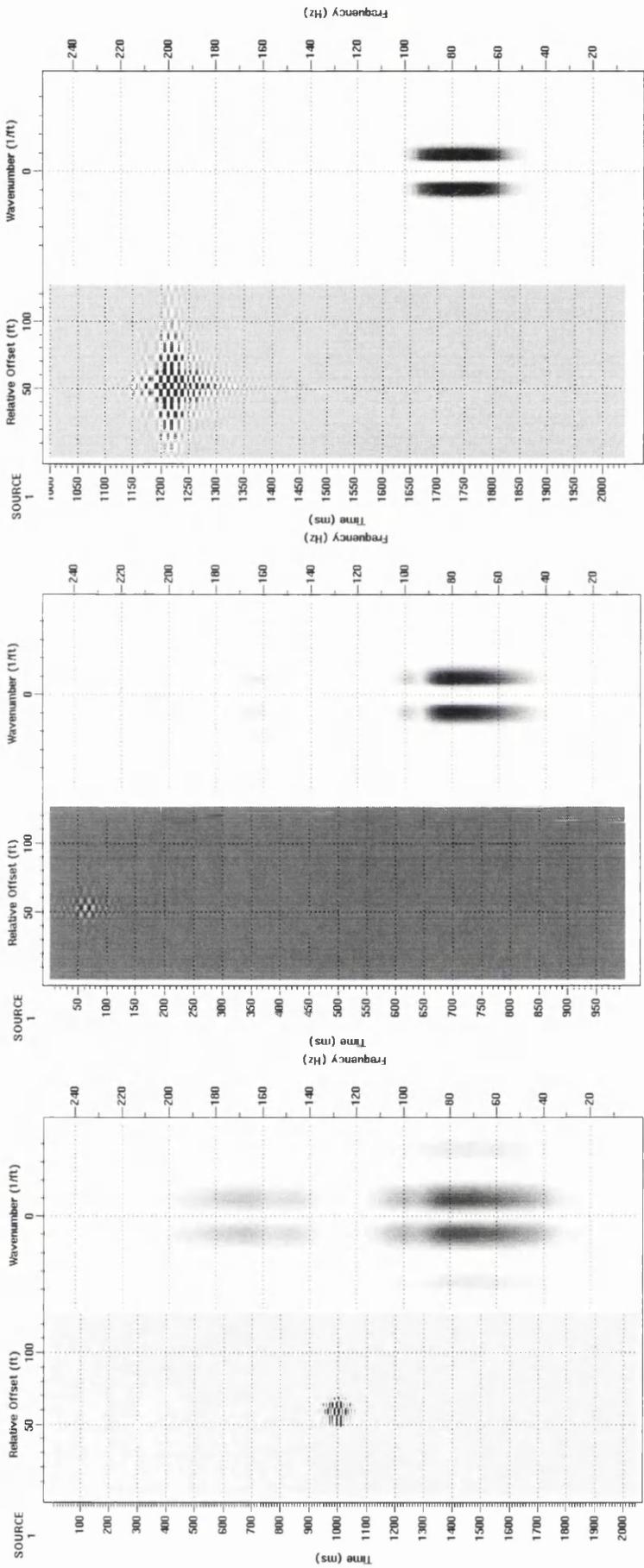
Figure 5-17 X-t and f-k form of two dimensional wavelet packets at the same packet scale generated from (a) four coefficient Daubechies kernel wavelet (b) twenty coefficient least asymmetric Daubechies kernel wavelet and (c) quintic spline Battle-Lemarié kernel wavelet.

process we will use two different kernel wavelets, one to optimise speed, the 20 coefficient least asymmetric Daubechies kernel wavelet and the other to optimise filter performance, the quintic spline Battle-Lemarié kernel wavelet, and we will compare the results at the filtered shot record level.

5.7 Wavelet Packet Transform Filtering

Figure 5-18 shows the two dimensional wavelet packet transform of the common shot record shown in **Figure 5-12(a)** after four iterations of the transform process. Again, as for all the transform processes discussed before we have defined a edge handling and array padding technique. We can see from this figure that the representation of the two dimensional discrete wavelet packet transform is even more cumbersome than for the wavelet transform. The partitioning of the original wavelet coefficients into smaller and smaller groups means that we soon find it difficult to associate coefficients with particular signals, a process that was fairly straightforward in the one-dimensional case. Therefore, when filtering in the two-dimensional wavelet packet domain we select regions in the wavelet packet domain in association with the associated zones of the signal in *f-k* space as we did for two dimensional wavelet transform filtering.

This selection procedure is not the optimum technique as we are not using fully the information that the wavelet packet coefficients could give us. However, until we can develop an efficient method of representing the four-dimensional wavelet packet transform space, this technique allows us to specify areas in the wavelet packet domain to filter.

As for the two-dimensional wavelet transform filtering we filter by muting the selected area in the transform domain, and invert the transform step. **Figure 5-19** shows the result of this filtering process on the shot record we transformed previously, and the corresponding *f-k* spectrum. From this Figure we can see that the wavelet packet transform processing has provided the best filtered shot record so far, with reflectors becoming apparent below 3 seconds at short offsets. There is some residual noise left after the filtering process, especially at shallow depths, which could possibly be removed in the future if we can develop a better technique for picking the coefficients to be filtered. **Figure 5-20** shows the difference section for the filtered record, where we

can see the noise removed by the filtering process. From this figure we can also see that there also seems to be very little signal removed.

Figure 5-21 shows the result of the filtering process using the least asymmetric 20 coefficient Daubechies wavelet. From this we can see that the filter has not performed as well in suppressing the low velocity noise with this kernel wavelet. The filtering parameters were identical to before, telling us that the differences in performance are due to the x-t and *f-k* form of the kernel wavelet. When we alter the filter zones in the wavelet packet domain to improve the resultant filtered section we do not observe a significant improvement in the filtered section. This must be partly due to the crude technique of selecting the wavelet packet coefficients to be filtered. However, from this we can state that the selection of kernel wavelet influences the performance of the wavelet packet filtering process.
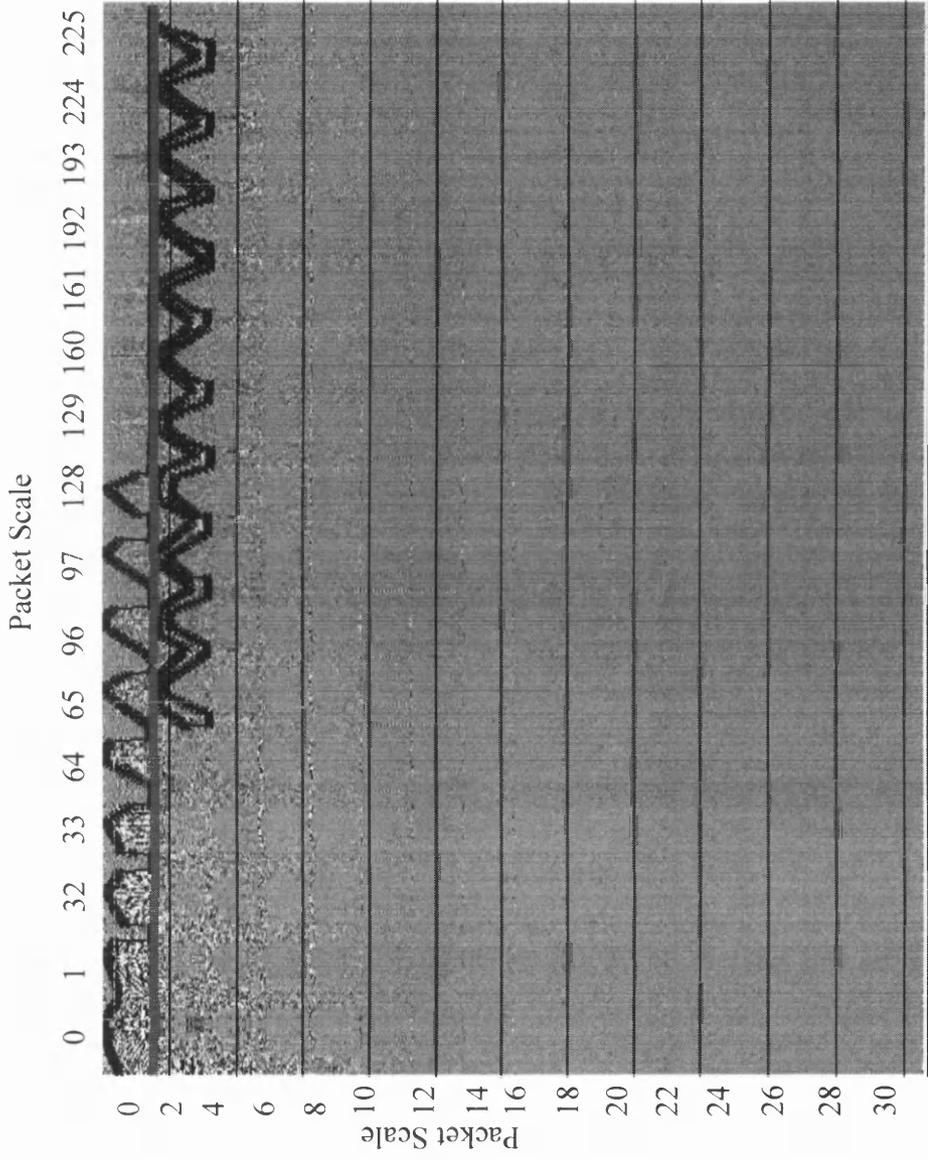
Figure 5-18 Wavelet packet transform of the shot record in Figure 5-12(a) after four iterations of the transform process. The thick black and grey lines outline the areas of the wavelet packet domain that are muted in the filtering process.
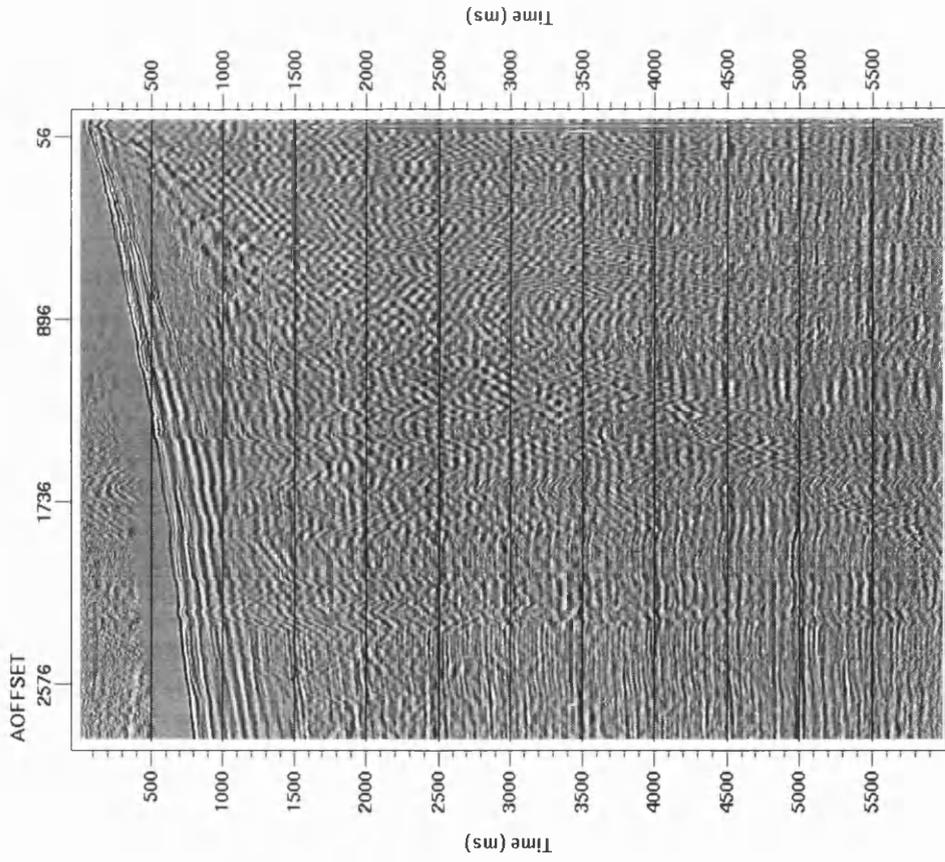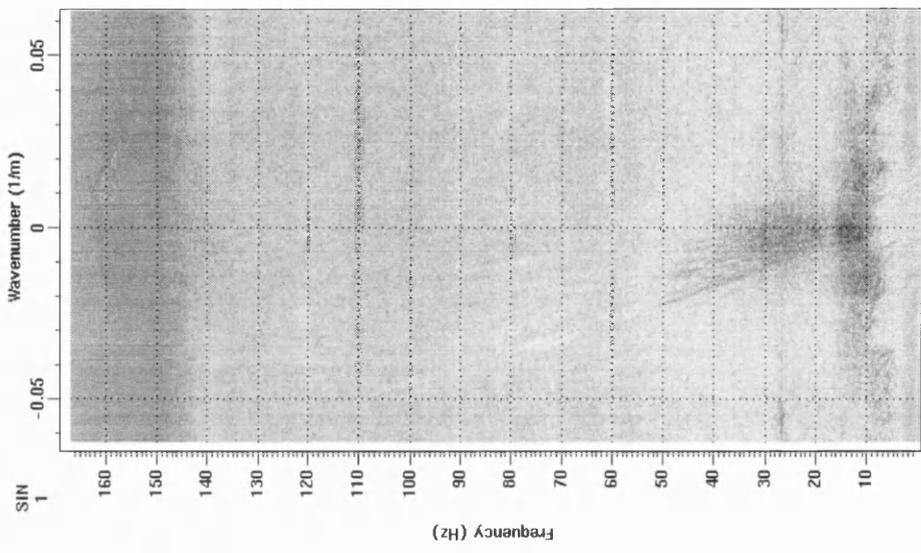
Figure 5-19 (a) The common shot record shown in Figure 5-12(a) after filtering using the two dimensional wavelet packet transform and (b) the corresponding f-k spectrum.
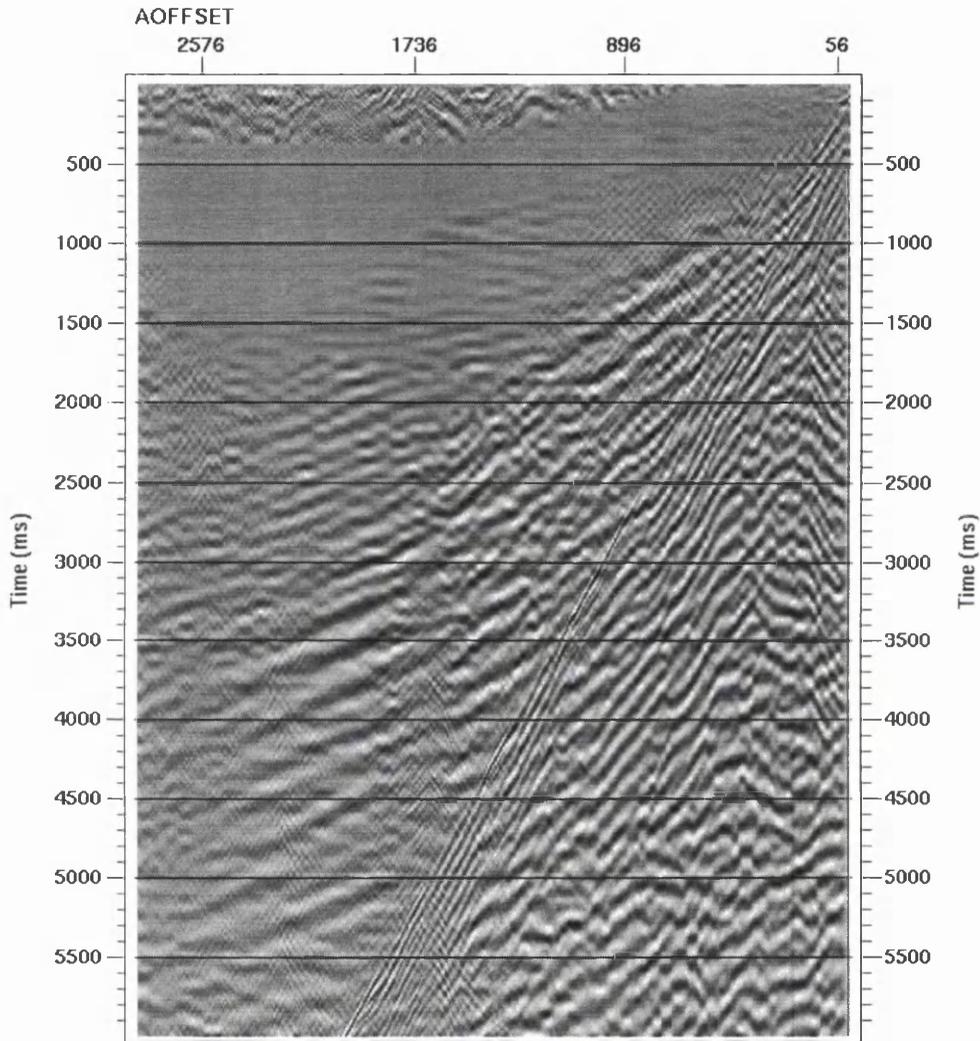
Figure 5-20 Difference section between the raw shot record shown in Figure 5-12(a) and the two dimensional wavelet packet filtered record shown in Figure 5-19(a). The transform used the quintic spline Battle-Lemarié kernel wavelet.
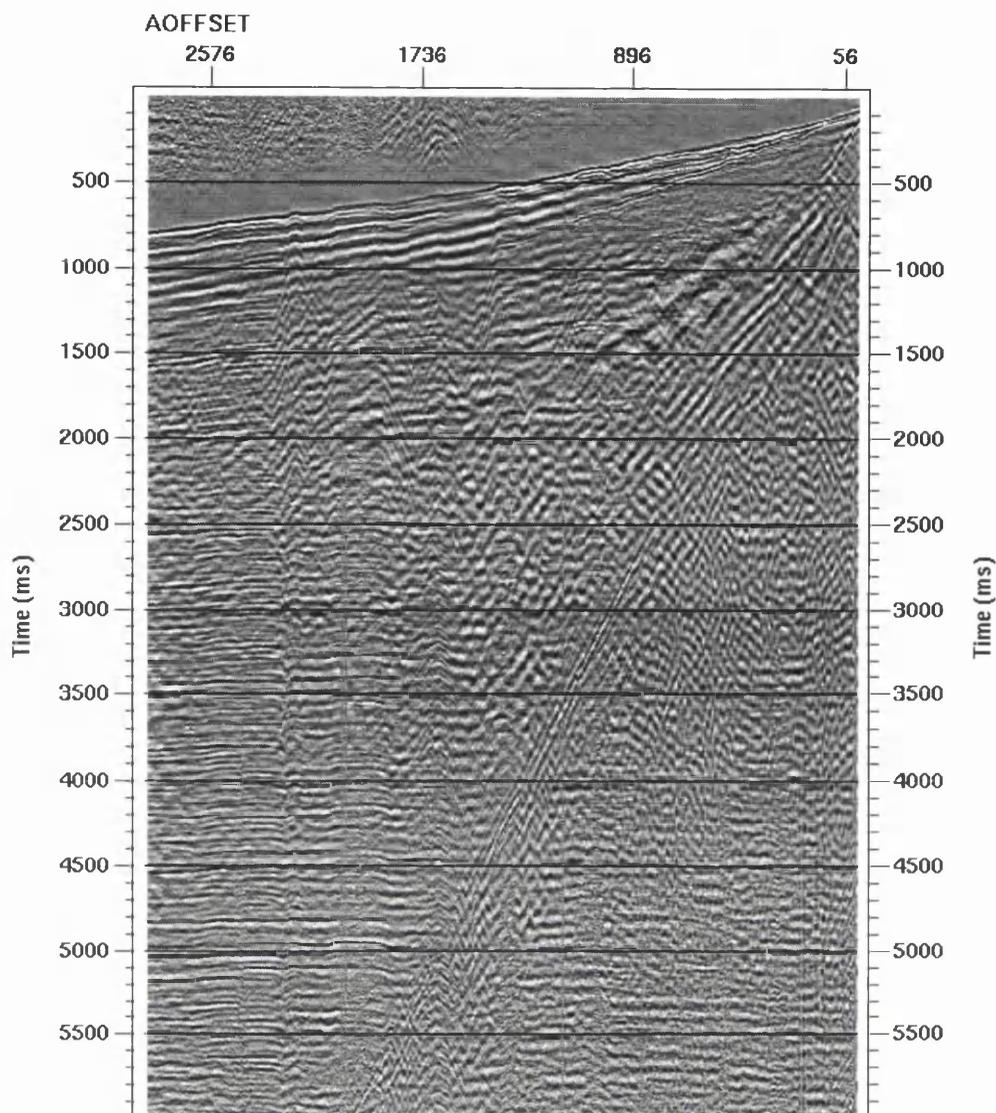
Figure 5-21 The common shot record shown in Figure 5-12(a) after filtering using a two dimensional wavelet packet transform. The 20 coefficient least asymmetric Daubechies kernel wavelet was used in the filtering process.

5.8 Amplitude Considerations

In the filtering processes considered in this chapter, the amplitude of reflected signals will be preserved if we have perfect signal-noise separation in the wavelet or wavelet packet domain, as with most filtering procedures. However, in most cases this is unlikely and therefore, the degree of amplitude distortion will depend on the degree of signal-noise overlap. When there is signal-noise overlap in the wavelet domain, a suitable weighting of the wavelet or wavelet packet coefficients may provide better amplitude preservation than the muting process. With the improvement that a simple weighting procedure brings in one-dimensional filtering, weighting procedures in the

two-dimensional wavelet domains are an avenue that should be investigated in the future, especially those based on information provided by the coefficients themselves.

5.9 Spatial Aliasing

A common problem with applying any two-dimensional filter to suppress low velocity signals from seismic gathers is spatial aliasing. As we mentioned earlier, the spatial axis is undersampled compared to the temporal axis. This undersampling leads to the aliasing of low velocity signals such that in *f-k* space, the signals are wrapped round from the positive wavenumber quadrant to the negative one. In extreme cases this can lead to overlap of signal and noise as shown schematically in **Figure 5-22**. When using wavelet or wavelet packet transforms to filter seismic signals we have the additional problem that in *f-k* space, the basis wavelets at any given scale or packet scale are symmetrical about the frequency axis. However, with the wavelet and wavelet packet transforms, we have additional time and offset variables present so the aliasing problem will only arise if the spatially aliased data overlaps with signal in *f-k* and x-t space.

To study the effects of spatial aliasing on the performance of wavelet based filters we filter a common shot record, **Figure 5-23(a)**, after several degrees of trace decimation. The shot record was acquired using an explosive source into geophone clusters with a 3 metre group interval. After trace decimation, the shot record has 12 metres (**Figure 5-24(a)**) and 24 metres (**Figure 5-25(a)**) receiver group spacing. The increasing degree of spatial aliasing is apparent in the accompanying *f-k* plots (**Figure 5-23(b)** to **Figure 5-25(b)**). The records were filtered with two dimensional wavelet packet filters using quintic spline Battle-Lemarié basis wavelets. The result of the filtering process on the shot records are shown in **Figure 5-23(c)** to **Figure 5-25(c)**. From these figures we can see the increasing degree of spatial aliasing leads to an increase in the degree of 'worminess' of the filtered record at short offsets. This is due to the filtering process being unable to separate the reflected signal from the low velocity noise due to increasing spatial aliasing. Another limitation associated with increased spatial aliasing is the corresponding decrease in the number of traces per gather. As the number of traces per gather decreases, so the number of scales in the offset domain decreases. This in turn decreases the resolution of the filter as it becomes more difficult to separate signal from noise in the limited number of scales.
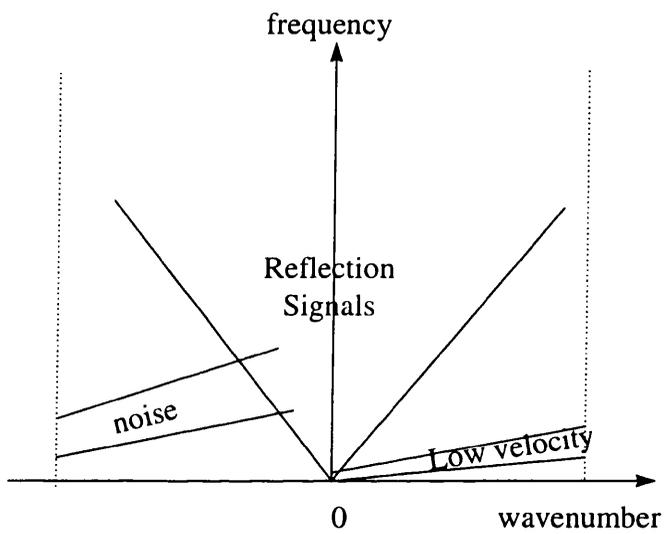
Figure 5-22 Schematic diagram showing the wrap around of low velocity noise in the f-k domain caused by spatial aliasing.
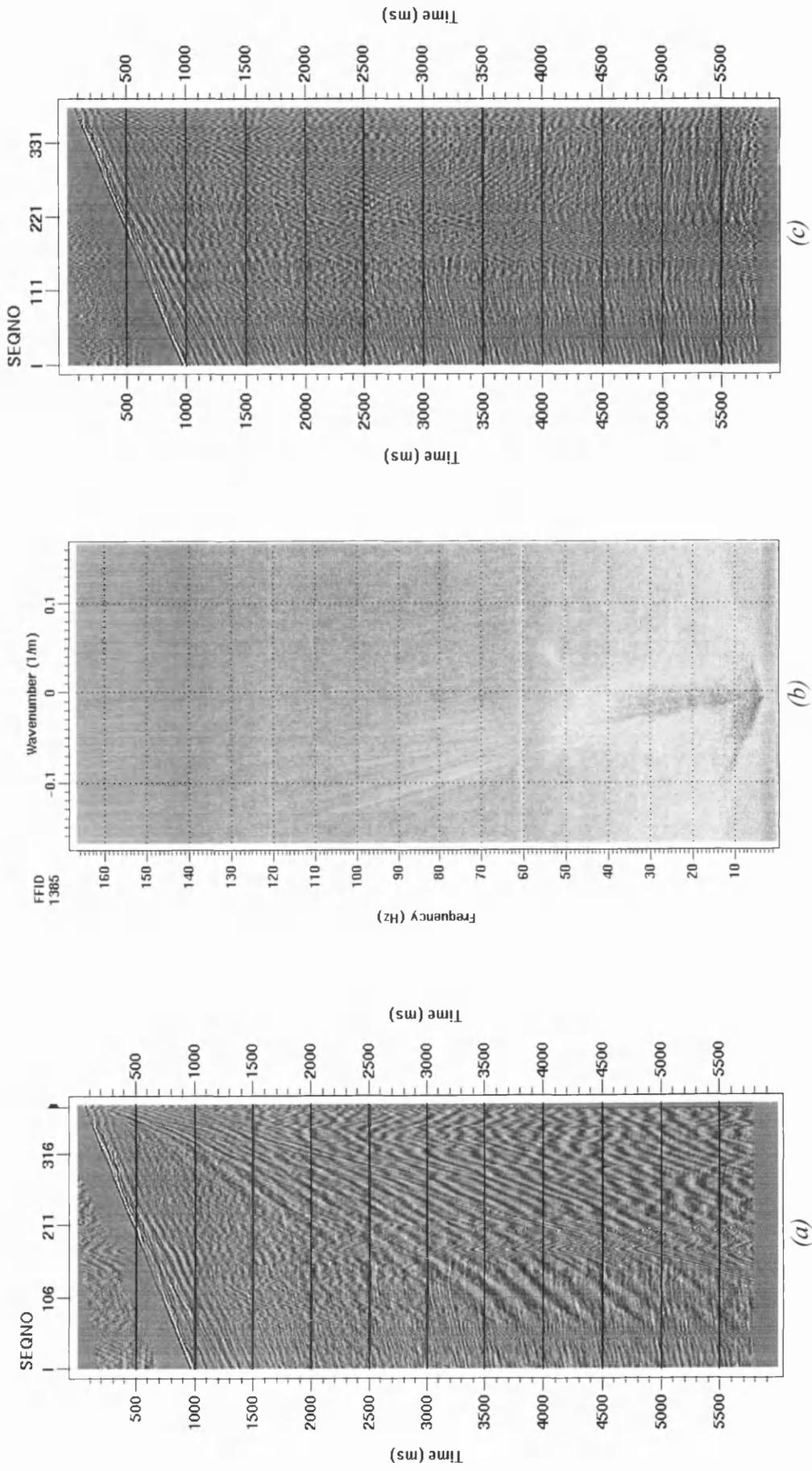
Figure 5-23 (a) A common shot record (3 m trace spacing) with (b) the corresponding f-k spectrum and (b) after filtering using the wavelet packet transform based filter. The data are not spatially aliased.

Chapter 5: Two Dimensional Filtering
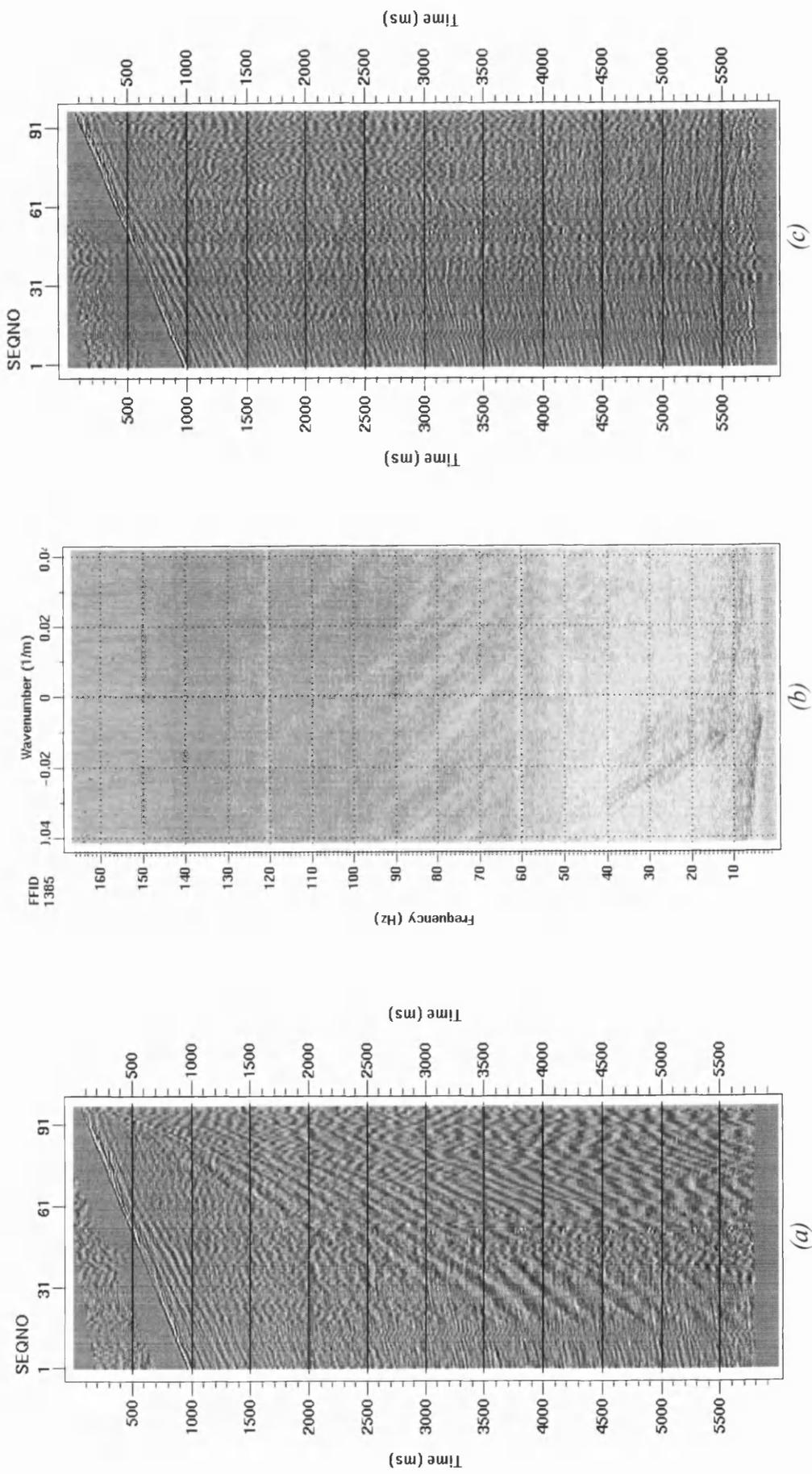
Figure 5-24 (a) The common shot record from Figure 5-23 after trace decimation such that the trace spacing is 12 m with (b) the corresponding f-k spectrum and (c) after filtering with the wavelet packet transform based filter. The data are slightly spatially aliased.
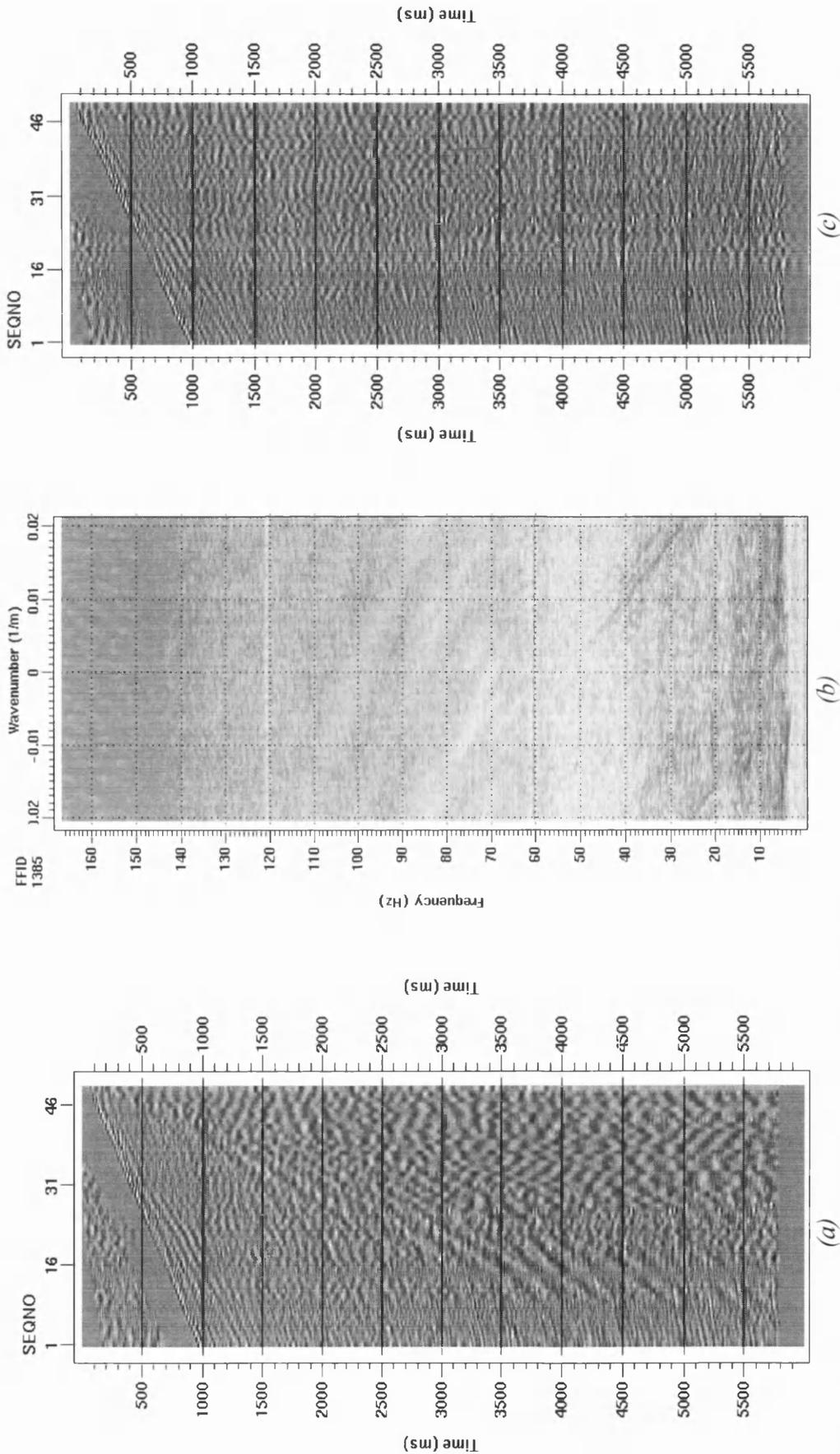
Figure 5-25 (a) The common shot record from Figure 5-23 after trace decimation such that the trace spacing is 24 m with (b)the corresponding f-k spectrum and (c) after filtering with the wavelet packet transform based filter. The data are spatially aliased.

Chapter 5: Two Dimensional Filtering

5.10 Filtering Influence on the Stack

To evaluate the benefit of the filtering process on the stack, we have filtered a series of common shot records and taken the processing sequence through to a brute stack to see if the improvement observed at the common shot level is seen on the stacked section.

Figure 5-26 shows a common shot record from a land-based seismic survey acquired in Northern Ireland. The data were acquired using a source array consisting of three Failing Vibroseis trucks with 12.5 m spacing generating six linear sweeps of 10-80 Hz per shot point. The geophones were deployed in linear arrays of 12 geophones with 1.25 m spacing and 15 m group interval. The survey area overlay a sequence of basalt layers near the surface which has lead to poor data quality, with the dominance of ground roll, guided waves and other coherent noise such as converted waves on the shot record as can be seen from **Figure 5-26**. From this survey we processed a sequence of 57 shot records to determine the effect of two dimensional wavelet packet filtering on the quality of the final stack.

Each shot record was filtered using the wavelet packet transform process described in previous sections. The quintic spline Battle-Lemarié kernel wavelet was used in the decomposition process. The data were padded out to the next power of two by mirroring the data about the last sample/trace, and the periodic boundary condition was used in the transform process. **Figure 5-27** shows the result of the filtering process on the shot record shown in **Figure 5-26**. We can see that the coherent noise has been suppressed considerably. **Figure 5-28** shows the brute stack obtained form the raw shot records. The stack is contaminated by steeply dipping noise which at earlier times disrupts the continuity of reflectors, and at later times obscures reflectors. **Figure 5-29** shows the stack of the filtered shot records from which we can see that there is considerable improvement in data quality. The shallow reflectors are more continuous, and reflections appear stronger at later times due to the suppression of noise. **Figure 5-33** and **Figure 5-30** show the stacks after the application of an AGC to enhance weaker reflectors. We can see that at later travel times, reflection energy is more coherent and the steeply dipping noise has been attenuated. For the two stacked sections the processing stream was identical and is shown schematically in **Figure 5-32**. As the processing sequences for the two brute stack sections are identical we have not taken

full advantage of the filtering result, and the corresponding improvement the filtered records would give in static solutions and velocity picks. However, comparing the records does show that without these subsequent improvements, improvement in the quality of the stacked section is easily apparent.

Figure 5-31 shows the difference section between the two stacks. From this we can see the noise removed from the stack and we can also observe that at earlier travel times, some reflected energy has been removed also. This is a result of poor signal noise separation in the wavelet packet domain and of poor filter design. The poor filter design results from using a mute in the wavelet packet domain rather than weighting the wavelet packet coefficients which, when we do not have good signal noise separation, may reduce the amount of reflection signal removed.
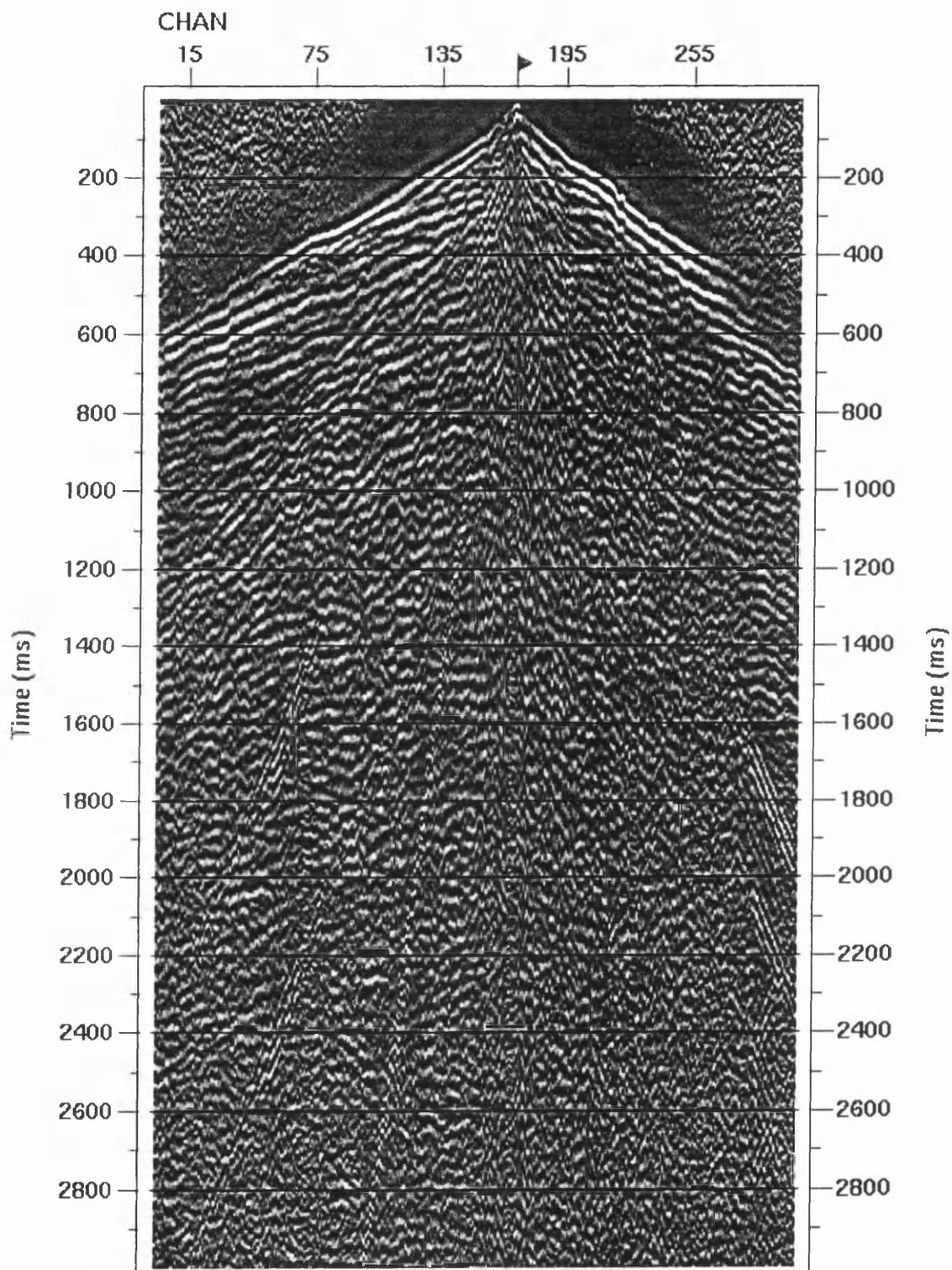
Figure 5-26 A common shot record from a seismic survey shot over basalt in Northern Ireland. The shot record is contaminated by steeply dipping coherent noise, a combination of ground roll, guided waves and mode conversions.
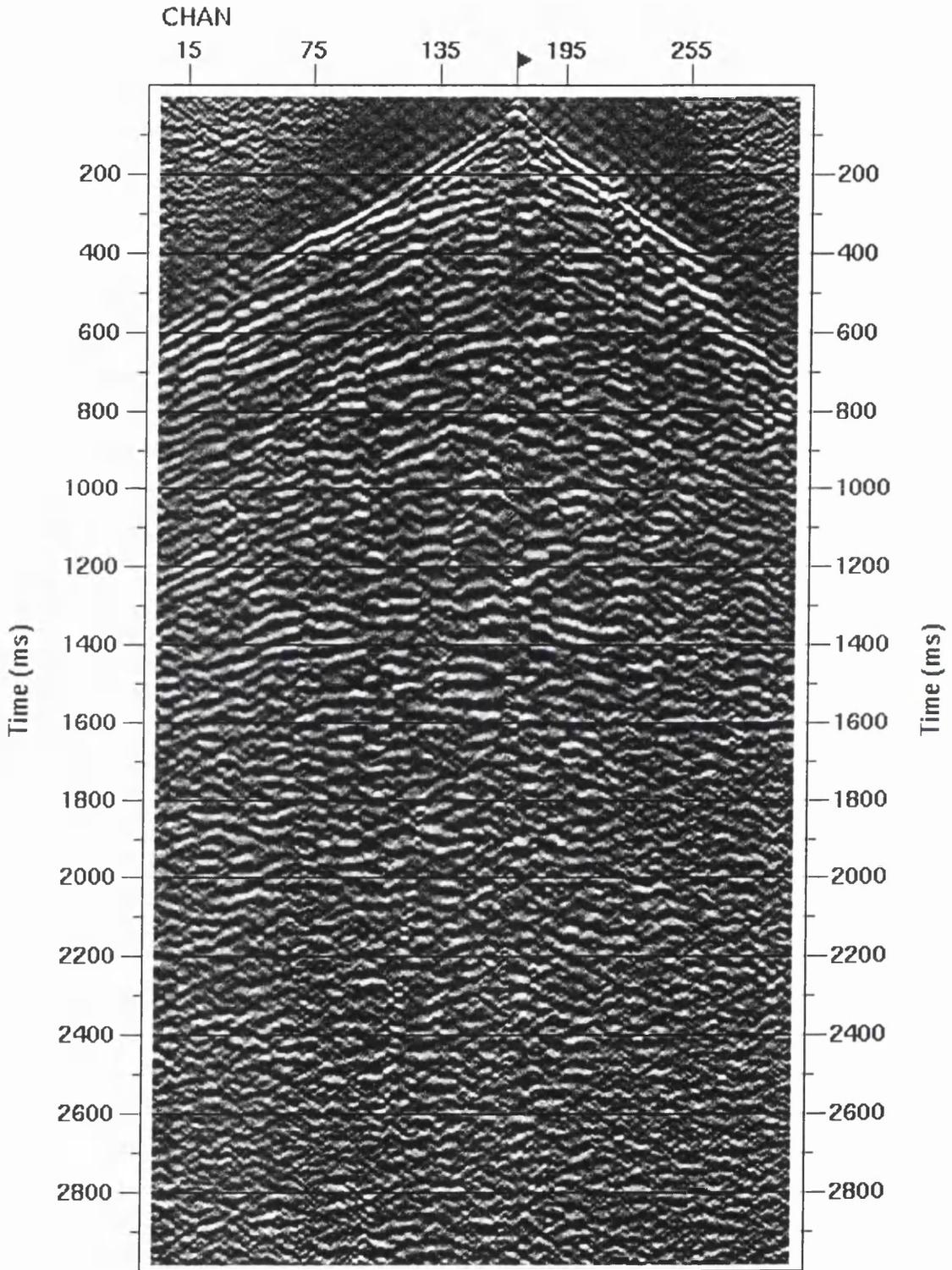
Figure 5-27 The shot record in Figure 5-27 after filtering using a two dimensional wavelet packet transform based filter. Comparing the figures we can see that the steeply dipping noise has been successfully suppressed.
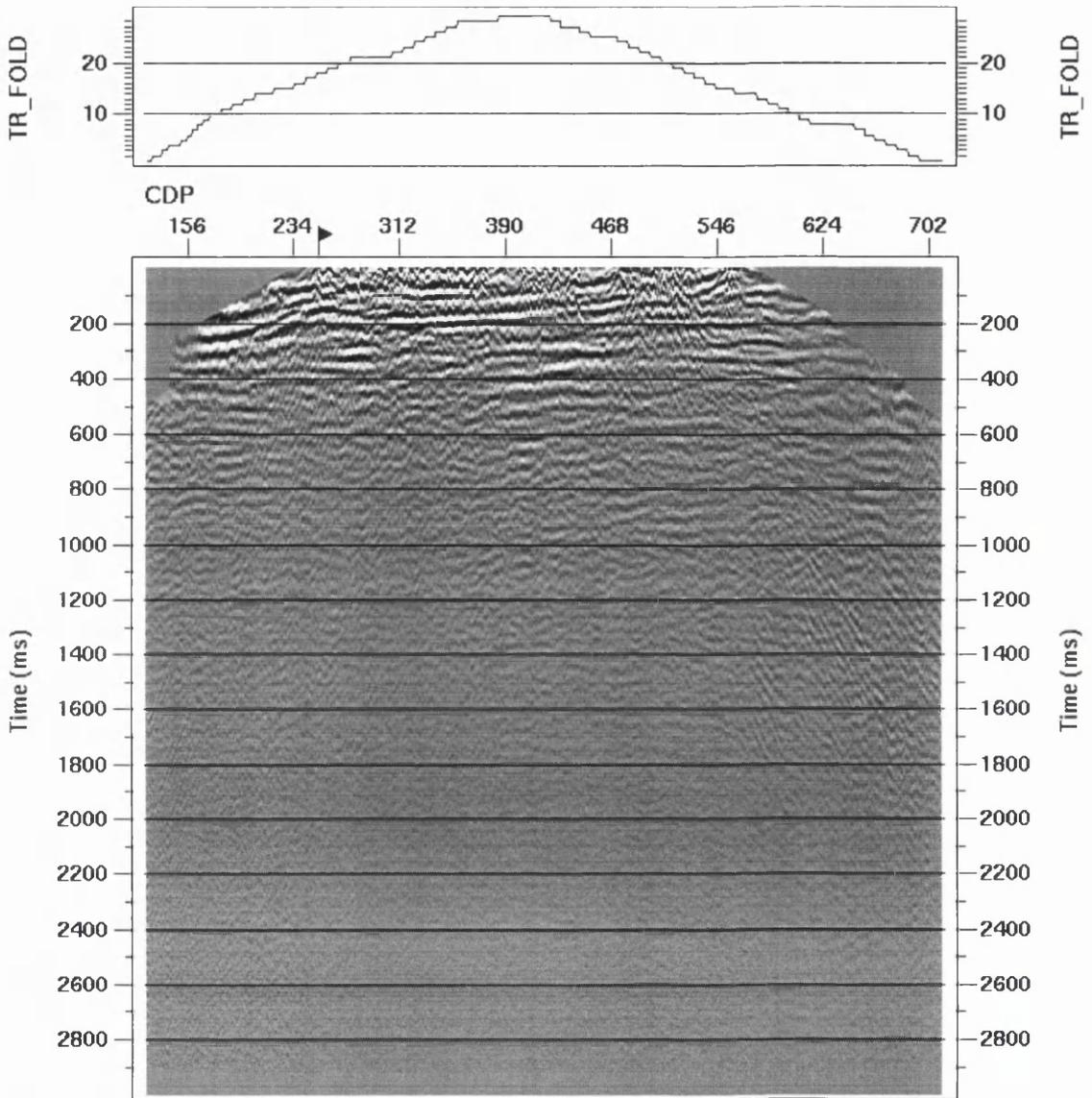
Figure 5-28 The brute stack of the line containing the shot record shown in Figure 5-26 with the corresponding trace fold (TR_FOLD). The processing sequence is shown in Figure 5-32.
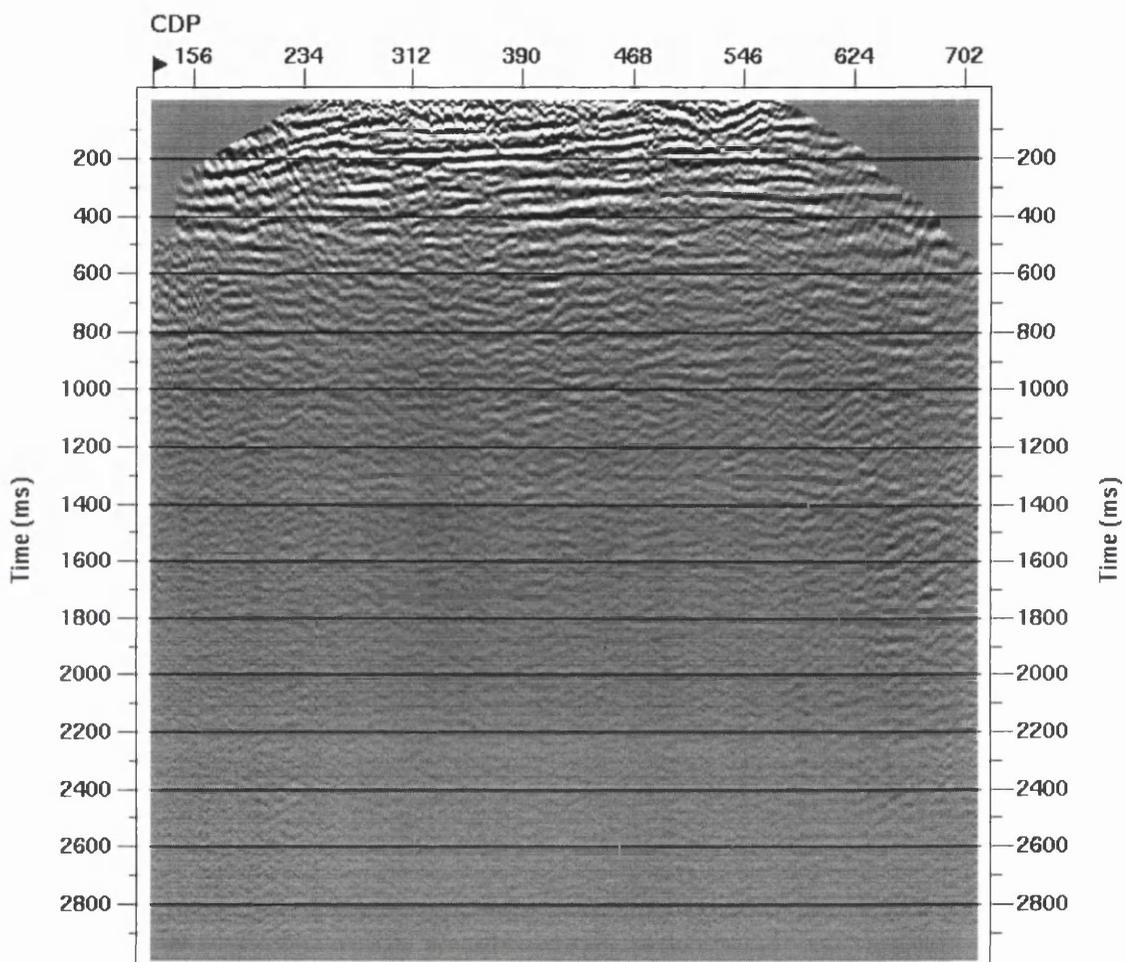
Figure 5-29 The brute stack of the seismic line shown in Figure 5-28 where the shots have been filtered using a two dimensional wavelet packet transform based filter. The processing sequence is shown in Figure 5-32.
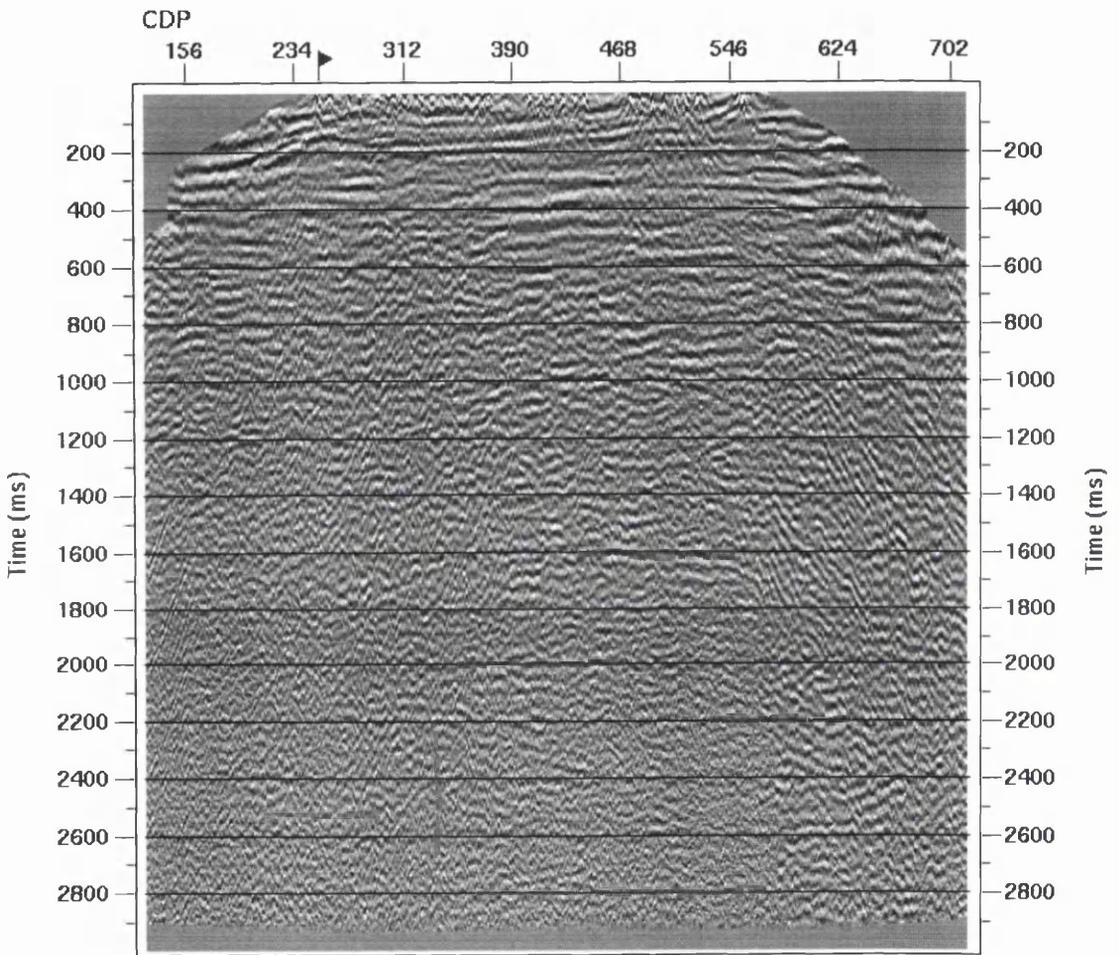
Figure 5-30 The brute stack after the application of a 500 ms AGC to enhance weak reflectors.
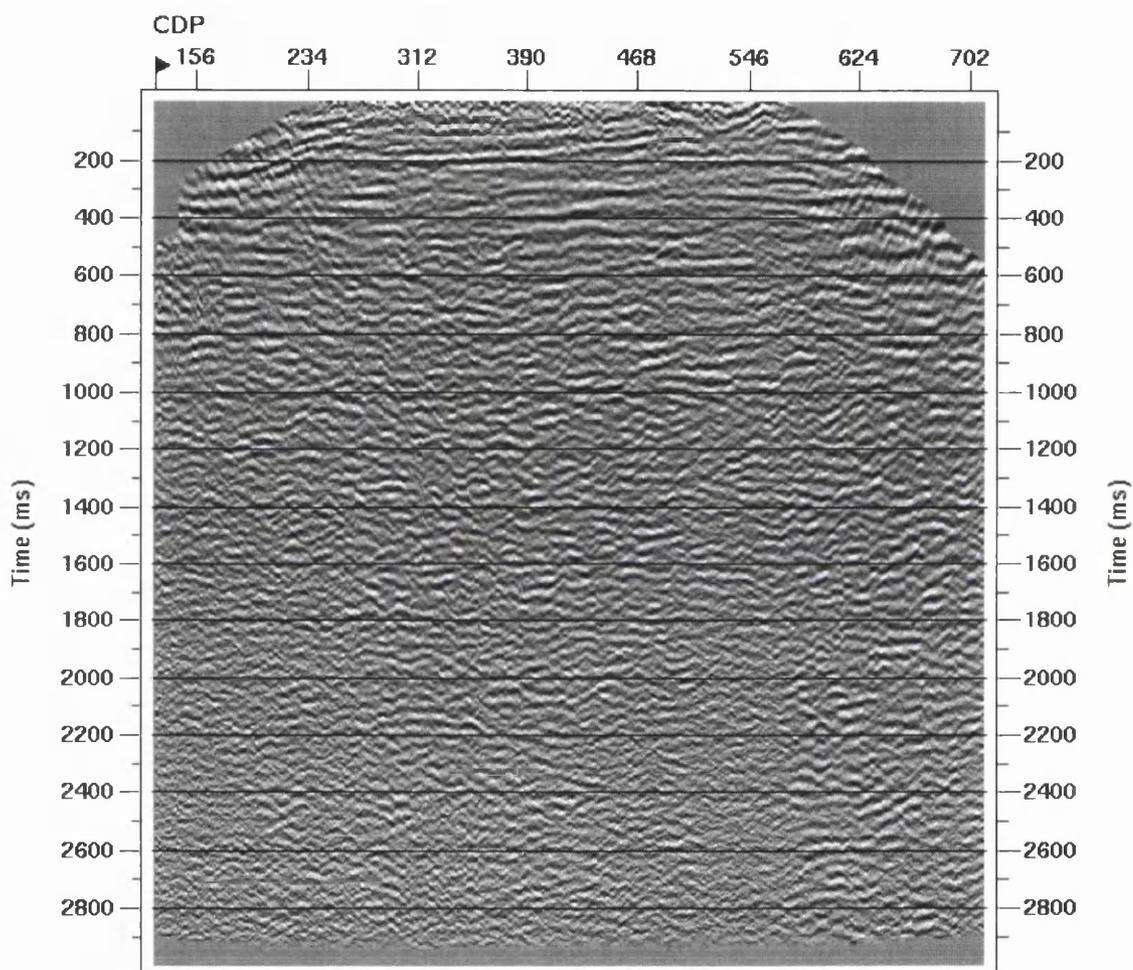
Figure 5-31 The filtered stack after the application of a 500 ms AGC to enhance weak reflectors.
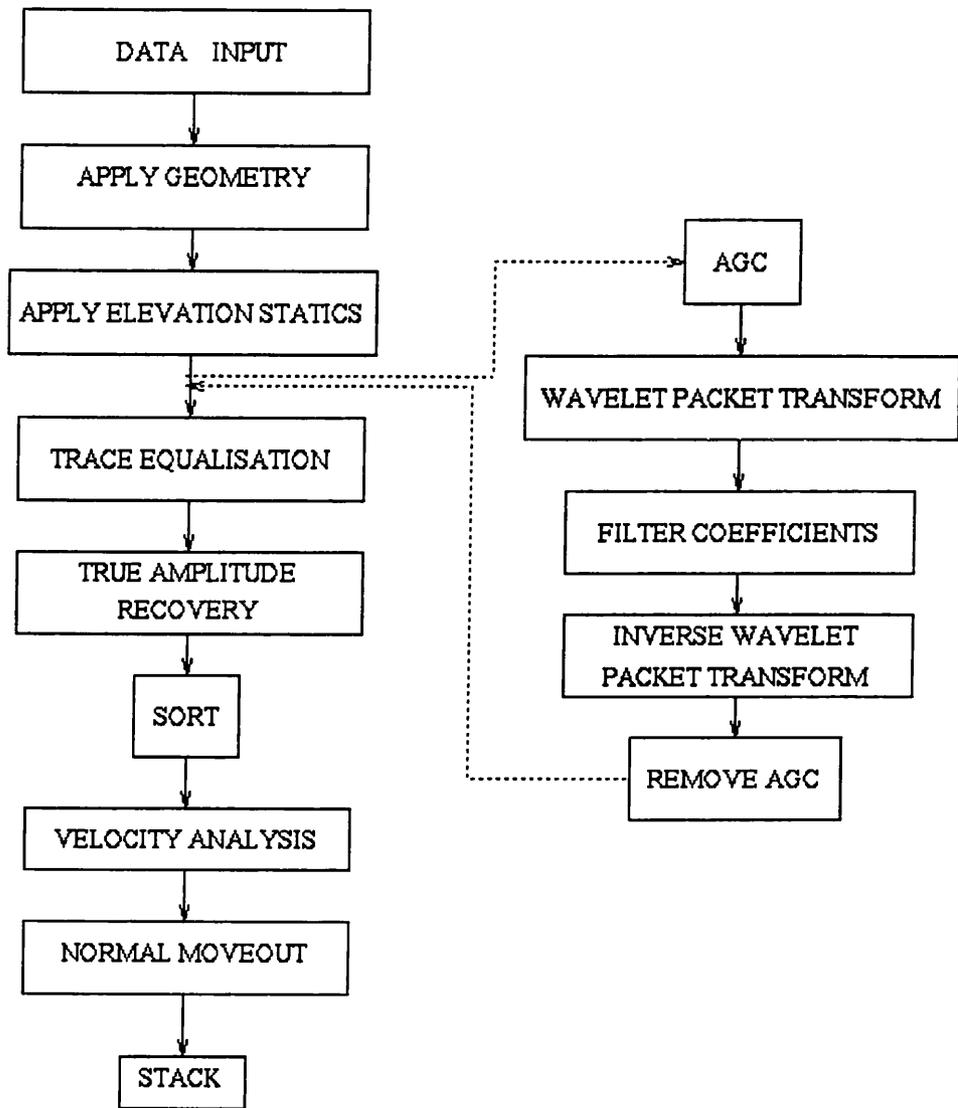
Figure 5-32 Processing flow for the Northern Ireland stacked section. The wavelet packet transform filtering procedure was inserted after elevation statics.
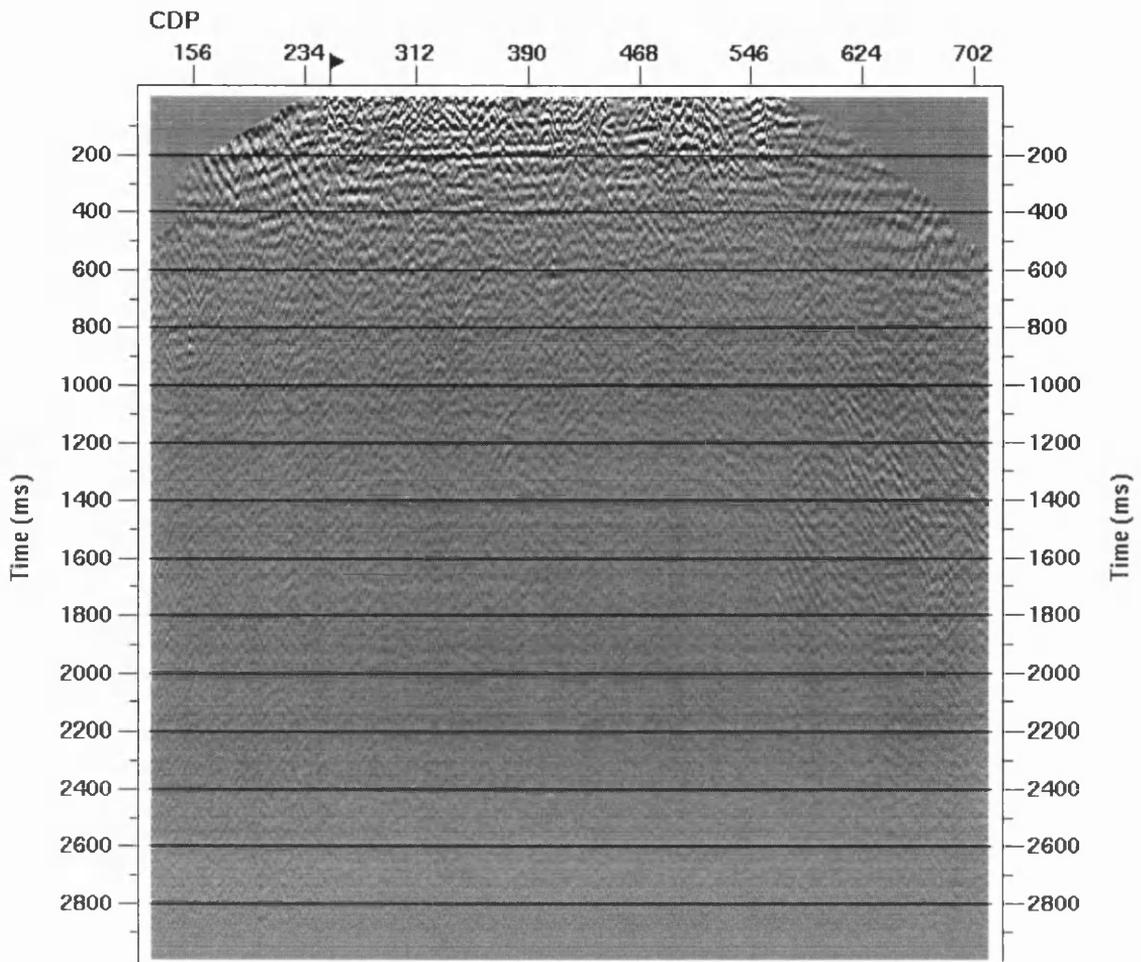
Figure 5-33 The difference section between the brute stacks with and without the application of the wavelet packet based filter.

5.11 Conclusions

By the extension of the discrete wavelet and wavelet packet filtering techniques to two dimensions we have shown that we can successfully suppress coherent noise from seismic records. The use of these transforms allows filtering in a time-frequency-wavenumber-offset sense, by the analogy of the scales and packet scales to frequency and wavenumber.

The wavelet transform decomposes *f-k* space into a series of octave segments, which we have shown may be too coarse a decomposition due to the positioning of seismic signals and noise in the *f-k* domain. We have shown that to make the two-dimensional wavelet transform effective as a filter, the addition of a shifting step (in this case a normal moveout) applied before and removed after the filtering process improves filter performance. The shifting moves the noise and signal into separate areas of *f-k* space which results in the performance improvement.

filter performance. The shifting moves the noise and signal into separate areas of f-k space which results in the performance improvement.

The extension of the filtering process to wavelet packets improves the resolution of the filtering in the f-k domain, but correspondingly decreases the resolution in the x-t domain by increasing the effective support of basis wavelet in offset-time. However, we have shown that using wavelet packets eliminates the need for the shifting process but with a corresponding increase in the transform time due to the extra iterations in the transform.

In both the wavelet and wavelet packet based filtering processes, the simple muting of coefficients in the transform domain effectively suppresses the coherent noise but, as we have shown, this can lead to the removal of signal as well as noise from seismic records, particularly at lower travel times. This is due to poor signal and noise separation in the transform domain. The addition of a weighting technique, which was successful in one-dimensional filtering processes, may lead to improved amplitude preservation.

We have shown that the choice of kernel wavelet is important in two-dimensional wavelet and wavelet packet transforms. The large trace to trace dynamic range of seismic data requires that a larger number of quadrature mirror filter coefficients need to be used for good signal reconstruction when using basis wavelets that are not perfectly compact, such as the Battle-Lemarié kernel wavelets. This can lead to a considerable increase in processing time compared to using perfectly compact kernel wavelets in the transform process. This increase in processing time can be avoided by the application of a gain function before filtering which can be subsequently removed after the filtering process.

In this chapter we have shown that two-dimensional wavelet based transforms can be used effectively to filter seismic records. In the next chapter, we investigate the use of wavelet transforms to filter seismic data by the application of the transform in the offset domain.

6. Velocity Filtering

6.1 Velocity Filtering of Seismic Data

Filtering of coherent noise from seismic data based on velocity has been a fundamental technique in seismic data processing for many years. These processes are used to discriminate against energy modes, such as ground roll or guided waves, that have a lower seismic velocity than primary seismic reflections on conventional shot and CMP records. Several different techniques are based on this principle which can be classified into global and local filtering techniques.

Global filtering techniques, such as those based on the *f-k* and τ-*p* transforms, use the separation of signal from noise in the transform domain to discriminate against noise. In these techniques a shot or CMP record (a 2-D array in offset-time) is transformed into the appropriate domain and filters designed by the selection of appropriate zones in transform space.

Local filtering techniques such as the delay and sum array filter (Dudgeon, 1977) are designed to maximise the degree of attenuation of undesired coherent energy and to minimise signal distortion caused by filtering. Most local techniques are applied across sensor arrays (that is in the offset domain of a shot/CMP record), after linear data shifting according to velocity, and amplitude scaling in an attempt to remove undesired signals travelling across arrays such as downgoing signals on VSP data or surface waves.

Lichmann and Northwood (1997) discuss the benefits and accompanying pitfalls associated with these techniques and their application. Local techniques design filters which have a small number of filter coefficients are speedy to implement, while global filters are easier to design, especially if there is clear signal and noise separation in the appropriate domain. Global filters which are applied to entire data arrays are difficult to vary in time whereas local filters, which are applied across an array in the offset domain, can be easily applied in a time varying fashion. Due to the amplitude scaling required in local filtering techniques, these methods tend not to be amplitude preserving whilst global filters preserve the relative amplitude of the data to a greater degree. The resolution and dynamic range of global techniques is dependent on the number of

sensors in the array and the number of time samples, unlike local techniques which can be implemented even when there are only a few sensors in the array.

A common problem associated with *f-k* filtering is that high resolution velocity notch filters cannot be designed effectively as they lead to the introduction of Gibbs oscillations into the filtered record. The amplitude of Gibbs oscillations, governed by the length of the pass-to-reject transition band in *f-k* space, can be minimised by increasing this ratio, leading to a corresponding reduction in resolution of the filter. Lichman and Northwood (1997) introduced a technique of high resolution velocity filtering based on Fourier techniques which does not introduce Gibbs oscillations and demonstrated its use for the suppression of ground roll, tube waves and multiples from seismic records. This technique involves the design of the filter in *f-k* space as an infinite continuous function which has the property of becoming a short discrete function in the x-t domain after the inverse Fourier transform. The short filter in the time domain has the same properties as the infinite function in the *f-k* domain. When combined with a linear data shift before transform these techniques can be used to discriminate against coherent events with relatively faster velocities compared to the primary reflected signal where the separation of events in transform space is more subtle.

The introduction of a hyperbolic shift rather than a linear shift allows hyperbolic velocity filtering which can be used for multiple suppression, a technique that is commonly used in combination with *f-k* filters. Multiple suppression is an important part of the seismic processing stream and current techniques are based on the distinct moveout of multiples when compared to primaries. The removal of multiple energy from the shot record is desirable for several reasons; to minimise multiple energy in the stack, to improve the quality of AVO analysis, and to allow the application of pre-stack migration which incorrectly treats multiples as primaries, leading to degradation of the final stack. The level of multiple suppression, measured as the improvement of primary-to-multiple amplitude ratio depends on the ability of the chosen transform to map the primaries and multiples to separate regions in the respective domain. This, in turn, depends on the moveout difference between the primaries and the multiples. The improvement in primary-to-multiple amplitude on traces prestack is not the only goal. The techniques must also preserve the amplitude of the primary amplitudes which can have fundamental importance in AVO interpretation.

Schuster and Sun (1993) introduced a technique of velocity filtering based the discrete wavelet transform and demonstrated its ability to suppress guided waves on VSP records and ground roll on land based seismic shot records. In this chapter, this local filtering technique has been developed further and adapted for the suppression of noise associated with direct arrivals and guided waves in land based seismic data and multiples from marine seismic data, investigating the effects of the technique on the stack and relative amplitudes of the data

6.2 Filter Methodology

As an illustration of Schuster and Sun's methodology consider the synthetic CMP gather shown in **Figure 6-1a**. The record consists of three events, one minimum phase linear event and two zero phase hyperbolic reflected events. The linear event has 2-4-12-16 Hz corner frequencies and a velocity of 3,000 ft/s. The reflected events consists of Ricker wavelets with 30 Hz centre frequency and velocities of 7,000 ft/s and 11,000 ft/s. The amplitude ratio of linear event to hyperbolic events is 4:1. The gather consists of 96 traces with a 30 ft trace spacing. Filtering of the seismic record to suppress linear events using Schuster and Sun's technique involves five steps:

1) the application of a linear moveout corresponding to the velocity of the linear event aligning the signal to a constant arrival time.

2) apply a discrete wavelet transform in the shifted offset domain at each sample.

3) zero the wavelet and scaling coefficients at the lowest scales.

4) apply the corresponding inverse discrete wavelet transform.

5) apply an inverse linear moveout to the data, restoring the data to its original position.

Schuster and Sun state that only a partial wavelet transform need be applied in the shifted offset domain, decomposing to the 3rd scale (for a record of 96 traces) and the remaining scaling coefficients muted. **Figure 6-1b** shows the synthetic CMP record after filtering using this technique. From this figure it is apparent that the technique has successfully suppressed the constant velocity event whilst retaining the character of the hyperbolic events.
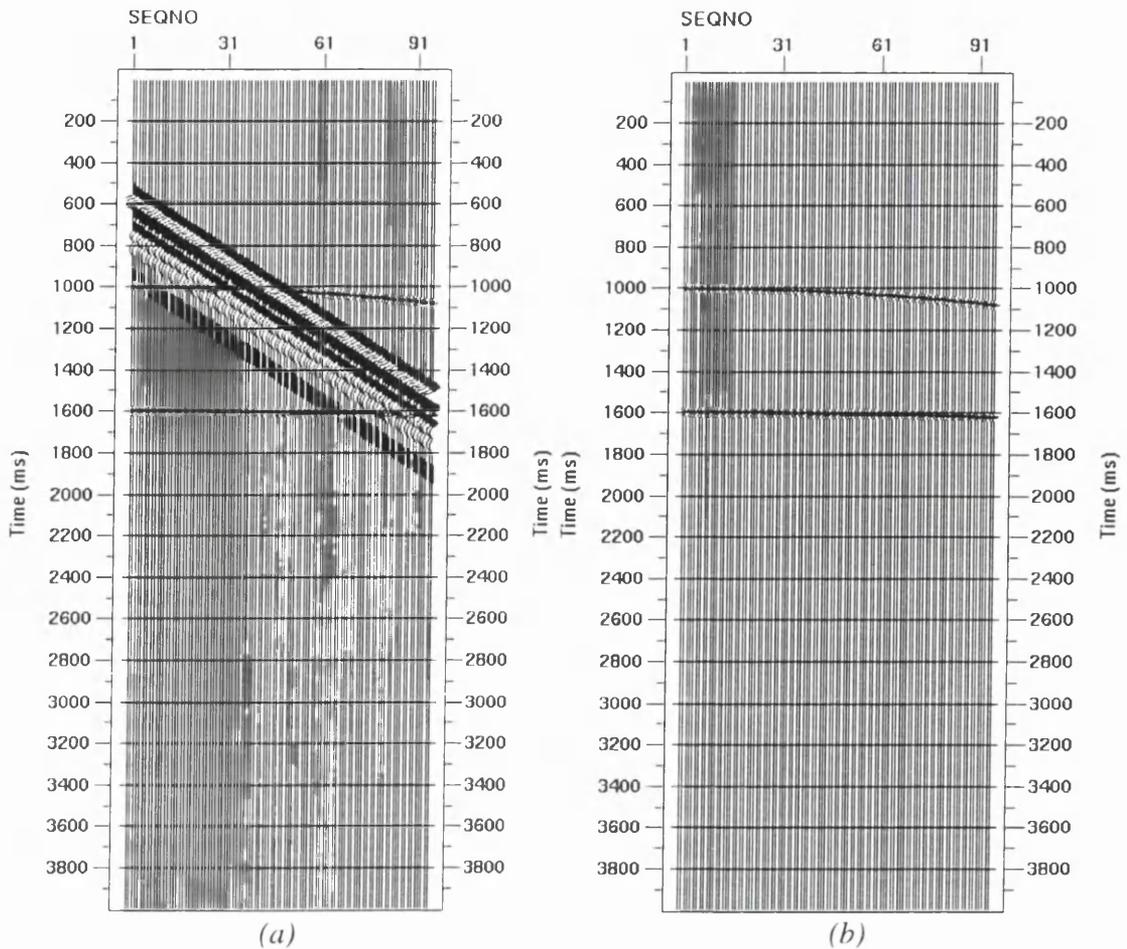
Figure 6-1 A synthetic CMP gather containing two hyperbolic and one linear event used to study the wavelet transform filtering technique of Schuster and Sun (1993) (a) before filtering and (b) after filtering.

The application of the this technique as a velocity filter is illustrated by the *f-k* spectra of the record before and after filtering as shown in **Figure 6-2**. After filtering, the energy in *f-k* space aligned along a band about the velocity of the linear moveout has been suppressed. The degree of suppression in this band is related to the amount of energy muted in the wavelet domain as well as the basis wavelet used in the decomposition. The signal is removed along the specified velocity and the filter slopes off with increasing/decreasing velocity in a similar fashion to the fall-off in frequency amplitude when filtering in time. The smaller the secondary lobes and the larger the slope of the scaling function in frequency space of the basis the faster the fall-off of the velocity filter in *f-k* space. This is indicated schematically in **Figure 6-3**. The width of the filtered band is determined by the amount of energy removed in the wavelet domain. The fewer iterations in the transform step the wider the band.
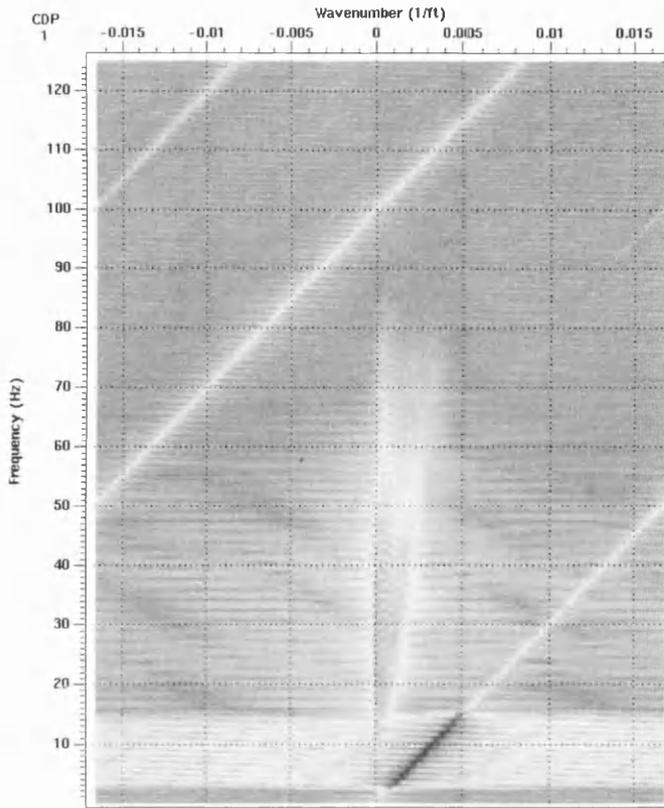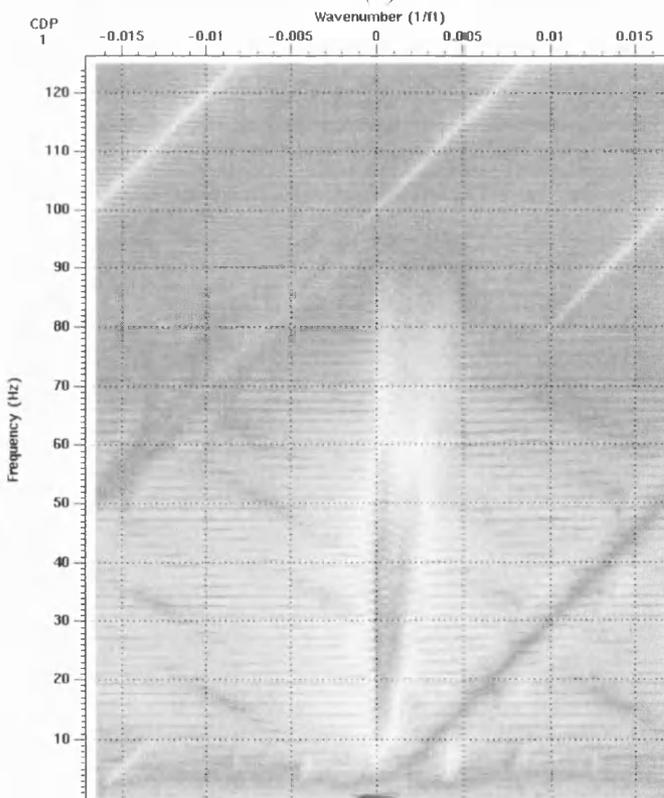
Figure 6-2 F-k spectra of Figure 6-1 (a) before and (b) after filtering using the wavelet transform technique. The spectrum has been suppressed along the velocity filtered.
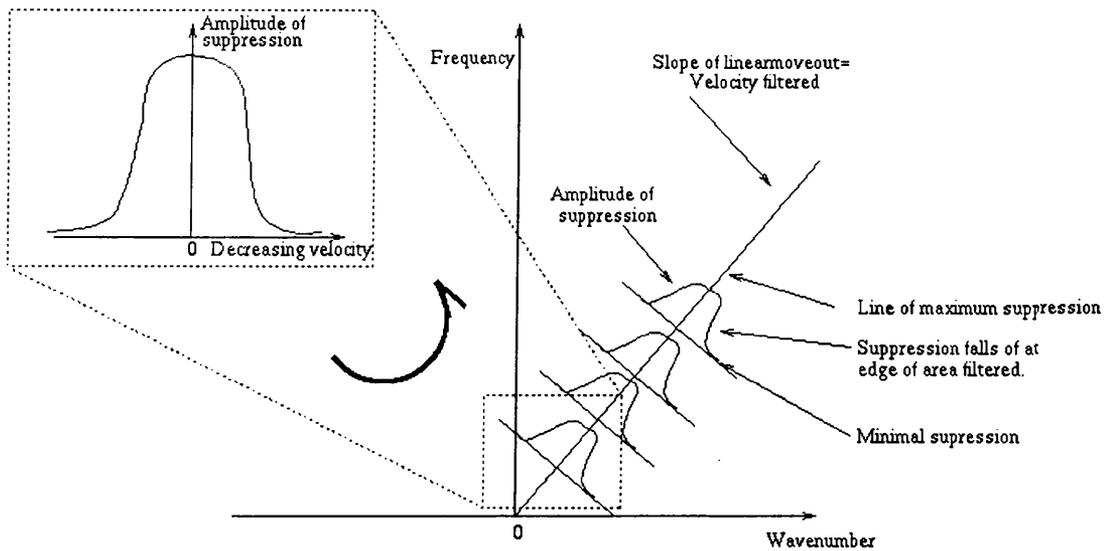
Figure 6-3 Schematic diagram showing the suppression of energy in f-k space by the wavelet transform filtering technique. The f-k spectrum is attenuated along the line of maximum suppression. The suppression depends on the kernel wavelet used and which coefficients are zeroed in the wavelet domain.

In this study, spline biorthogonal kernel wavelets (Chui, 1992) were used in the transform process. This was found through trial and error to give the best results in terms of filtering and amplitude preservation which are discussed in the next section.

6.3 Amplitude Considerations

The next step in ascertaining the effectiveness of the filter is to examine how the filter has affected the amplitudes of the hyperbolic reflected arrivals. To analyse the effect on amplitudes, a NMO corresponding to the hyperbolic event velocities was applied to the data and the event amplitude plotted against offset. **Figure 6-4a** shows the amplitude of the 1000 ms hyperbolic event of **Figure 6-1** before and after filtering. The large variations in the pre-filter amplitude at offsets less than 1700 ft are due to energy from the linear event superimposed on the hyperbolic event. Comparing the post-filter amplitude to the ideal amplitude response of 1 at all offsets, we can see that the linear event has been effectively suppressed whilst preserving the original reflector amplitude. **Figure 6-4b** shows more closely the post-filter amplitude in comparison with the ideal result of the synthetic gather without linear noise. Apart from relatively large amplitude variations at the edges of the signal, the amplitude of the event is distorted only very slightly, the maximum difference being approximately 1% of the ideal output. The distortion varies with offset yet no systematic
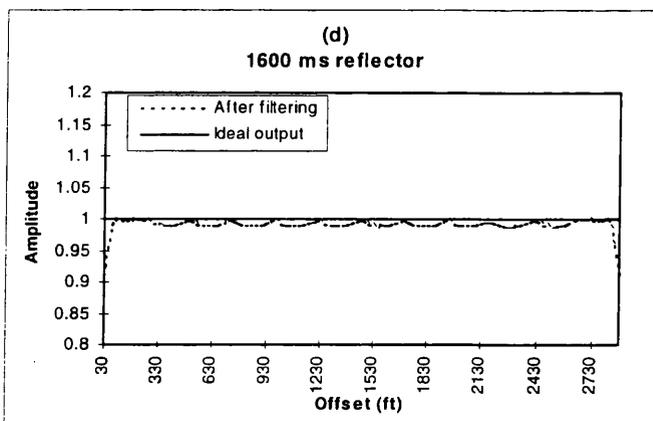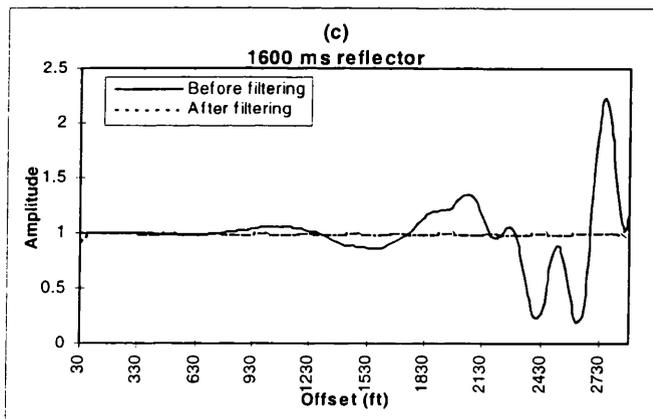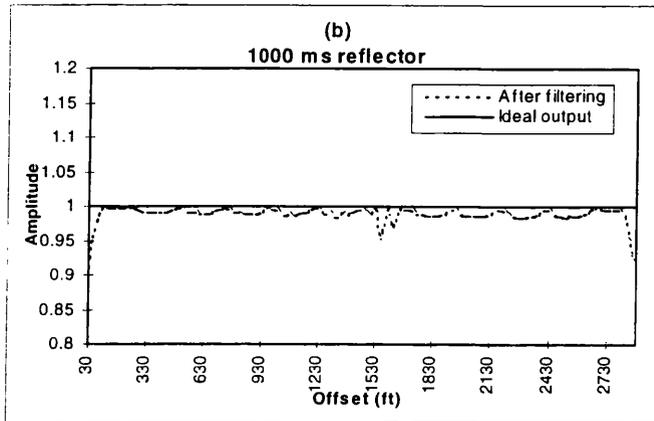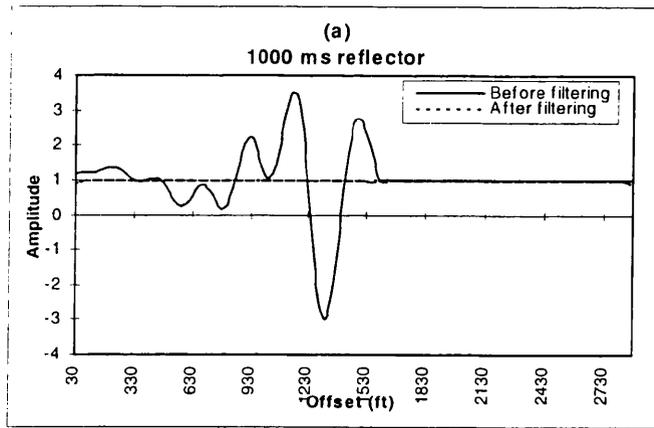
Figure 6-4 AVO relationships before and after filtering: 1000 ms reflector (a) before and after filtering, (b) after filtering compared to the ideal output: 1600 ms reflector (c) before and after filtering, (d) after filtering compared to the ideal output.

trends are introduced into the data, which may mimic AVO effects. **Figure 6-4c** and **d** show the corresponding graphs of amplitude versus offset for the reflector at the 1600 ms arrival time in **Figure 6-1**. For this signal we see less amplitude distortion than for the 1000 ms reflector.

To minimise distortion, the wavelet decomposition can be applied to a greater degree than the third scale, before removing the remaining scaling coefficients. **Figure 6-5** shows the amplitude of the 1000 ms and 1600 ms reflector after transforming to the 6th scale followed by muting of the remaining scaling coefficients. The distortion has been considerably reduced.

6.4 Amplitude Considerations and Faster Velocities

The model considered in the previous section can be considered to be simplistic in that there is perfect separation of the signal and noise in the *f-k* domain (**Figure 6-2**). Ideally with any global or local velocity filtering technique we should be able to get near perfect amplitude recovery of the reflection signal when there is minimal signal overlap in *f-k* space. In terms of 2-D Fourier analysis, any hyperbolic signal can be considered as the sum of a series of dipping linear events in x-t space with decreasing velocity as indicated schematically in **Figure 6-6**. Suppressing any linear event using this technique will result in the removal of one of the constituent components of the hyperbolic signal if there is signal overlap in *f-k* space. Extending this technique to the removal of linear events with faster velocities, for example in the case of guided waves or noise associated with first break energy, the amplitude of this energy overlap in *f-k* space will be more significant and larger amplitude distortions introduced. This is demonstrated in **Figure 6-7** and **Figure 6-8**. **Figure 6-7** shows a synthetic seismic record containing two zero-phase hyperbolic reflected events with velocities of 7500 ft/s and 11,000 ft/s and one linear event with a velocity of 15,000 ft/s and the corresponding *f-k* decomposition (the trace spacing, number of traces, and event frequencies are identical to the previous model). From the *f-k* spectrum we can see that there is considerable signal/noise overlap. **Figure 6-8** shows the CMP gather and corresponding *f-k* spectrum after filtering using Schuster and Sun's third scale criteria. The large area of *f-k* space about the linear event velocity which has been removed leads to severe distortion of the
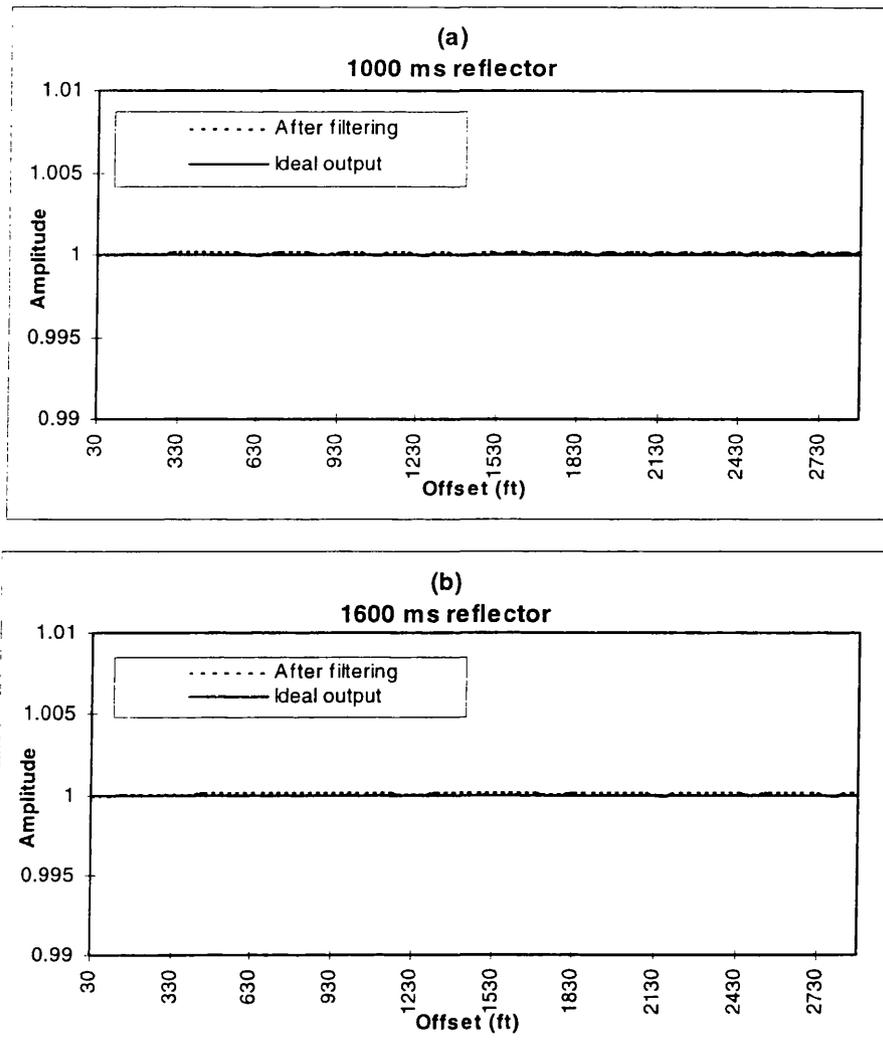
Figure 6-5 AVO relationships for the synthetic reflectors and ideal output of Figure 6-1 filtered in the wavelet domain after the application of the transform to a higher resolution for the (a) 1000 ms reflector and (b) 1600 ms reflector.
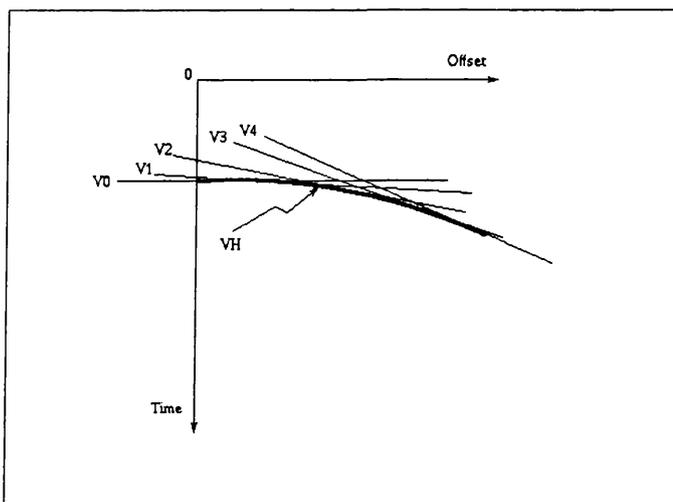
Figure 6-6 Schematic diagram showing a reflector of velocity VH as the sum of a series of linear events with velocities V0 to V4 with increasing dip (decreasing velocity).
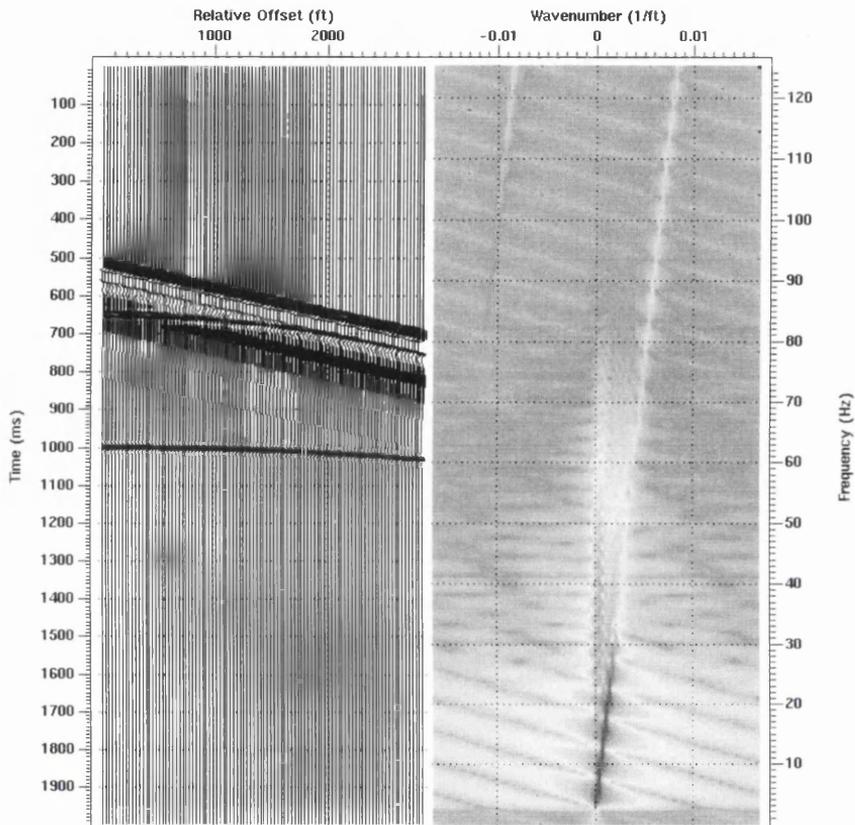
Figure 6-7 A synthetic CMP gather containing a linear event with faster velocity approaching reflectors tangentially and the corresponding f-k spectrum showing signal overlap in f-k space.
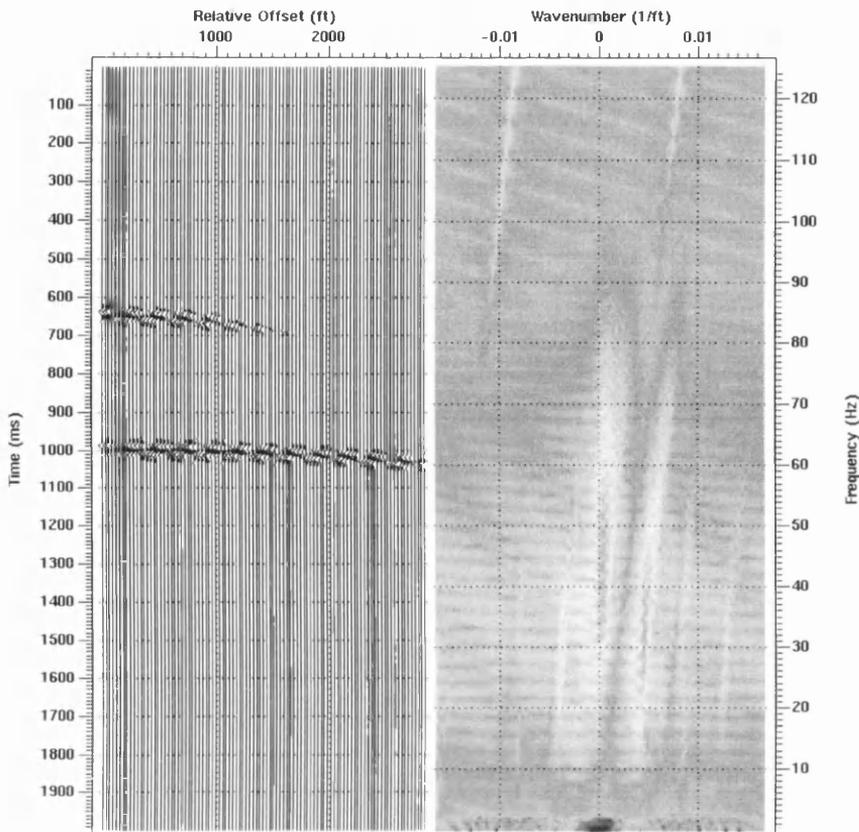
Figure 6-8 The synthetic CMP gather from Figure 6-7 after filtering using Schuster and Sun's criteria with the corresponding f-k spectrum.

reflected signals. To maximise the resolution of the filter, and so minimise reflected signal distortion, the transform can be applied to the largest possible scale (the tenth scale for a 1024 sample vector) and the remaining scaling coefficients muted. **Figure 6-9** shows the record and corresponding *f-k* representation after filtering using this maximum resolution criteria. Although some signal deterioration is still apparent, the character of the reflectors have been more faithfully retained than previously and the linear event suppressed.

Analysing this result more carefully, **Figure 6-10a** shows the amplitude versus offset relationship for the 650 ms reflector before and after filtering. Comparing the filtered amplitude compared to the ideal output, we can see the amplitude distortions (up to 35% of the ideal amplitude) at offsets greater than 1000 ft. **Figure 6-10b** shows the AVO relationships for the 1000 ms reflector and from this we can see that, the distortion is less pronounced (~8% of the ideal amplitude), and is concentrated mostly the signal edges. Distortions at large offsets for shallow reflectors may not be a major concern as these may be muted during NMO due to stretching of the seismic wavelet, and in AVO analysis due to large incident angles. Nevertheless, these amplitude distortions indicate

that a more robust technique for filtering the wavelet coefficients is required, the simple muting of coefficients is too simple.

In cases of signal and noise overlap in *f-k* space a weighting must be applied to the coefficients rather than muting. This weighting must retain the information in the scaling coefficients corresponding to the reflected event and discriminate against the information corresponding to the linear event. The problem to be addressed is how can we derive the correct weighting?

As discussed in Chapter 2.4.5 the discrete wavelet transform is translation invariant due to the downsampling process inherent in the transform. This downsampling process leads to differences between wavelet transforms of identical signals translated by one sample as shown schematically in **Figure 6-11**. Taking the wavelet transform of a signal scale by scale, where one scale of the transform contains two signals which overlap in frequency space, the variation in wavelet transform coefficients with signal translation resulting from downsampling will be minimised for the lowest frequency component of the scale, as amplitude changes with time are slower. Yet for the higher frequency components, the variation with translation will be more pronounced. To determine the component of a wavelet transform coefficient corresponding to these lower frequencies of a we can *calculate the mean of the values of the scaling coefficient for wavelet transforms of the input signal at all translations*. For a signal with a very low frequency component, such as the aligned linear events we are processing, the translations should have minimal effect on the value of the transform corresponding to this component. The higher frequency components will vary due to the translation. The scaling coefficient left when the wavelet transform has been applied fully, contains the very low frequency signal and some higher frequencies related to the hyperbolic signal. Calculating the mean of this coefficient at all translations of input signal will give an estimate of the low frequency component which can then be subtracted leaving an approximation to the remaining components.
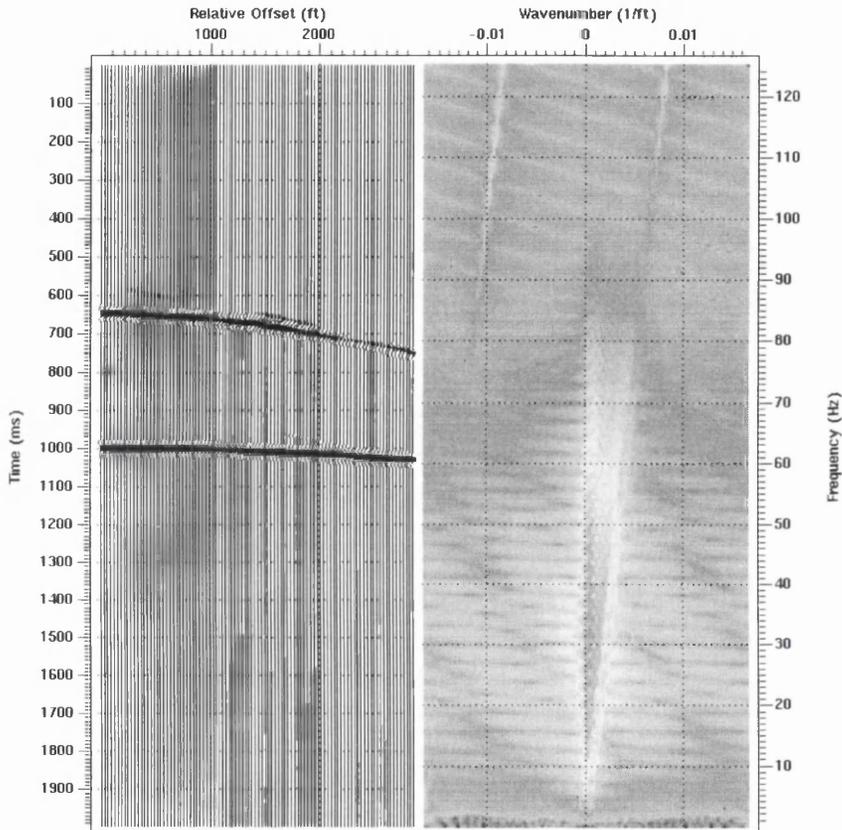
Figure 6-9 CMP gather from Figure 6-7 after filtering using the maximum resolution criteria with the corresponding f-k spectrum. The reflector signals have been retained to a greater extent than in Figure 6-8.
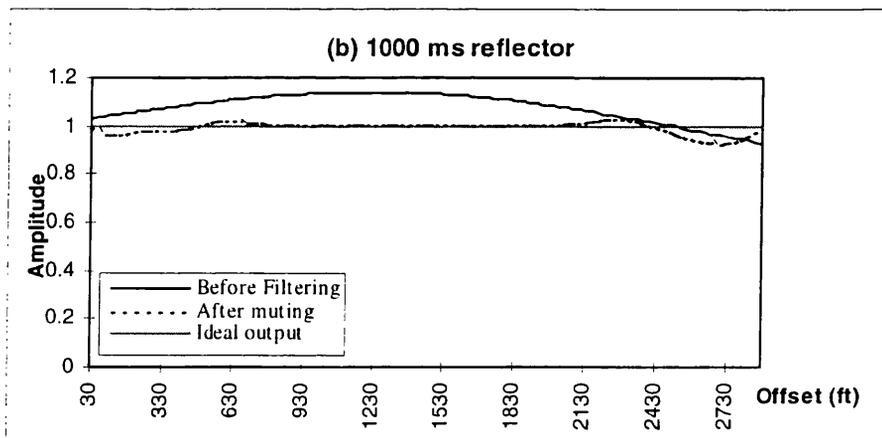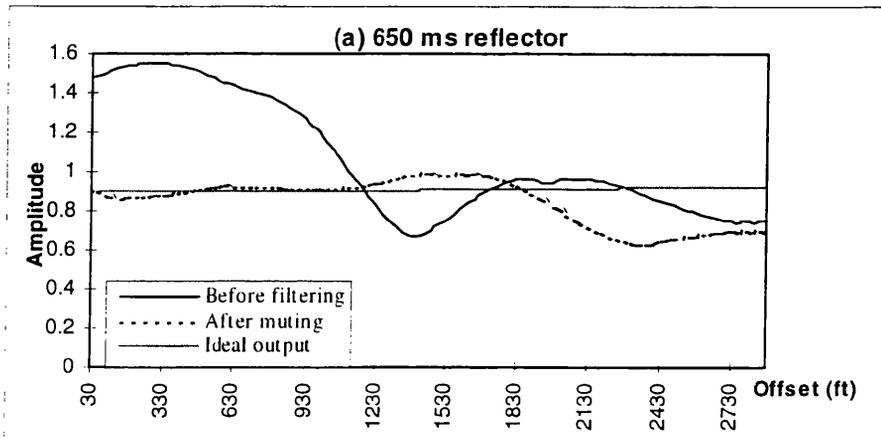
Figure 6-10 Amplitude versus offset relationship for the (a) 650 ms and (b) 1000 ms reflector before, after filtering and compared with the ideal output.
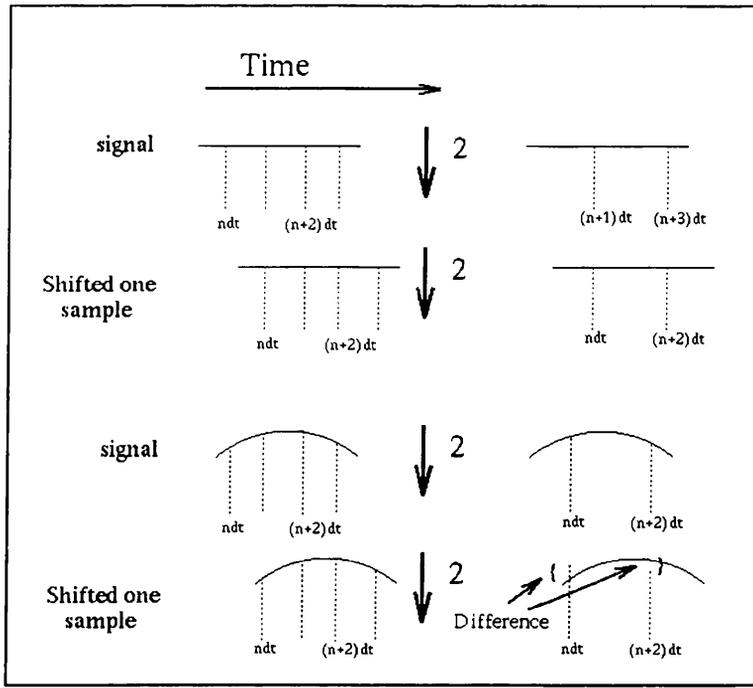
Figure 6-11 A schematic explanation of how the downsampling process in the discrete wavelet transform results in differences in transform when identical signals are offset by one sample.

Figure 6-12 shows the CMP gather and corresponding *f-k* spectrum after filtering using the weighting technique described previously. **Figure 6-13** shows the effect on reflector amplitudes. The resulting amplitude distortions have been reduced considerably. A drawback of this technique, however, is that the wavelet transform has to be applied multiple times to each shifted time sample, increasing the processing time of the filter. Alternatively, the fast algorithm described by Beylkin (1992) can be used, which when applied to a signal calculates the value of the wavelet transform of a signal at all translations. This would reduce computational times considerably. The design of these filters is subsequently governed by the choice between speed of processing and amplitude preservation. In seismic records where there is signal overlap in *f-k* space, the slower weighting procedure is required to prevent distortion of amplitudes. Where the signal to noise moveout difference is large, and so there is good separation in the *f-k* domain, filtering can be performed by muting at a suitable scale in the wavelet domain. This is also the case where amplitude preservation is not the overriding factor. This filtering technique is effectively a local filter and so can be implemented in a time varying fashion, adapting to the nature of the noise at specific times.

6.5 Filtering Examples: Real Data

The performance of filters has so far been tested only on synthetic data. The next step in performance evaluation is to apply the filters to real seismic data.

Figure 6-14 shows two typical shot records from a land-based seismic survey shot over the Longannet coalfield in Eastern Central Scotland (data courtesy of IMCL). The survey was shot using an explosive source into ninety receiver stations of six geophones deployed in a linear array . The group spacing was 12m and shot spacing was 15m. The first record was shot with end on geometry and the second with a split-spread geometry. The shot records have been muted to remove any signal that will not contribute to the stack, that is, data that had greater than 30% NMO stretch was muted. The filtering technique relies on trace to trace continuity of events and so elevation and maximum power autostatics based on the brute stack have been applied to the shot records before filtering.
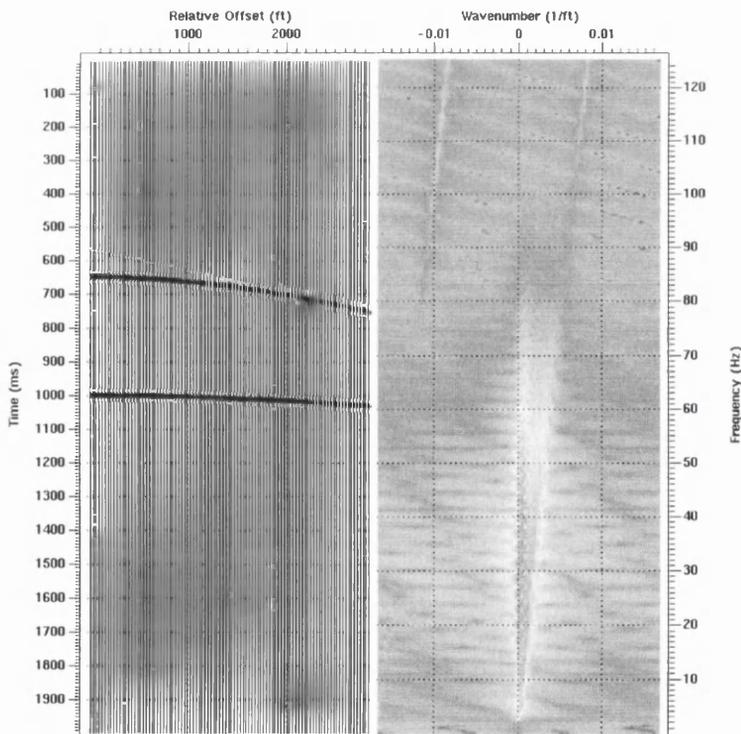
Figure 6-12 The synthetic CMP gather from Figure 6-7 after filtering using the weighting procedure and the corresponding f-k spectrum. The reflector signal has been preserved more accurately.
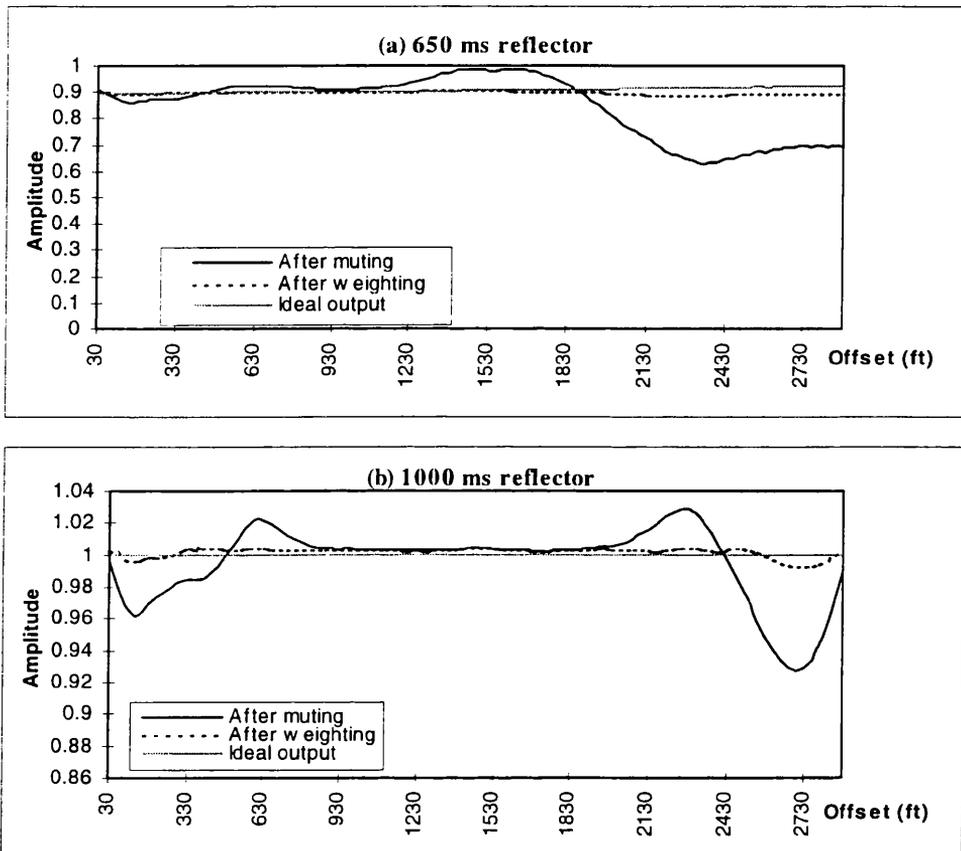
Figure 6-13 Comparison of the AVO relationships after filtering using the weighting and muting techniques with the optimum ideal output. The weighting procedure minimises any distortions resulting form the filtering process.
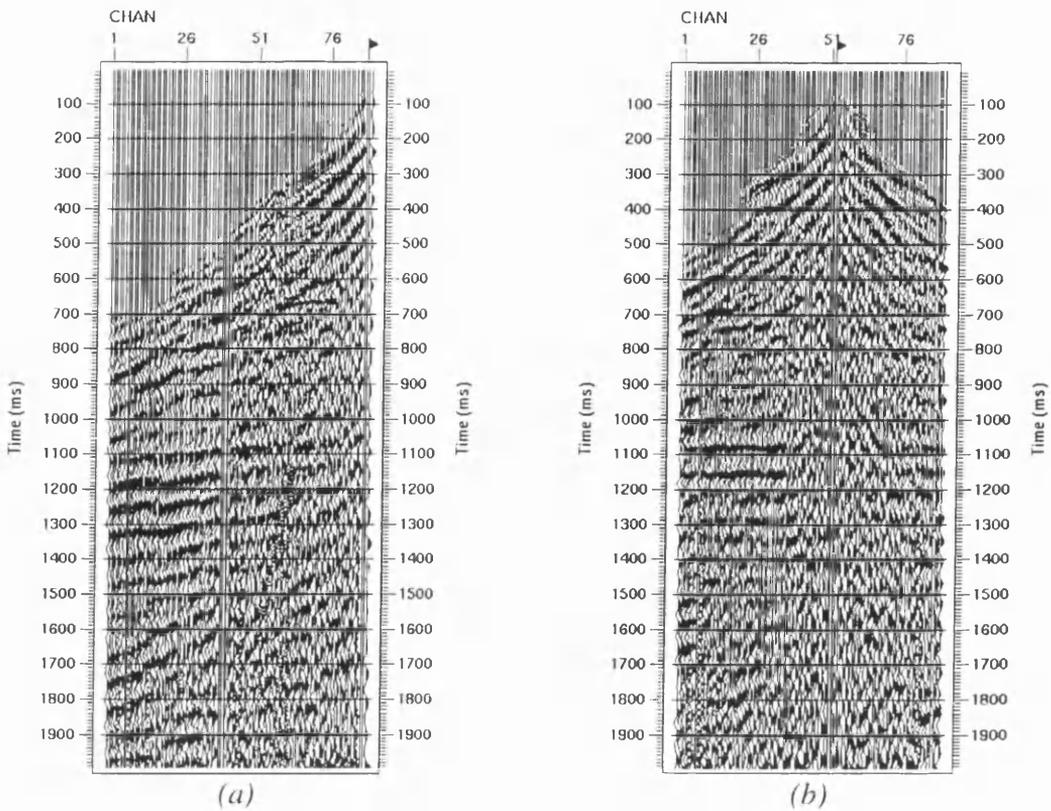
Figure 6-14 Two shot records (a) end-on and (b) split-spread from a survey over the Longannet coalfield. A 500 ms AGC has been applied to enhance reflectors. The records are heavily contaminated by linear noise.

From inspection of the shot records it is apparent that there is a great deal of linear noise present in the record, especially at 0-800 ms and 1300-2000 ms intercept times **Figure 6-14a** and 0-800 ms on **Figure 6-14b**. The moveout difference between the linear noise and the target reflection events was large, allowing the mute filtering of scaling coefficients in the wavelet domain after decomposition to the fifth scale. The filtered records are shown in **Figure 6-15** where an improvement in the shot records can be clearly observed. **Figure 6-16** shows difference sections for these shots at the same scale. In both records, reflections between 600-900 ms have been enhanced and the continuity of deeper reflections improved. From the filtered records we can see that the velocity of the target reflectors (now more apparent) seems fast enough to justify the muting of coefficients rather than weighting and so no redesigning of the filter was required.
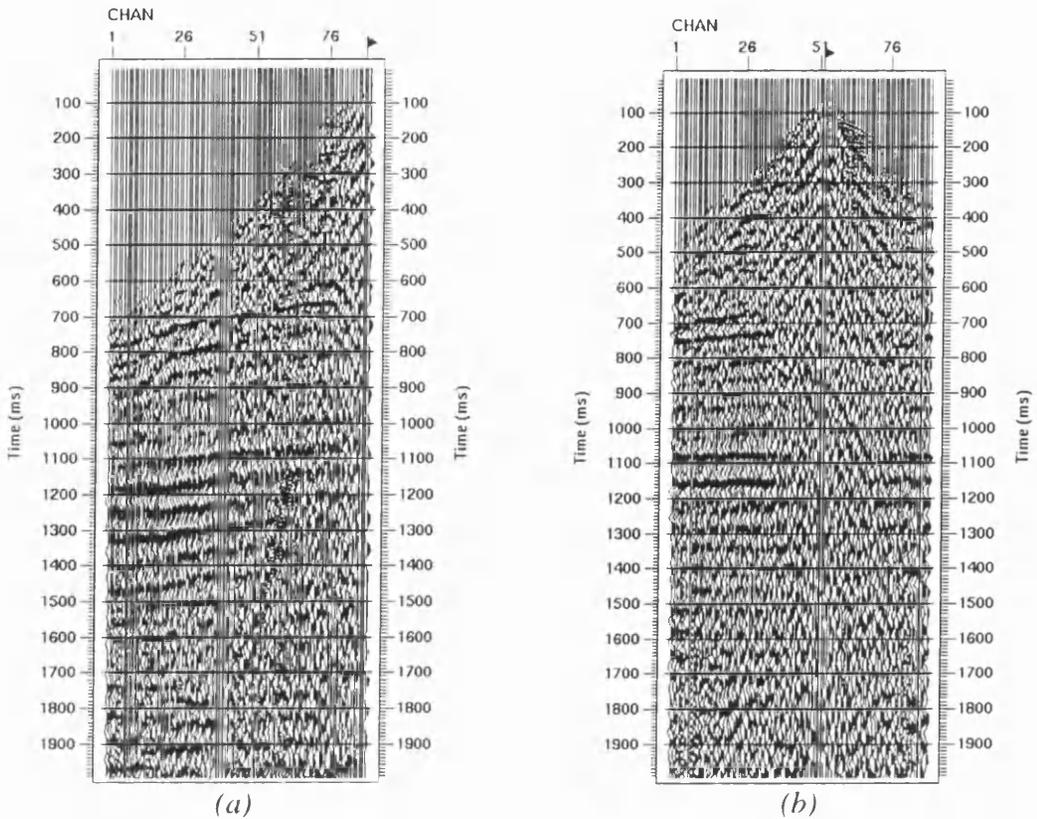
Figure 6-15 Shot records shown in Figure 6-14 after filtering using the wavelet transform technique. A 500 ms AGC has been applied to enhance reflectors. The filtering consisted of muting coefficients in the wavelet domain.

Figure 6-17 shows brute stacks of the corresponding seismic line before and after filtering using the wavelet transform technique. Continuity of deeper reflections (below 1300 ms) has been improved and the continuity of reflections above 600 ms has also been improved. **Figure 6-18** shows the difference section for the two stacks. The linear noise removed from the sections is apparent. Some shallow reflection energy is also apparent in the difference section (above 400 ms), indicating that reflected energy was removed by the muting process in the wavelet domain. This would indicate that the weighting process is required at earlier travel times. Nevertheless, we can conclude that this is a powerful technique for suppressing linear noise from seismic records. The two stacked sections were processed using the same processing flow, apart from the wavelet transform filter. The filtering technique will further improve stacked sections by subsequent improvements in velocity analysis and static solutions for any data set.
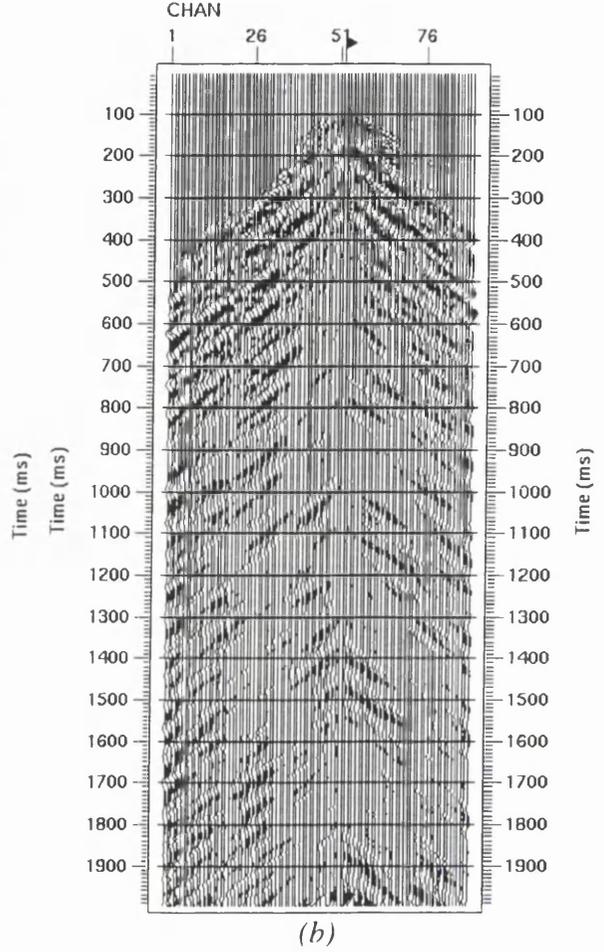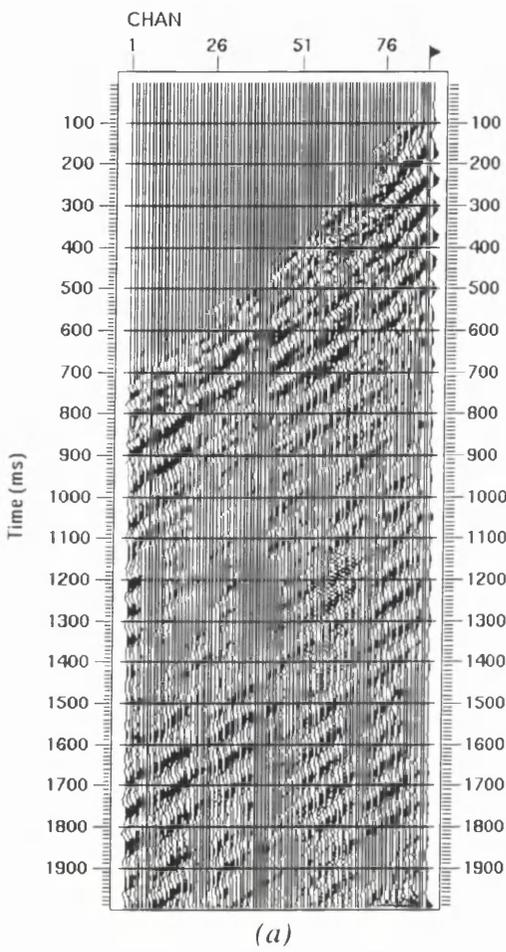
Figure 6-16 Difference sections between Figure 6-14 and Figure 6-15 showing the linear noise removed from the shot records. The scaling of these records is the same as for the previous Figures.
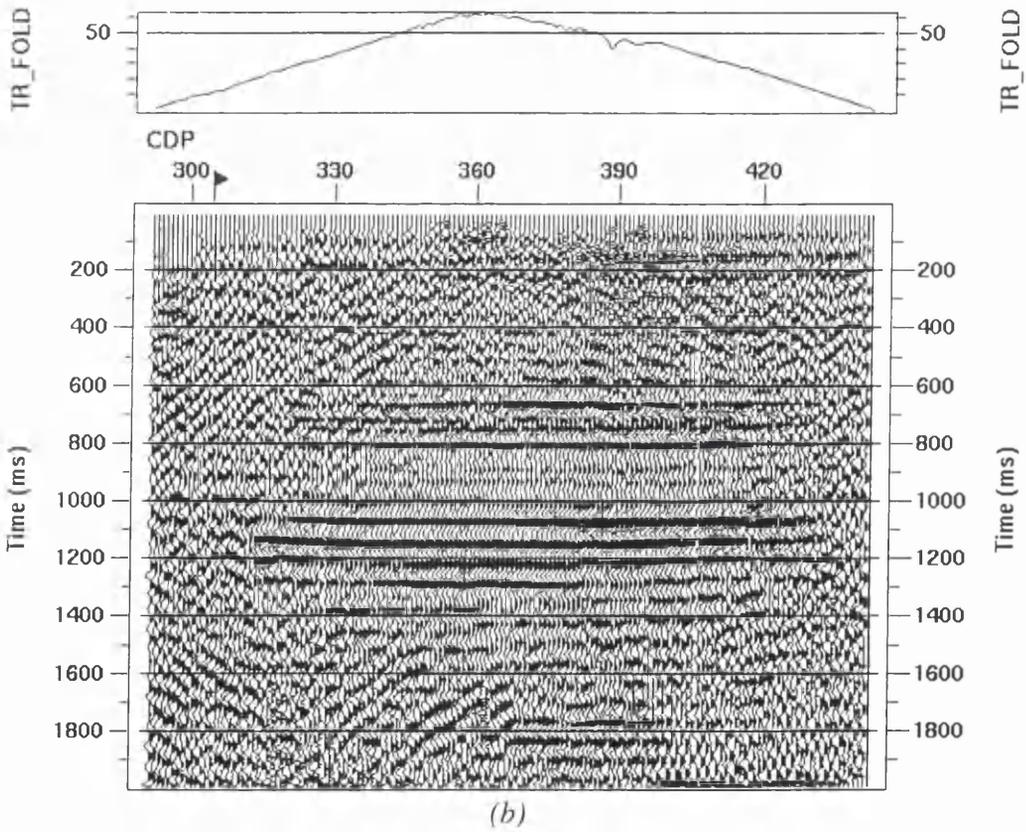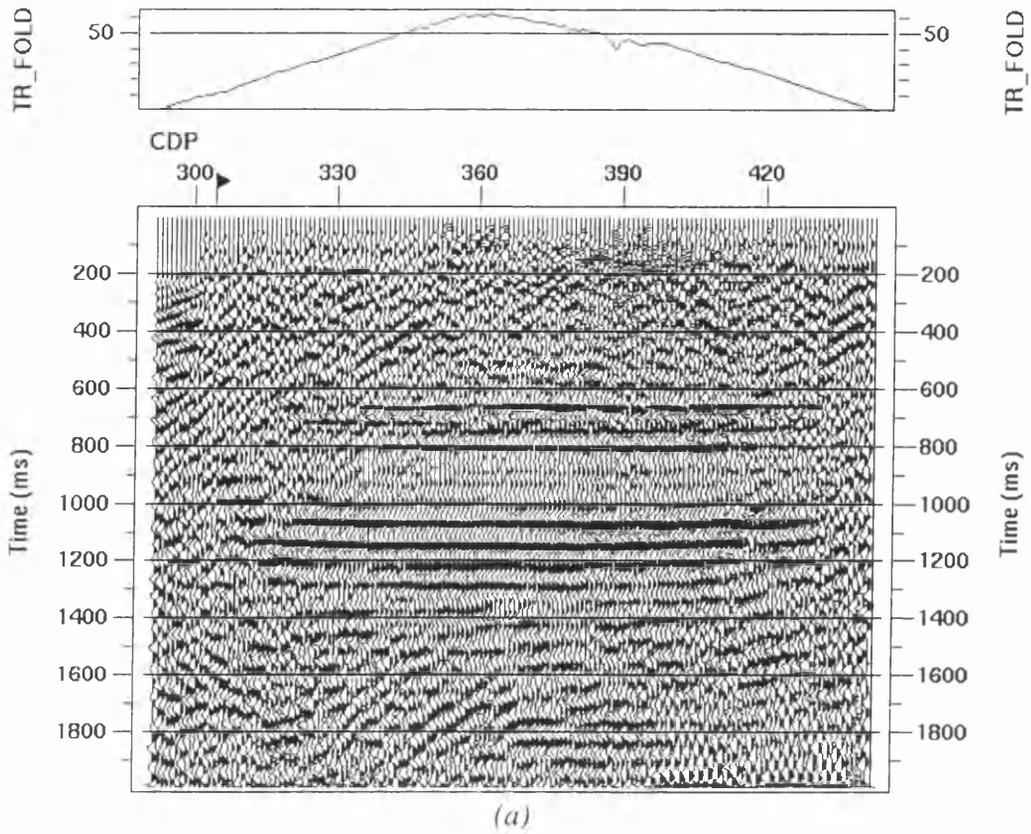
Figure 6-17 Brute stacks of the seismic line over the Longannet coalfield (a) without wavelet transform filtering and (b) with wavelet transform filtering of the shot records.
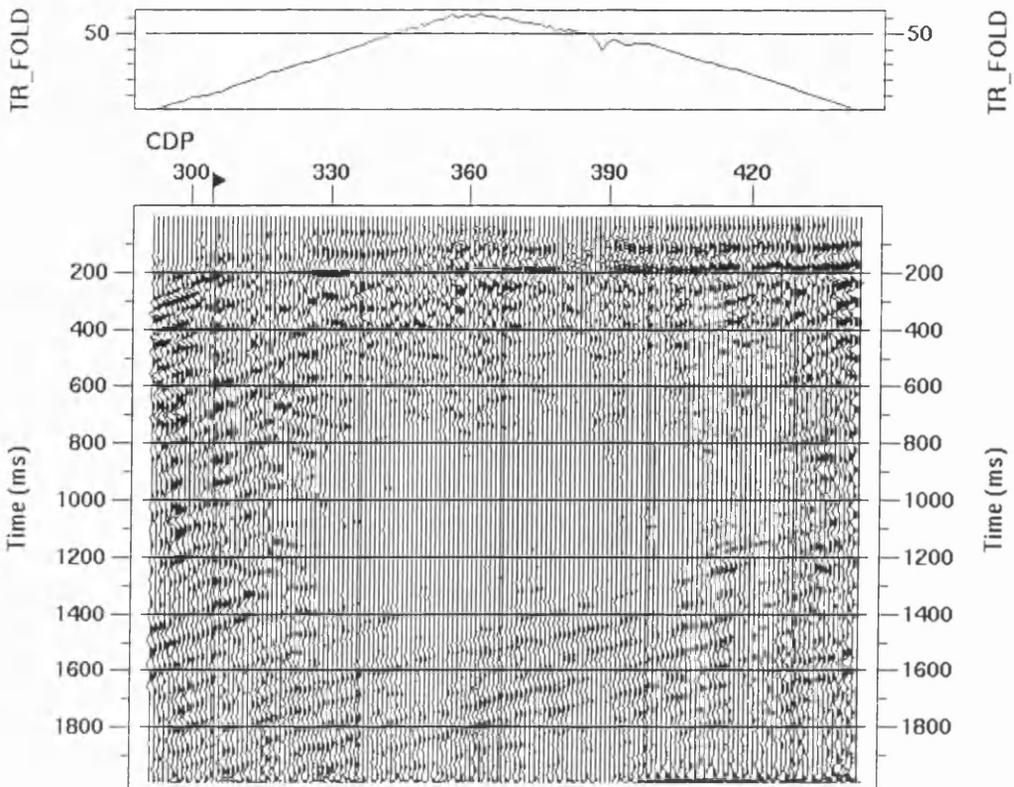
Figure 6-18 Difference section between Figure 6-17 (a) and (b) showing the noise removed from the stack by the filtering process.

6.6 Swell Noise Suppression

A further application of this technique is the suppression of swell noise from marine common shot records. Swell noise appears on shot records when sea conditions deteriorate and wave action manifests itself on shot records. The hydrophones, which under calm conditions would be at a constant depth with respect to the sea surface, detect changes in pressure due to wave action which change the effective depth of the hydrophone. This change in depth leads to a corresponding change in pressure which is detected by the hydrophone. The frequency and amplitude of this pressure change due to the wave action is dependant on wave speed, amplitude, angle of incidence with repect to the hydrophone streamer, and wavelength. This is indicated schematically in **Figure 6-19**. On a common shot gather, this pressure change will appear as a linear seismic event as it travels across the streamer (assuming that there is no streamer feathering). Therefore we can consider swell noise to be the superposition of linear events on a shot record which is also represented schematically in **Figure 6-19**.

Figure 6-20 shows a marine common shot record contaminated by swell noise from which the superposition of linear events is apparent. This is also demonstrated in

Figure 6-21 which shows the linear events on the record more clearly after low pass and high pass filtering. **Figure 6-22** shows the record after removal of some of the linear events, and an improvement in the shot record is apparent. The quality of the shot record is still degraded by the presence of very steeply dipping linear events which are very highly spatially aliased. These highly spatially aliased events, after linear moveout, leave a signal with very few samples in the offset direction. Therefore, to remove these events the wavelet transform technique must be adapted to handle signals with very few samples and any subsequent amplitude effects investigated. This work is planned for future developments. The initial removal of a subset of the linear events has shown that the technique is viable, the next stage in development is the adaption of the technique to the harder task of removing the highly spatially aliased noise.

(a)

(b)
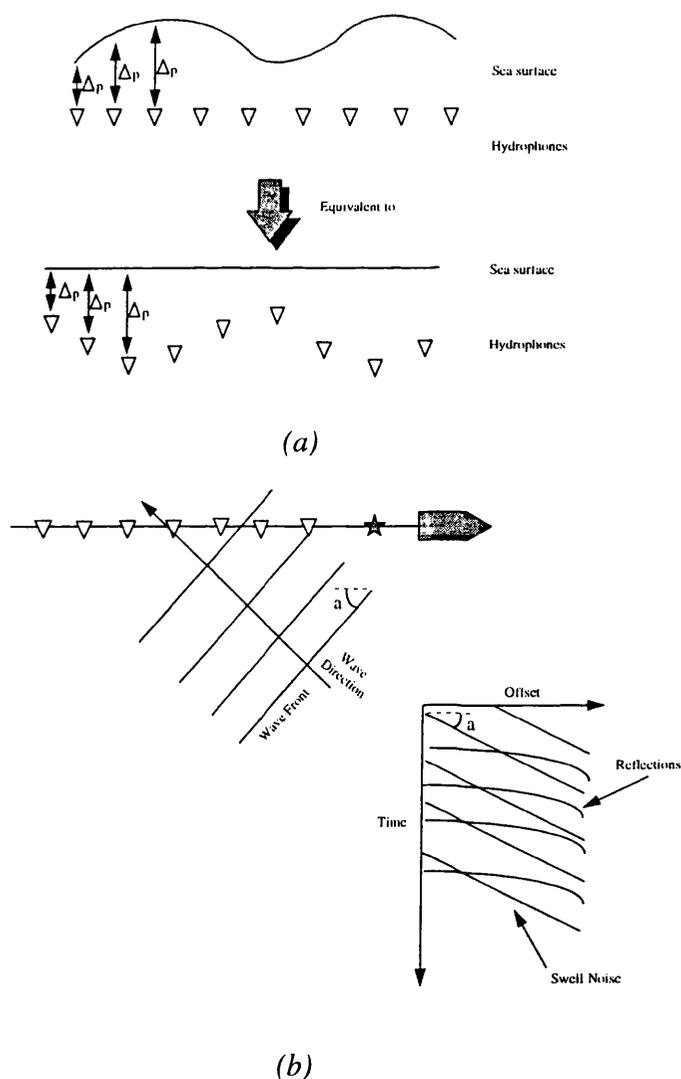
Figure 6-19 Scematic diagram showing (a) the generation of swell noise on marine records caused by changes in hydrophone depth due to waves on the sea surface and (b) the resultant pressure wave then manifests itself on seismic records as a linear event, the velocity of which is dependant on the angle of incidence of the wavefront.
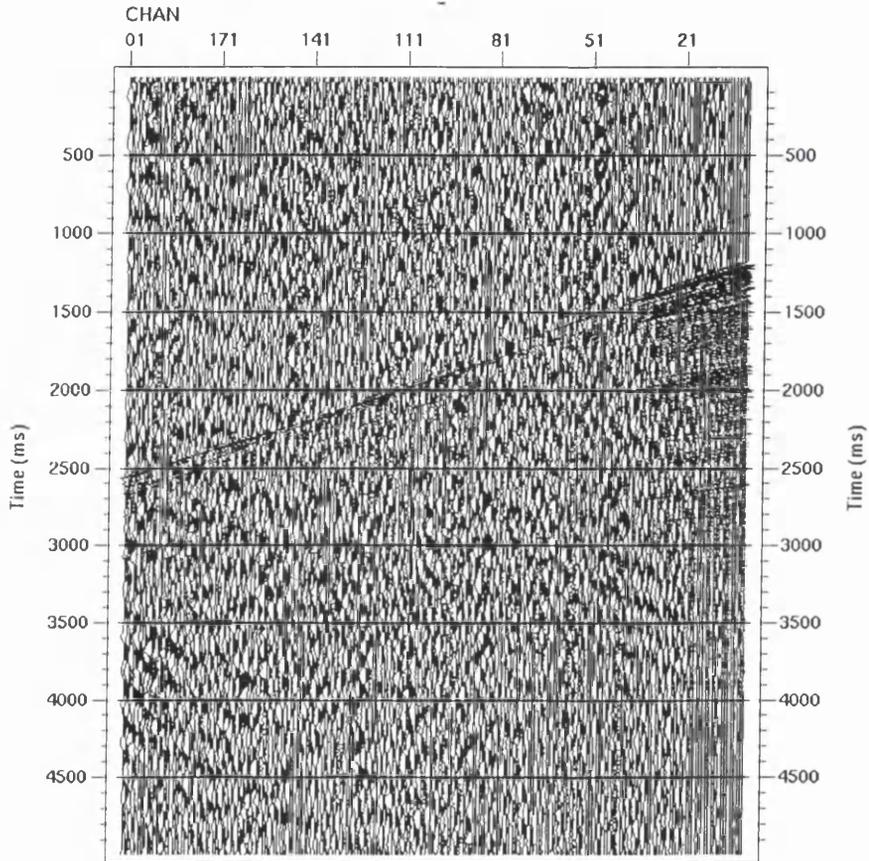
Figure 6-20 A common shot record from a marine survey contaminated by swell noise which obscures most reflected signals.
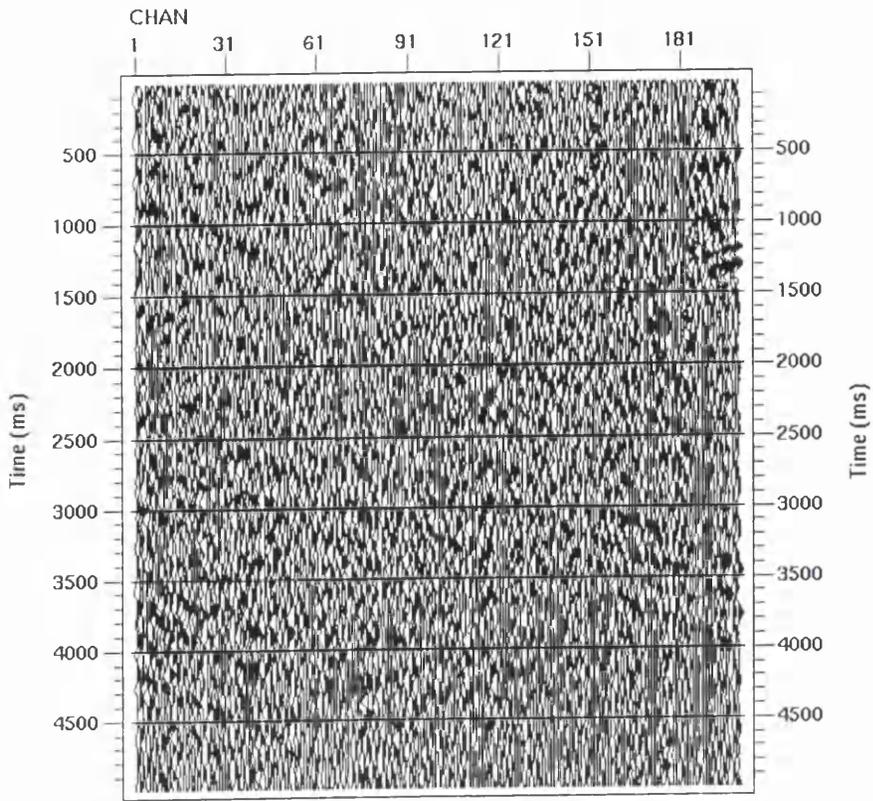
(a)
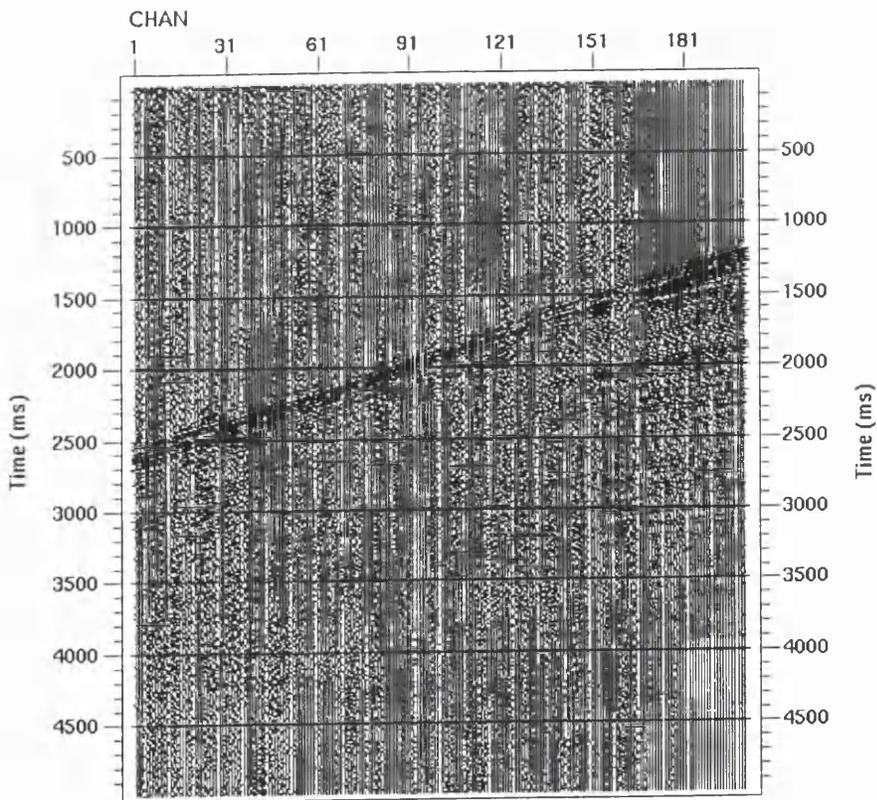
(b)

Figure 6-21 Figure 6-20 after (a) a low pass filter and (b) a high pass filter showing that the swell noise is the superposition of several sets of linear noise, which in (a) are gently dipping and in (b) are very steeply dipping and so very highly spatially aliased.
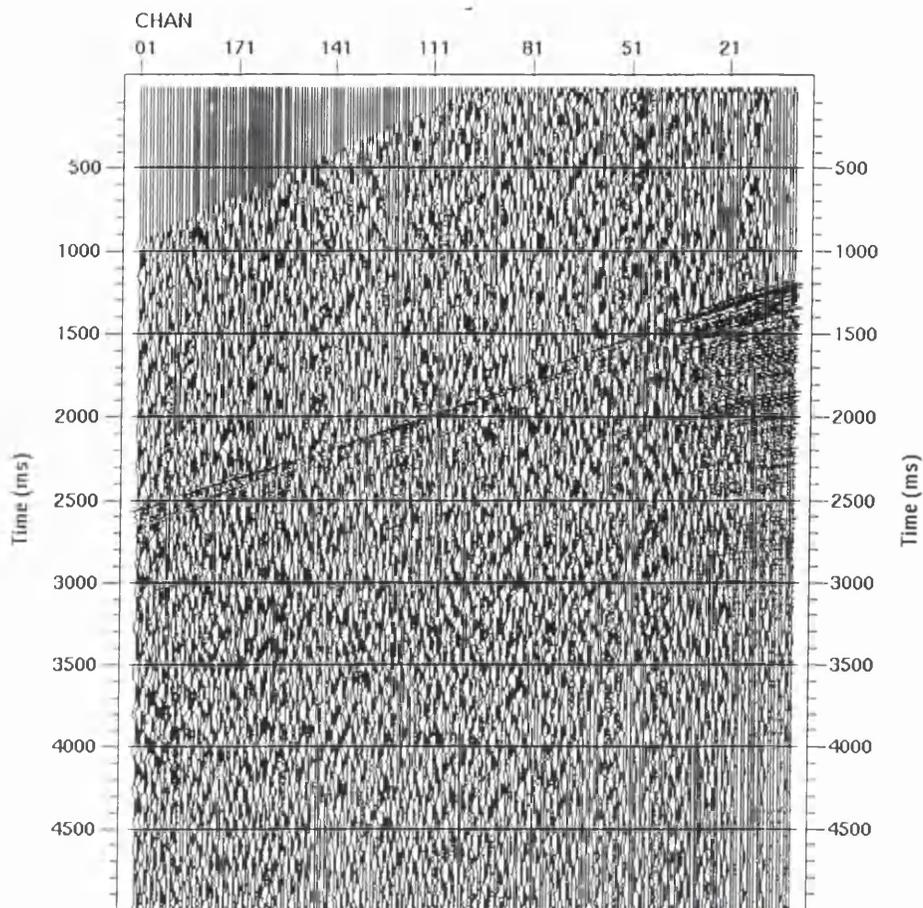
Figure 6-22 Figure 6-20 after filtering with the wavelet transform linear velocity filter to suppress some of the swell noise. The gently dipping noise has been suppressed. The removal of the steeply dipping noise will need further investigation.

6.7 Hyperbolic Filtering of Seismic Data: Multiple Suppression

In this section the wavelet transform technique has been adapted by replacing the linear moveout with a normal moveout. This technique has applications in the suppression of multiples from CMP records. The modified technique will be evaluated on a synthetic model (**Figure 6-23**), taken from Alvarez and Larner (1996), and will also be evaluated on marine data at the CMP stage and the resultant stack analysed.
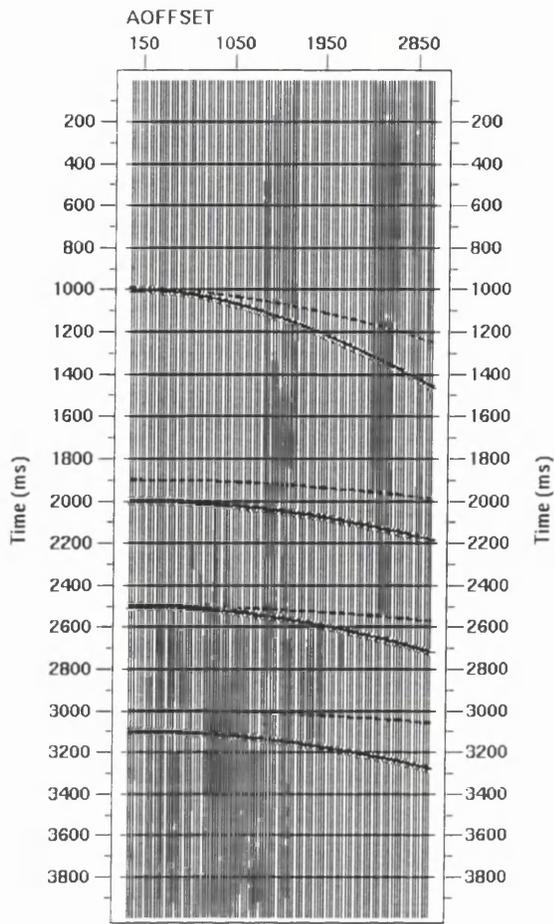
Figure 6-23 The synthetic model used for evaluating multiple suppression technique, taken from Alvarez and Larner (1996). The velocities are given in Table 6-1.

Alvarez and Larner (1996) compared the suppression of multiples using the *f-k* transform to the radon (τ-*p*) transform approach of Hampson (1986), and a hybrid approach that combines Hampson's method with a variant of the statistical method of Harlan *et al.* (1984) which discriminates between focused energy and defocused energy in the τ-*p* domain. The results of this study will be compared to the results published in Alvarez and Larner's extended abstract. The model consists of four primary events and four multiple events. Two of the multiples are coincident with the primaries and the other two are not. The former are intended to show contamination by residual multiples and the latter to show distortion of primary after attempts at multiple suppression. The amplitude of reflection with offset will be varied to form three models. The first will have a primary to multiple amplitude ratio of 4:1. The second model will have a primary to multiple ratio of 1:1 and the final model will be the same as the second with the addition of a linear decrease in amplitude with offset such that the far trace has half the amplitude of the near trace. The rms stacking velocities for the events are given in **Table 6-1**. The gather consists of 100 traces with 30 m trace spacing. The primary and

multiple events consist of Ormsby wavelets with 8-12-50-70 Hz corner frequencies. These parameters were kept constant for each model.

	Time (ms)	Velocity (m/s)	Amplitude Model 1	Amplitude Model 2	Amplitude Model 3
Reflector 1	1000	4000	400	100	100-50
Multiple 1	1000	2800	100	100	100-50
Reflector 2	1900	5000	400	100	100-50
Multiple 2	2000	3400	100	100	100-50
Reflector 3	2500	5000	400	100	100-50
Multiple 3	2500	2800	100	100	100-50
Reflector 4	3000	5000	400	100	100-50
Multiple 4	3100	2800	100	100	100-50

Table 6-1 Parameters for the synthetic CMP models used in study of multiple suppression.

Extending the linear filtering technique to the suppression of multiples simply replaces the linear moveout with a normal moveout which is used to flatten the multiple arrivals. These are then suppressed by applying the wavelet transform filtering technique, and an inverse NMO applied to restore the data. Compared to the first linear model, the moveout difference between any multiple and primary arrival is very small and so simple muting of coefficients in the wavelet domain will not be sufficient, therefore, weighting will be necessary. When this technique is used for multiple suppression, the problem of distortion of reflectors observed at larger offsets with the linear filtering will occur at near offsets where it is essential to preserve the amplitudes of the primary events.

The synthetic gather for model one, after the application of NMO to flatten primaries, is shown in **Figure 6-24a**. The gather was initially filtered by applying a NMO corresponding to the multiple rms velocities, applying the wavelet transform in the shifted offset direction and weighting the scaling coefficient. The transform was then inverted and the NMO removed. The result of the filtering process is shown in **Figure 6-24b** where we can see that, apart from some residual energy at large offsets, the lower three multiples have been suppressed effectively, while at the shallowest reflector a large amount of multiple energy remains. This energy is due to NMO stretch

associated with flattening of the multiples. **Figure 6-25** shows the shallowest reflector after the application of the NMO and it is apparent that the time width of the event increases with offset. Where the multiple event has aligned the event perfectly (that is where the amplitude is consistent across the traces i.e.: at the peak amplitude) the filtering technique has worked effectively, yet, where the data is not perfectly aligned due to seismic wavelet stretching, multiple energy has been retained. To minimise this residual energy problem for the first reflector, the scaling coefficient and lowest scale of the wavelet transform for the data was muted for the first multiple. The residual energy for the other multiples is minimal and will be effectively suppressed by the stacking process and so the weighting process was applied to suppress the remaining three multiples. **Figure 6-26** shows the result of this filter. From this we can see that the remainder of the multiple energy has been reduced considerably.

Figure 6-27 and **Figure 6-28** show the synthetic gathers after NMO for models two and three and the corresponding result of the filtering process. For models two and three the filtering parameters were the same as for model one. From this we can see that, for these models, the multiple suppression technique was successful.
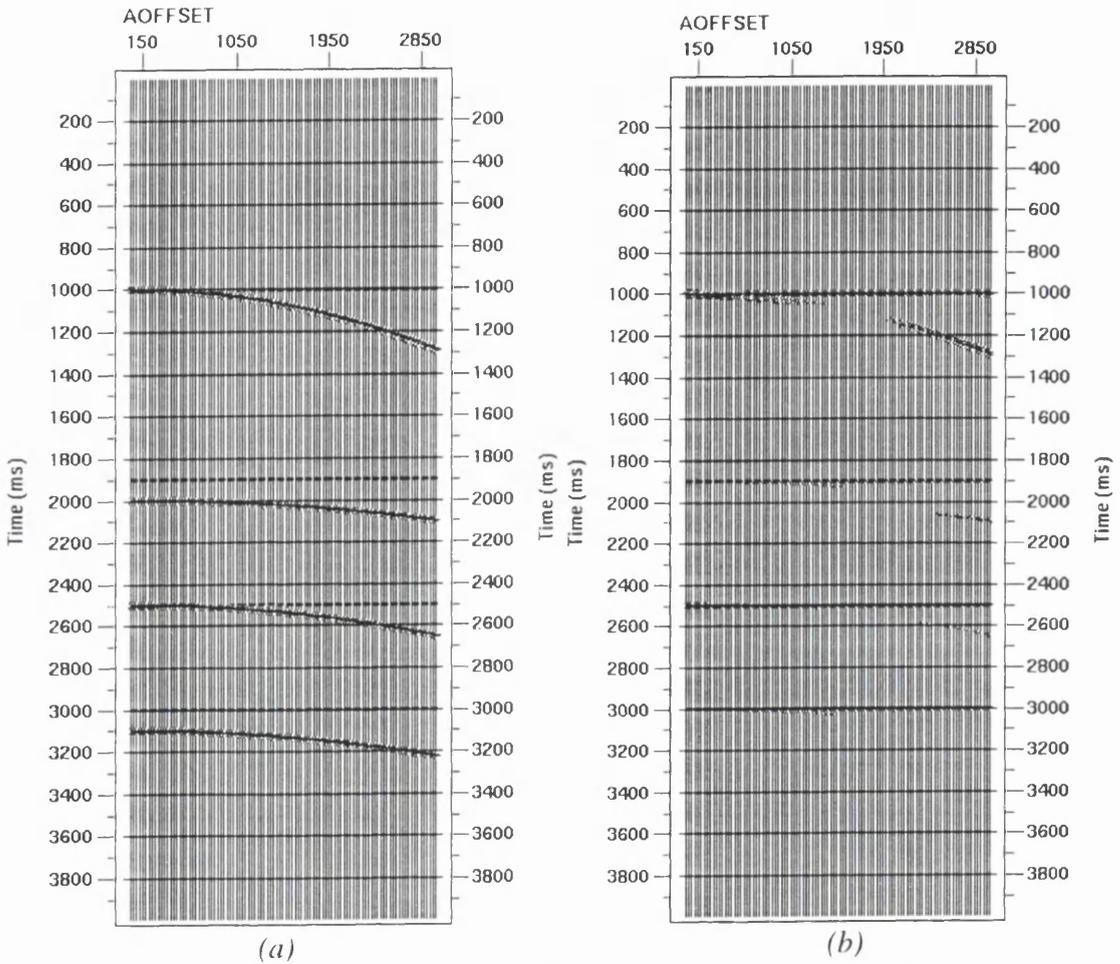
Figure 6-24 The synthetic gather from Figure 6-23 after (a) NMO corresponding to the primary velocity function and (b) filtering by the wavelet transform technique followed by NMO.
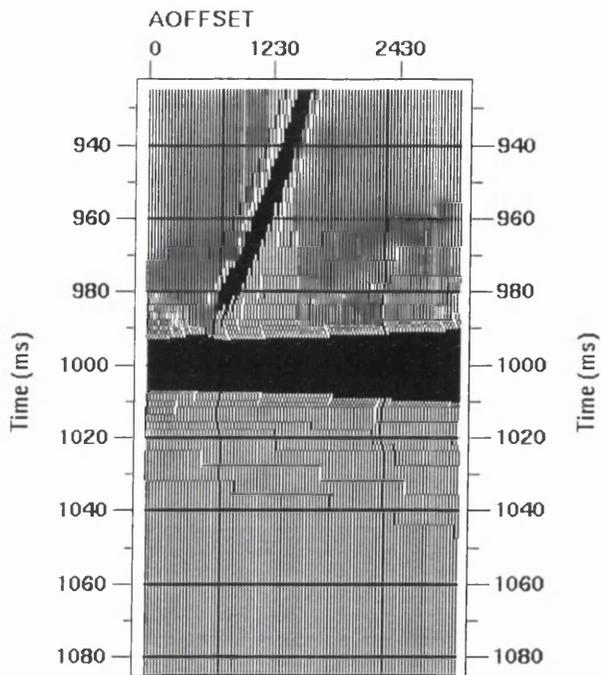
Figure 6-25 Reflector 1 at a larger scale showing stretch of the multiple seismic wavelet with offset caused by the application of the NMO.
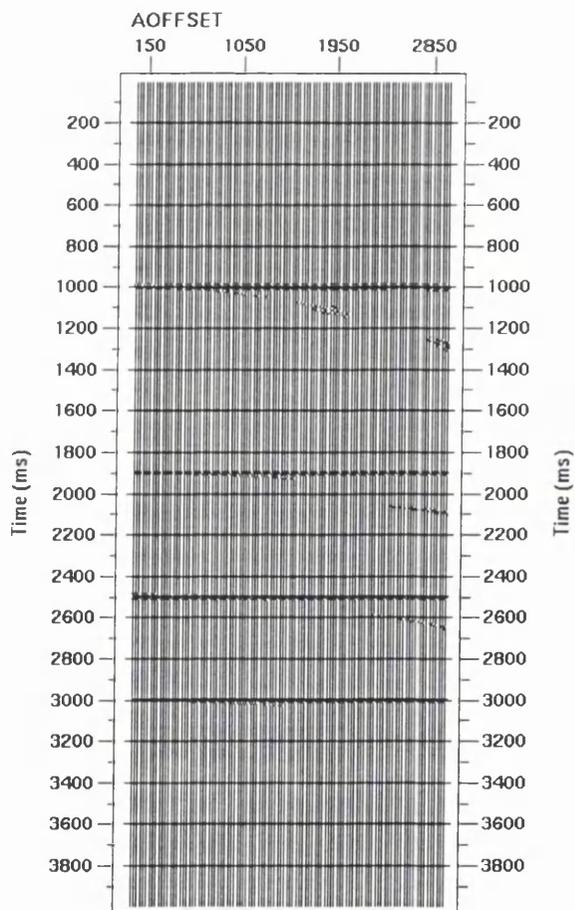
Figure 6-26 The synthetic gather shown in Figure 6-23 after filtering in the wavelet domain by a combination of muting (reflector 1) and weighting (reflectors 2-4) in the wavelet domain.
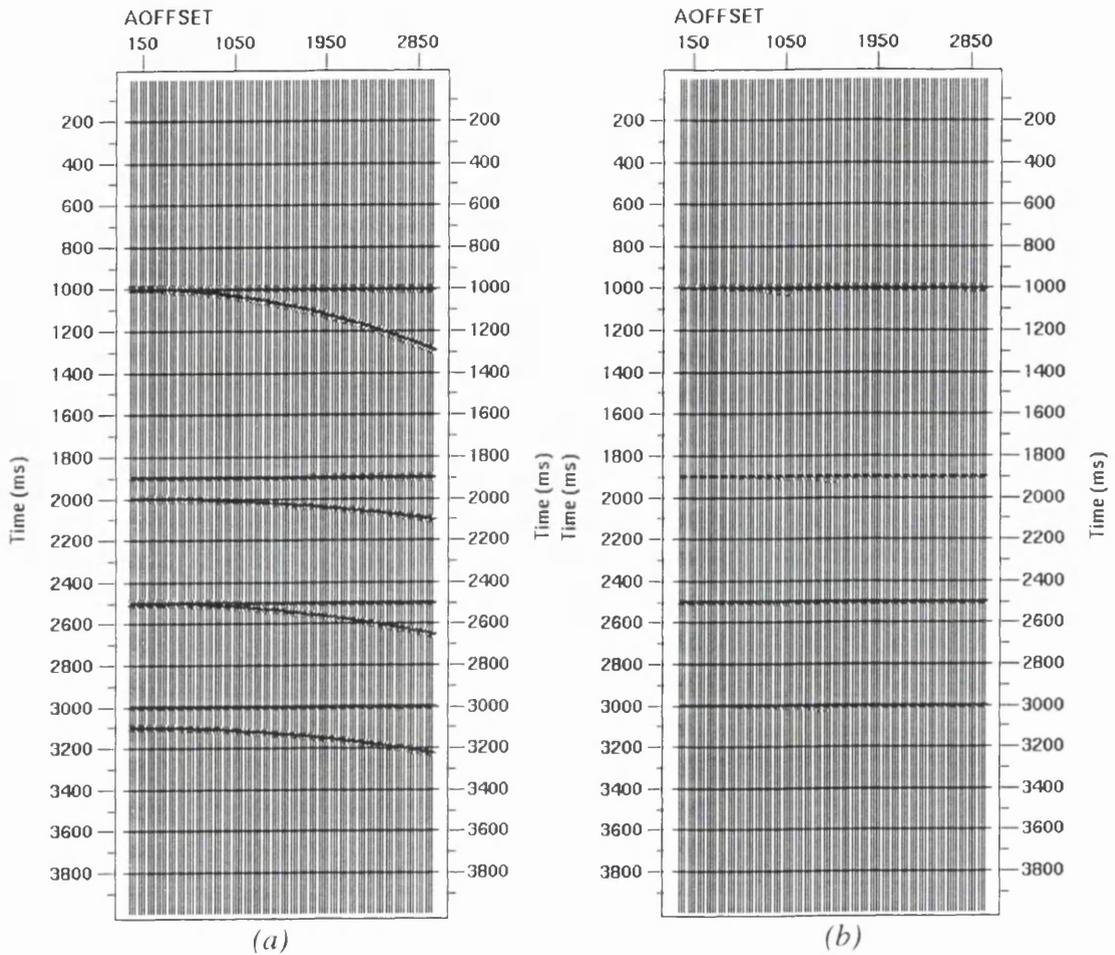
Figure 6-27 Model 2 synthetic CMP gather (a) before and (b) after filtering using the wavelet transform process. The filtering parameters are the same as in Figure 6-26.
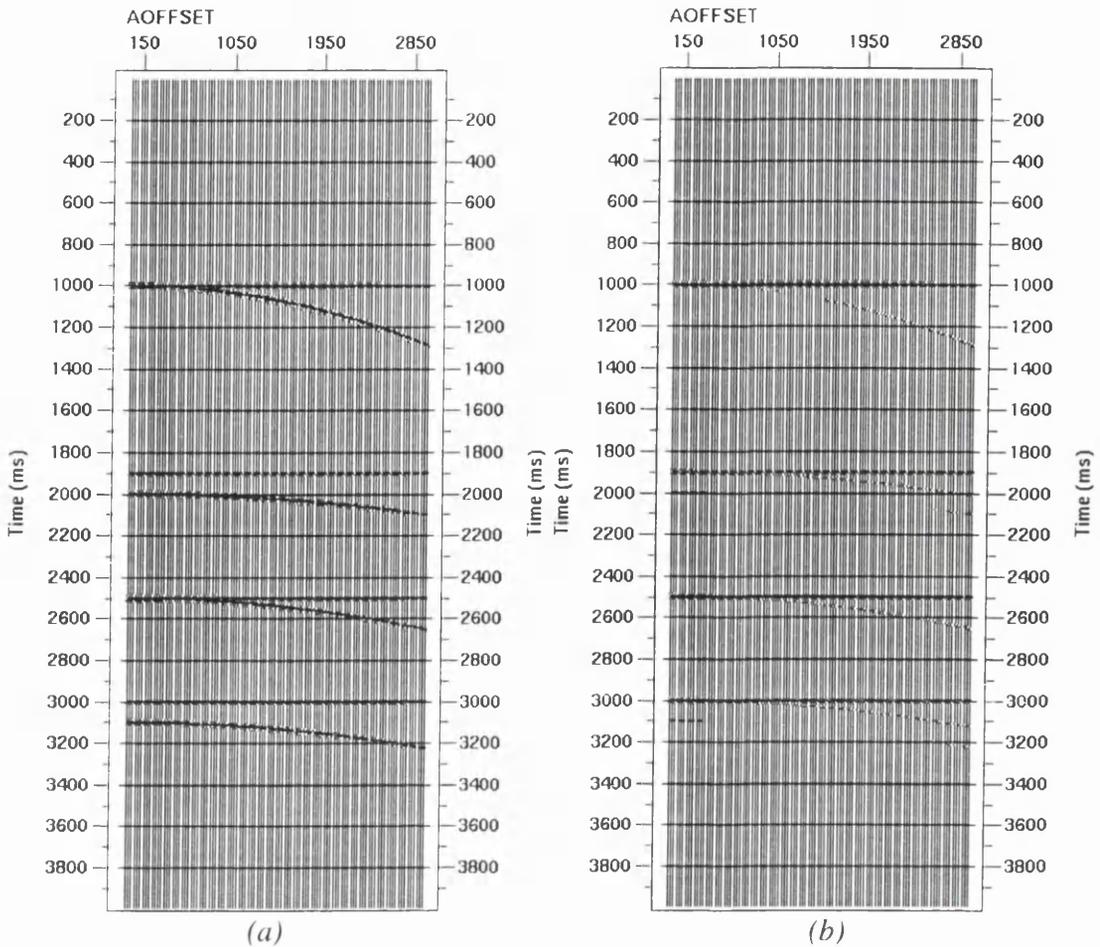
Figure 6-28 Model 3 synthetic CMP gather (a) before and (b) after filtering using the wavelet transform process. The filtering parameters are the same as in Figure 6-26.

6.8 Multiple Suppression: Amplitude Considerations

To compare quantitatively the effects of the filter on the amplitudes of the primary events, the amplitude variation with offset of each reflector before and after filtering has been plotted for model one (**Figure 6-29**) with the ideal output which is the primary reflectors with no multiples. From this diagram we can see that the amplitude is preserved within 72% of the ideal output at offsets less than 1000 m, and within 99% at greater offsets. Comparing **Figure 6-29** to **Figure 6-30** which shows the AVO effects from the study by Alvarez and Larner (1996) we can see that the technique produces better results than the *f-k* and Hampson techniques and results comparable with the hybrid technique at short offsets (< 1000m). At larger offsets (> 1000 m) the result of the wavelet transform process is superior, giving far less distortion than any of the techniques.

Figure 6-31 and **Figure 6-32** shows the AVO results for the remaining two models. From model two we see that the results are similar to those obtained with model one indicating that the primary to multiple amplitude ratio is not important when considering amplitude preservation at short offsets. For model three we see that the amplitudes are preserved to within approximately 80% of the original amplitude. This performance is not as good as for the previous models but, is to be expected when we consider that the weighting technique used was designed to discriminate against events which have reasonably consistent trace to trace amplitude. To improve the performance of the filter on the third model, an alternative weighting technique would be required. When applying this technique to real data, there will be very little prior AVO knowledge and so a weighting technique must be developed that calculates weights based on the data alone, eliminating the constant trace to trace amplitude assumption about the data. This would lead to a more adaptive weighting procedure.
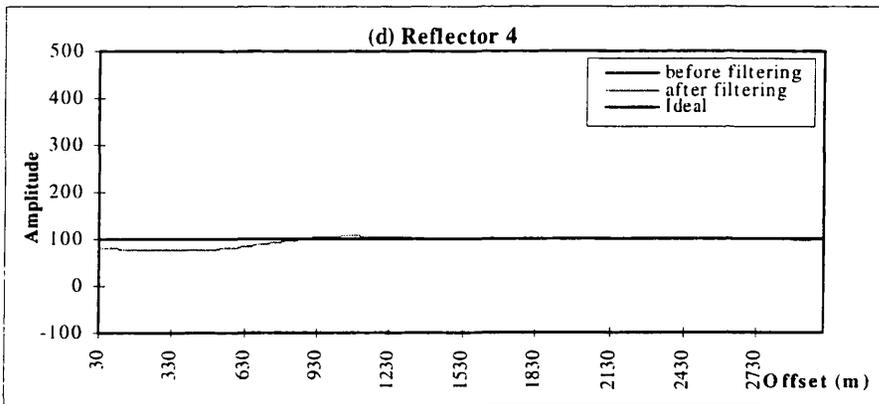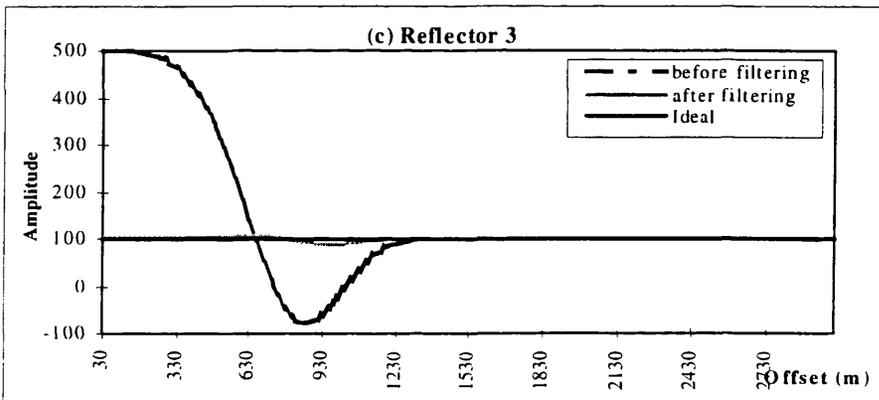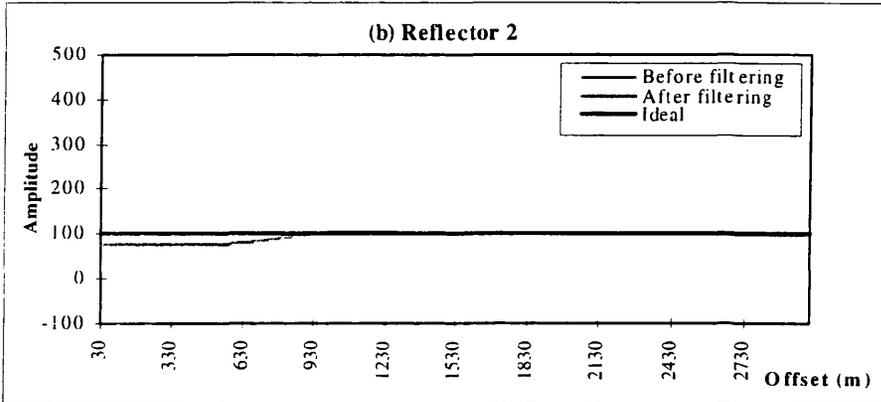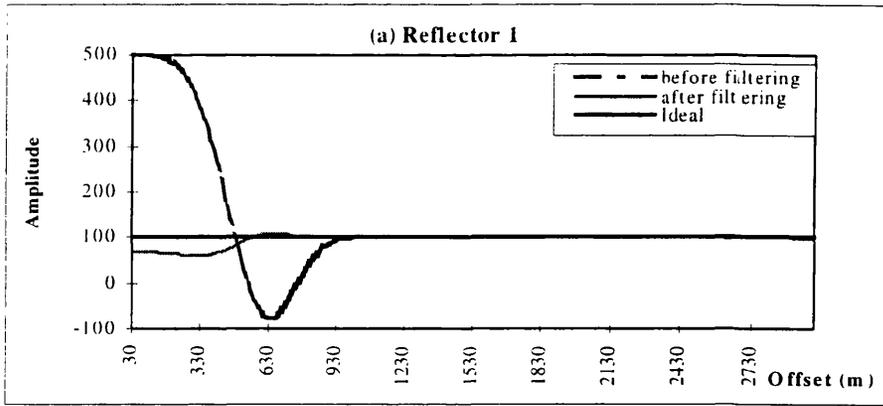
Figure 6-29 AVO relationships for the reflectors in model 1, showing the original signal, the filtered output and the ideal output for (a) reflector 1, (b) reflector 2, (c) reflector 3, and (d) reflector 4.
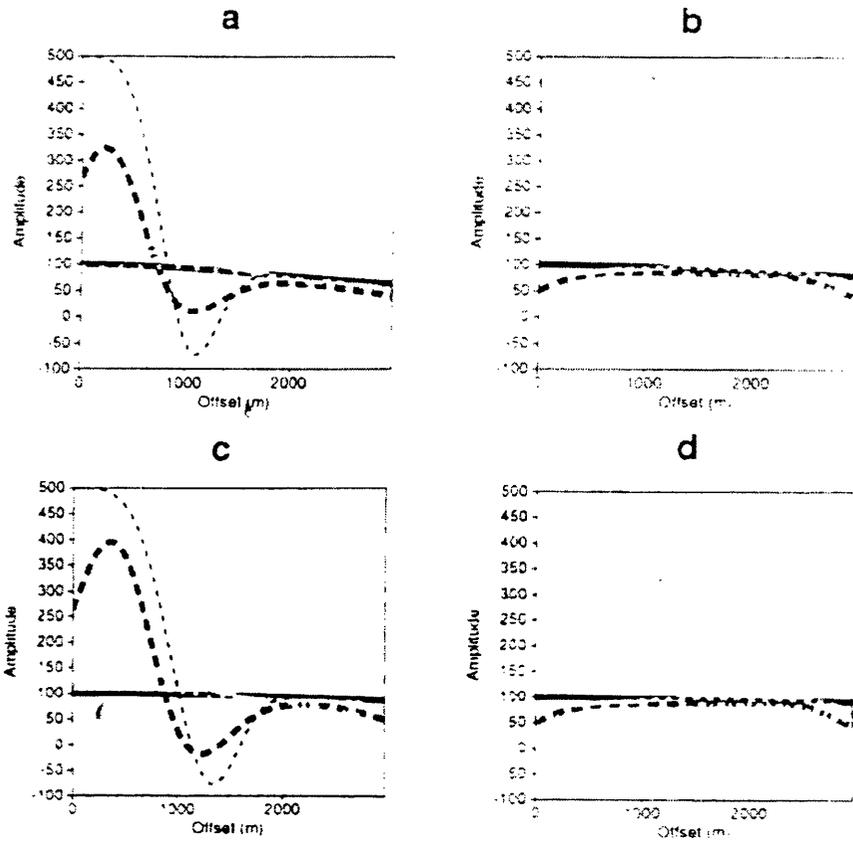
Figure 6-30 AVO relationships for the reflectors in model 1 for other multiple suppression techniques, taken from Alvarez and Larner (1996). (a) Reflector 1 (b) reflector 2, (c) reflector 3 and (d) reflector 4. The thin dotted line is the original unfiltered amplitude, the thick dotted line, the output from f-k multiple suppression, the solid grey line, the output from Hampson's technique and the dotted grey line the output from the hybrid technique. The solid black line is the ideal output
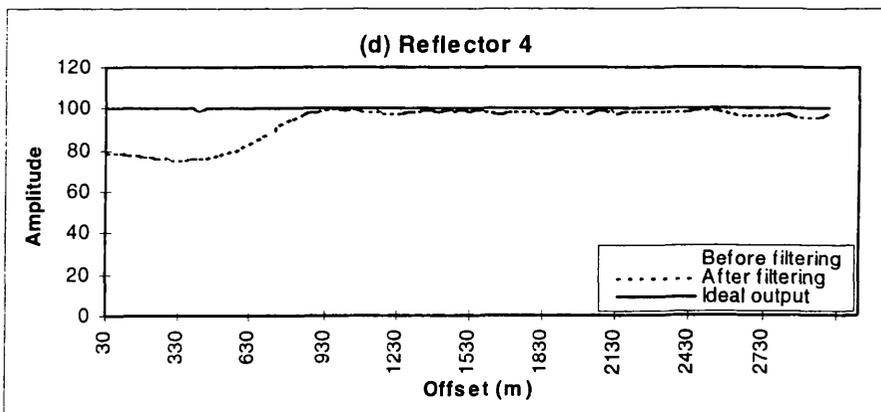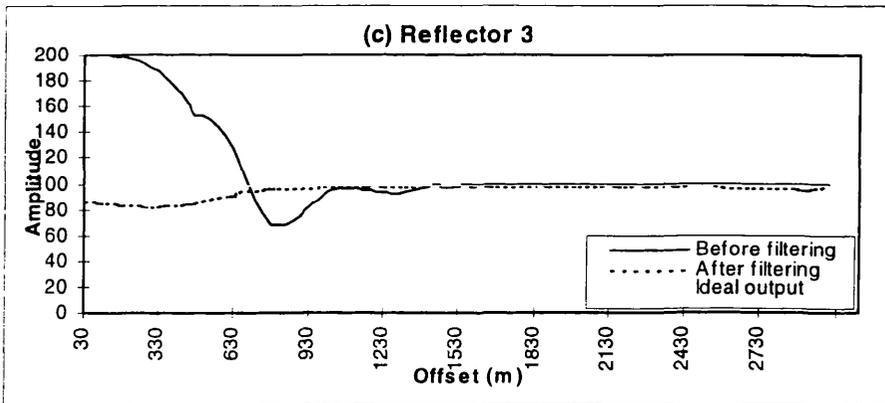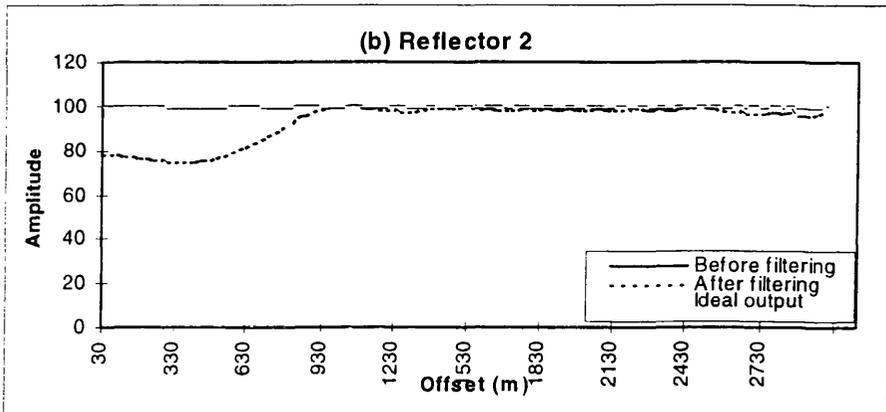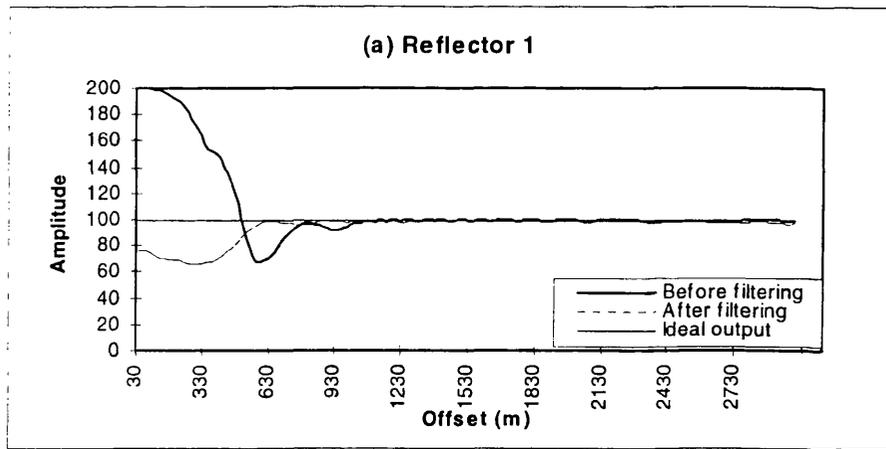
Figure 6-31 AVO relationships before and after wavelet transform multiple suppression for model 2 (a) reflector 1, (b) reflector 2, (c) reflector 3, and (d) reflector 4.
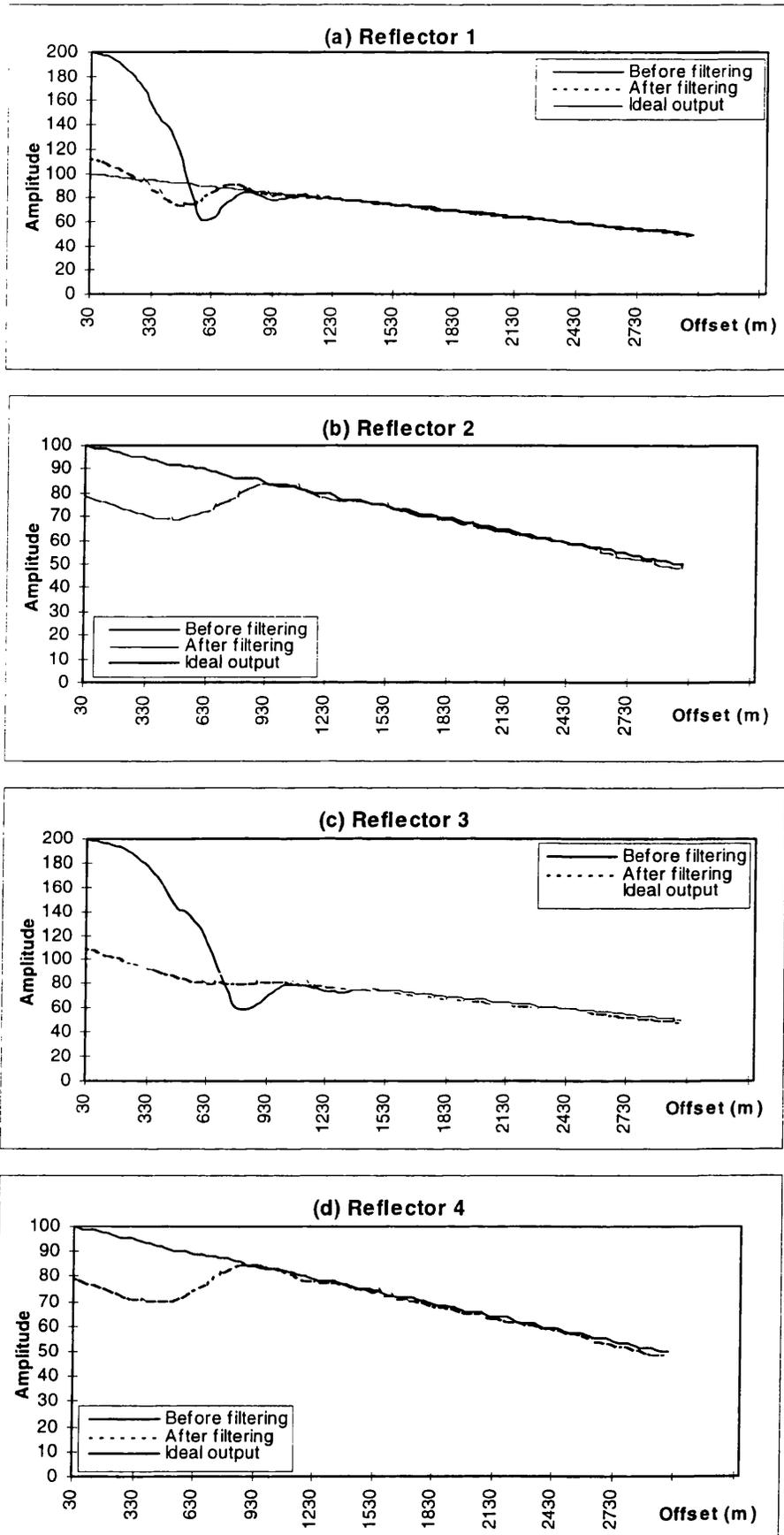
Figure 6-32 AVO relationships before and after wavelet transform multiple suppression for model 3 (a) reflector 1, (b) reflector 2, (c) reflector 3, and (d) reflector 4.

6.9 Multiple Suppression: Influence on the Quality of the Stack

One of the primary goals of multiple suppression is improvement of the stack. **Figure** 6-33 shows stacked traces for the model 1 CMP before and after filtering, and without any multiple energy. From this figure we can see that the technique is effective in reducing multiple energy in the stack. To quantitatively analyse the observed improvement, the ratio of peak to trough amplitudes of primary to multiple energy were analysed and compared with other techniques. The ratio calculated was the maximum primary peak-trough amplitude divided by the maximum multiple peak-trough amplitude. Where the multiples and primaries were at the same time the multiple ratio was calculated after subtraction of the ideal primary (the stacked trace of primaries only). This analysis is shown in **Table** 6-2. From this table we can conclude that in terms of stacked signal improvement, the technique performs as well as the hybrid technique, and gives better results than the *f-k* and Hampson techniques.
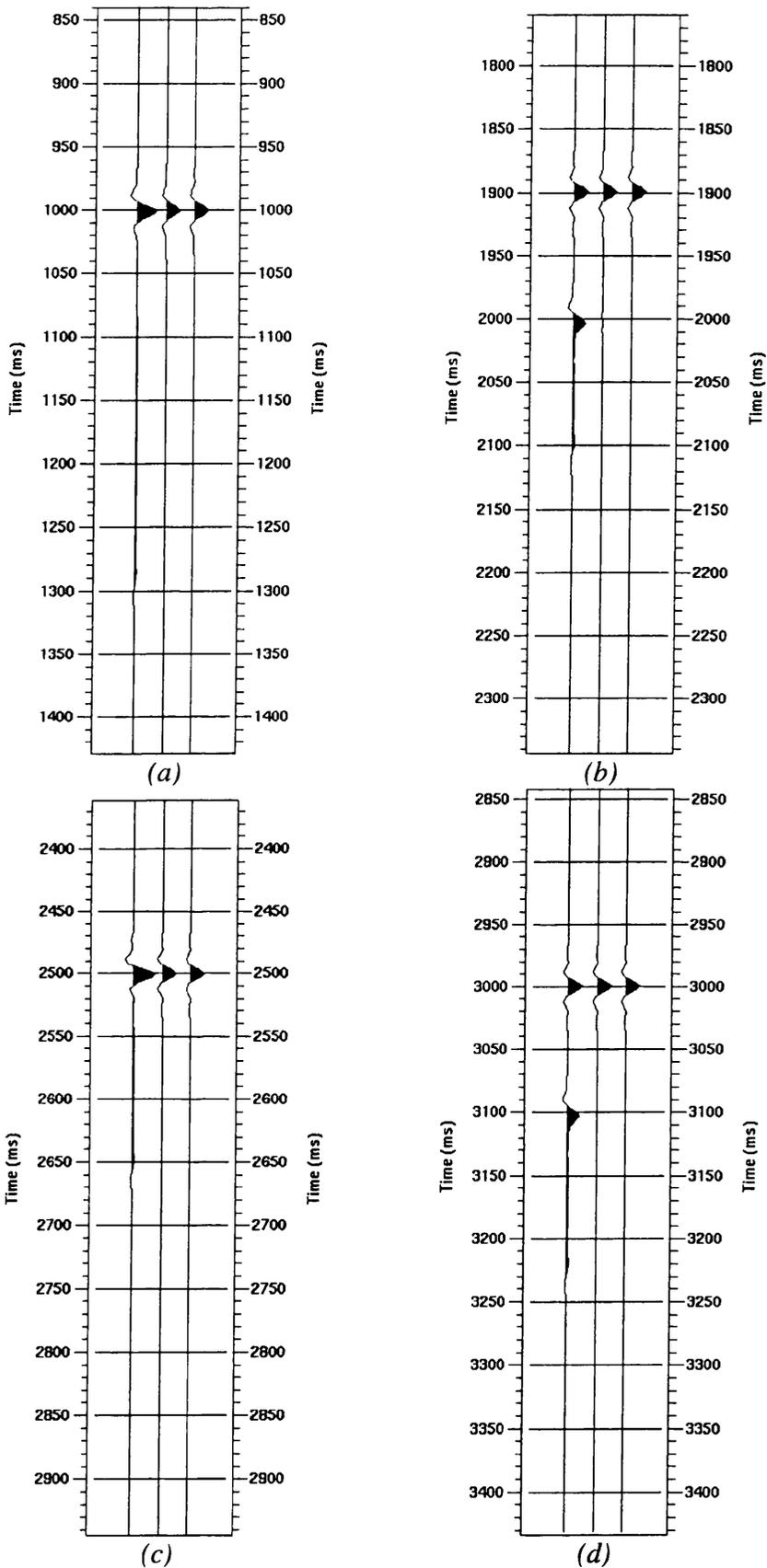
Figure 6-33 Stacked synthetics for model 1 showing the unfiltered stacked trace as trace 1, the filtered trace as trace 2 and the ideal output trace as trace 3, for (a) reflector 1 (b) reflector 2 (c) reflector 3 (d) reflector 4.

Model	f-k	Hampson	Hybrid	Wavelet
1	1.0	5.5	10.2	18.5, 18.3, 26.3, 47.2 (27.6)
2	4.3	18.0	40.0	19.9, 18.0, 25.8, 19.7 (20.8)
3	0.85	3.9	4.4	36.4, 9.9, 31.0, 10.5 (21.95)

Table 6-2 Table comparing the ratio of maximum primary peak-trough to maximum multiple peak-trough amplitudes for various multiple suppression techniques. The wavelet technique shows the improvement for each reflector and the mean in brackets. The values for the other techniques is the mean.

6.10 Multiple suppression: Examples

The wavelet transform multiple suppression technique was evaluated on a deep water marine seismic data set from the Porcupine Basin. **Figure 6-34a** shows a typical CMP gather from this seismic line. The data has been pre-processed with a band-pass filter to remove low frequency swell and cable noise. The gather is highly contaminated by multiples which can be seen from the gather and a semblance analysis of the gather shown in **Figure 6-35a**. The CMP gather was filtered using the wavelet transform technique the output of which is shown in **Figure 6-34b** with the corresponding difference section shown in **Figure 6-36**. A semblance analysis if the filtered gather is shown in **Figure 6-35b**. From these figures we can see that the multiples have been successfully suppressed and the primaries preserved. There is some residual multiple energy left at large offsets as was seen with the synthetic examples. When implementing the filter on this data set an AGC was applied before filtering and removed after filtering. Only very simple pre-processing had been applied to the data without the amplitude preserving steps normally associated with seismic data (Resnick, 1993) therefore the AGC was applied to minimise large trace to trace amplitude variations for coherent events.
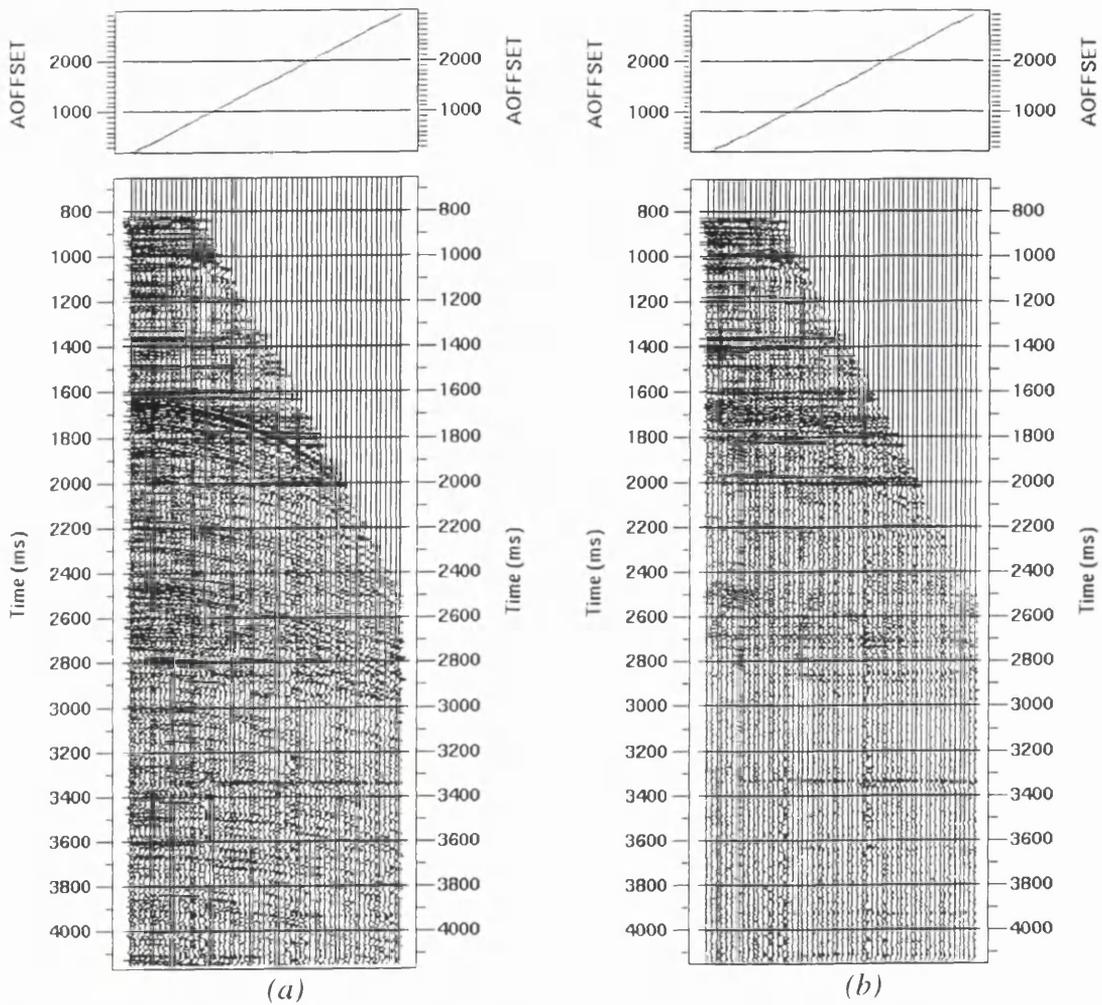
Figure 6-34 Marine CMP gather from Porcupine basin (a) before and (b) after multiple suppression using the wavelet transform technique.
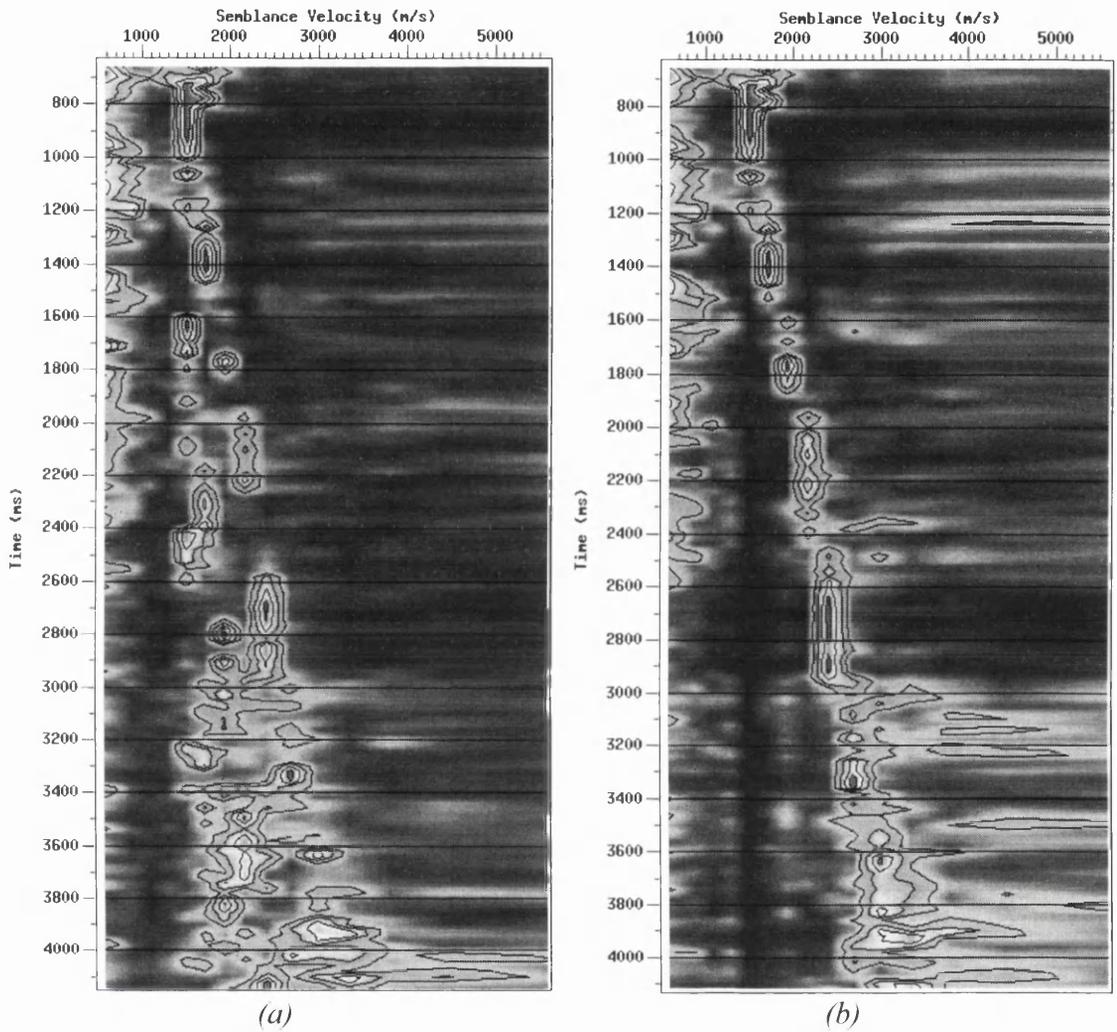
Figure 6-35 Semblance analyses corresponding to Figure 5.29(a) and (b) showing the suppression of the multiple velocity field.
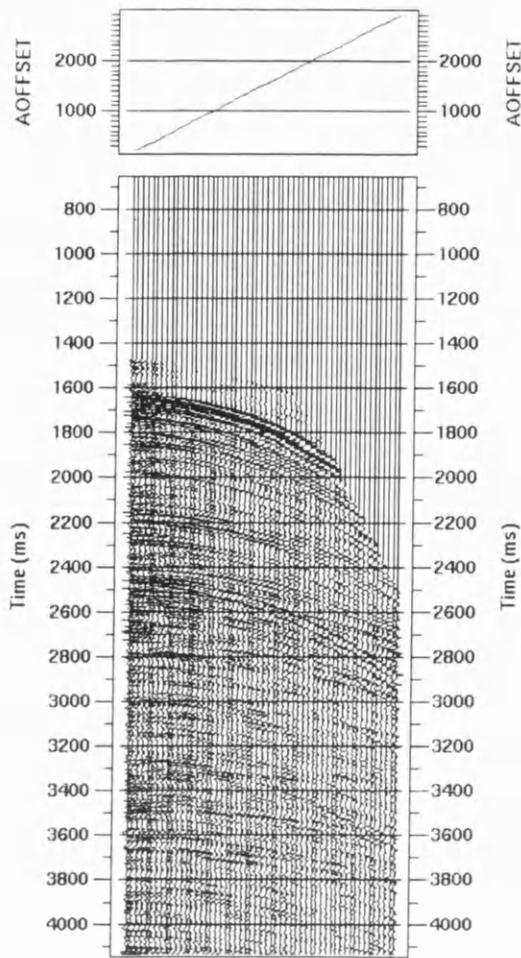
Figure 6-36 Difference section between Figures 5.28(a) and (b) showing the multiple field which has been removed.

Figure 6-37a and **Figure 6-37c** show a portion of the brute stack before and after multiple suppression. Apart from the wavelet transform filter, the stacked sections were processed with identical streams. **Figure 6-37b** shows the difference between the two stacked sections at the same scale. From these figures we can see that a considerable amount of multiple energy has been removed from the original stack. From the figure we can see that reflectors at 1650 ms and 2750 ms which previously contained multiple energy are now more clearly defined on the filtered section.

6.11 Conclusions

We have seen that the ability of the discrete wavelet transform to filter non-stationary data can be used as a powerful processing technique for the suppression of coherent linear and hyperbolic energy from seismic data. With the use of a simple weighting process the amplitude of primary reflected events can be preserved during linear filtering and minimised during hyperbolic filtering. Due to NMO stretch, the

performance of the technique for multiple suppression seems to be more effective for multiples with smaller residual moveout. This is also a problem with other multiple techniques. Further limitations include dead channels which must be interpolated and the presence of large trace to trace statics. As with other techniques, this filtering process is limited by knowledge of the velocity function. If the velocity function is known within acceptable errors, 2-10% (Schneider, 1971) the filtering process is an effective tool. The more accurately the velocity field is known, the better the performance of the filter.

Further study of the relationship of the wavelet transform when applied across gathers may reveal a more robust weighting technique that will preserve amplitudes to a greater degree in hyperbolic filtering and minimise the effects of NMO stretch. To further improve the wavelet transform technique and to produce the corresponding improvement on real data the wavelet transform technique must be modified to adapt to AVO changes in time as observed in the data. The choice of basis wavelet used in the transform is also crucial, as was discussed in Chapter 3.

Figure 6-37 Brute stack of the line containing the CMP gather shown in Figure 5.29, (a) before, and (c) after the multiple suppression technique. The difference section is shown in (b).

7. Conclusions

7.1 Discussion

The primary aim of this thesis was to evaluate the potential of the discrete wavelet transform as a potential tool for filtering seismic data. In chapter three we have shown that in one dimension, the multiresolution form of the discrete wavelet transform can be used effectively to suppress low frequency noise from seismic records in a time varying fashion. The simple process of muting coefficients in the wavelet domain works effectively to suppress ground roll from common shot records. However, we have demonstrated that in the wavelet domain, we do not always have perfect signal/noise separation, and so there is an associated loss of signal when we use this muting technique. By the development of a simple weighting procedure based on the Q-value of a seismic signal (which can be estimated from the signal and the corresponding wavelet transform coefficients), we have shown that it is possible to filter the signals in the wavelet domain, preserving more faithfully the bandwidth of the reflected signal.

For higher frequency noise contamination, we have shown in chapter four that the discrete wavelet transform does not perform well, due to the octave band splitting of the frequency domain inherent in the transform process. Replacing the discrete wavelet transform with the discrete wavelet packet transform allows a more flexible tiling of the time-frequency domain and so allows filtering of higher frequency noise whilst minimising any degradation of reflected signals. Again, we have shown that a simple muting technique in the wavelet packet domain effectively suppresses higher frequency noise such as airblast, while weighting of the coefficients allows noise suppression, preserving the bandwidth of reflected signals more faithfully. A drawback of this technique is that more transform iterations are required than for the wavelet transform, and that the transform produces a redundant representation in the transform domain.

In chapter five, we have shown that the filtering processes and results in one dimension can be extended to two dimensions where we filter a four variable transform space. These filters are analogous to time-offset varying *f-k* filters. We demonstrated that the discrete wavelet transform in two dimensions suffers from poor resolution in *f-k* space and to be effective for filtering, an additional shift is required to move the data in *f-k* space to areas that are covered with higher resolution by the transform. The wavelet

packet transform circumvents this problem, with the added cost of extra transform iterations. Using real data we showed that the simple muting of wavelet packet coefficients effectively suppressed low velocity noise from common shot records, leading to improvement in the shot records and subsequent stack.

In chapter six, by applying the one-dimensional discrete wavelet transform in the offset direction, we successfully suppressed linear and hyperbolic coherent noise from seismic records. Using prior knowledge of the corresponding velocity function for the linear and hyperbolic events in common shot records and common midpoint records, we used linear and hyperbolic moveout to flatten seismic events, which we then suppressed by using the wavelet transform to filter the data across traces. This was an example of using the non-stationarity properties of the wavelet transform to filter data. We demonstrated on field data that this technique could be used effectively to suppress noise from both field records and stack. By testing the technique on synthetic data we showed that for high velocity contrasts, primary reflector amplitudes were preserved, and by the design of a weighting technique based on wavelet transforms of shifted versions of the input trace, we could preserve primary reflector amplitudes where the velocity contrast was low. We have demonstrated that these techniques were shown to produce results comparable to current multiple suppression techniques.

Throughout the development of these processing techniques we have investigated how the choice of kernel wavelet in the transform process affects the performance of the filters. By analysis of these results and the frequency and temporal forms of the kernel wavelets we have developed criteria which helps assess the suitability of a kernel wavelet for filtering. The important properties are the flatness of the frequency spectrum across scales and the rate of fall-off in frequency amplitude at scale boundaries. These properties are of particular importance in wavelet packets where the frequency support of the wavelet packets is of more importance. Symmetry of the kernel wavelet is desirable if perfect signal/noise separation is unobtainable, but is not of such importance if the noise can be isolated in the wavelet/wavelet packet domain. In two dimensions we have the added requirement of temporal support which for certain kernel wavelets can lead to errors in the transform-inverse transform process. We have shown that the application of a gain function before filtering which is subsequently removed after filtering, minimises this effect.

Throughout this work we have investigated possible techniques of using the discrete wavelet and wavelet packet transform to filter seismic data. For these

techniques to be viable, they must provide filtering techniques and results that are at least comparable with current filtering techniques in terms of performance and processing speed.

7.2 Future work

The work in this thesis has developed throughout the course of the research. The techniques developed will, hopefully, continue to develop in the future and are in no way complete. The work has successfully ascertained whether the discrete wavelet transform is a viable filtering tool, and in some parts has developed the techniques further. It is impossible to fully develop these techniques within the scope of this thesis but it could be useful as a springboard for further development.

As an example, the weighting techniques developed in this work were used to determine whether weighting would be a viable alternative to muting the transform coefficients. The technique used, based on the Q-value of the seismic signal, in our opinion is not the best choice of technique as it is based on Fourier theory. A technique which would be far more satisfactory would be a weighting technique based on information extracted from the wavelet/wavelet packet coefficients themselves. This would maximise the information given by the transform. However, the Q based technique served a purpose in that it showed us that a weighting would be effective and therefore we can predict that a wavelet coefficient based technique will provide better results. Development of weighting techniques in all the filtering processes covered in this thesis will lead to improved results.

In using the wavelet packet transform for the filtering of seismic data, we have barely scratched the surface. In using a constant level of decomposition we have neglected the potential information that the data at other levels of resolution provides. We have demonstrated that it can be used as an effective tool for filtering, but further development of this technique must use a best basis selection procedure that maximises the information obtained from all levels of resolution. Development of a suitable best basis selection for filtering would probably involve an entire thesis of work, but would maximise the information provided by the transform and optimise any filtering results.

In ascertaining the criteria to assess the potential of a kernel wavelet for filtering, we looked at several wavelets. For the wavelet transform, and therefore the wavelet packet transform, there is an infinite number of possible kernel wavelets to choose from

that satisfy the necessary conditions. Obviously it is impossible to examine every possible wavelet, but approaching the problem from the opposite direction, developing a wavelet specifically for optimum filtering performance, is a realisable goal.

In this work we have concentrated on the suppression of noise from the pre-stack domain. This work can easily be extended to the post-stack domain, but, as more and more emphasis is being placed in pre-stack domain in terms of seismic attributes and derived reservoir information, the greatest potential is probably in the pre-stack domain. In addition, three-dimensional seismic is now a standard acquisition technique, the extension of these filters from one and two dimensional filters to three-dimensional filters could be advantageous.

References

Al-Husseini, M. I., Glover, J. B., and Barley, B. J., 1981, Dispersion patterns of the ground-roll in eastern Saudi Arabia: Geophysics, **43**, 121-137.

Alkemade, J. A. H., 1993, The finite wavelet transform with an application to seismic processing, *in* Koornwinder, T. H., Ed., Wavelets: An elementary treatment of theory and applications, 183-208.

Alvarez, G., and Larner, K., 1996, Implications of multiple suppression for AVO analysis and CMP stacked data: 66th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1518-1521.

Battle, G., 1987, A block spin construction of ondelettes. Part I: Lemarie functions: Comm. Math. Phys., **110**, 601-615.

Battle, G., 1988, A block spin construction of ondelettes. Part II: the QFT connection: Comm. Math. Phys., **114**, 93-102.

Beresford-Smith, G., and Rango, R., 1988, Dispersive noise removal in *t-x* space: Application to Arctic data: Geophysics, **53**, 346-358.

Beresford-Smith, G., and Rango, R., 1989, Suppression of ground roll by windowing in two domains: First Break, **7**, No. 2, 55-63.

Beylkin, G., 1992, On the representation of operators in bases of compactly supported wavelets: SIAM J. Num. Anal., **29**, No. 6, 1716-1740.

Bosman, C., and Reiter, E., 1993, Seismic data compression using wavelet transforms: 63rd Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1261-1264.

Bracewell, R., 1986, The fourier transform and its applications: McGraw-Hill Publ. Co.

Bunks, C., Saleck, F. M., Zaleski, S., and Chavent, G., 1995, Multiscale seismic waveform inversion: Geophysics, **60**, No. 5, 1457-1473.

Chakraborty, A., and Okaya, D., 1995, Frequency-time decomposition of seismic data using wavelet-based methods: Geophysics, **60**, 1906-1916.

Chui, C. K., 1992, An introduction to wavelets: Academic Press.

Cohen, A., Daubechies, I., and J.-C. Feauveau, 1992, Biorthogonal bases of compactly supported wavelets: Commun. On Pure and Appl. Math., **45**, 485-560.

Coifman, R. R., Meyer, Y., Quake, S., and Wickerhauser, M.V., 1989, Signal processing and compression with wave packets: Proceedings of the International

Conference on Wavelets, Marseille, Meyer, Y., Ed., Masson, Paris.

Coifman, R., Meyer, Y., and Wickerhauser, M. V., 1992, Wavelet Analysis and Signal Processing, *in* Ruskai, M. B., Beylkin, G., Coifman, R., Daubechies, I., Mallat, S., Meyer, Y., Eds., Wavelets and their Applications: Jones and Bartlett, 153-178.

Daubechies, I., 1988, Orthonormal bases of compactly supported wavelets: Commun. Pure Appl. Math., **41**, 909-996.

Daubechies, I., 1992, Ten Lectures on Wavelets: CBMS-NSF Series in Applied Mathematics, SIAM, **61**.

Daubechies, I., 1993, Orthonormal bases of compactly supported wavelets II. Variations on a theme: SIAM J. Math. Anal., **24**, 499-519.

Dessing, F. J., and Wapenaar, C. P. A., 1995, Efficient migration with one-way operators in the wavelet transform domain: 65th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1240-1243.

Dobrin, M. B., 1957, Dispersion in Seismic Surface Waves: Geophysics, **16**, 63-80.

Donoho, P. L., Ergas, R. A., Villasenor, J. D., 1995, High performance seismic trace compression: 65th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 160-163.

Dudgeon, D. C., 1977, Fundamentals of digital array processing: Proc. IEEE, **65**, 898-904.

Evison, F. F., 1956, Seismic waves from a transducer at the surface of stratified ground: Geophysics, **21**, No. 4, 939-959.

Faqi, L., Kabir, M. M. N., and Verschuur, D. J., 1995, Seismic processing using the wavelet and the radon transform: Journal of Seismic Exploration, **4**, 375-390.

Foster, D. J., and Mosher, C. C., 1994, Wavelet transform methods for geophysical applications: 64th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1465-1468.

Gabor, D., 1946, Theory of communication: Journal of I.E.E. (London), **93**, 429-441.

Goupillaud, P., Grossman, A., and Morlet, J., 1985, Cycle-octave and related transforms in seismic signal analysis: Geoexploration, **23**, 85-102.

Grossman, A., and Morlet, J., 1984, Decomposition of Hardy functions into square integrable wavelets of constant shape: SIAM. J. Math. Anal., **15**, No. 4, 723-736.

Grubb, H. J., and Walden, A. T., 1997, Characterizing seismic time series using the discrete wavelet transform: Geophysical Prospecting, **45**, No. 2, 183-206.

Haar, A., 1910, Zur Theorie der orthogonalen Funktionensysteme: Math. Annal., **69**, 331-371.

Hampson, D., 1986, Inverse velocity stacking for multiple elimination: J. Can. Soc. Expl. Geophys., **22**, 44-55.

Harlan, W. S., Claerbout, J. F., and Rocca, F., 1984, Signal/noise separation and velocity estimation: Geophysics, 49, 1869-1880, *in* Slant Stack Processing, Geophysics reprint series No. 14, Soc. Expl. Geophys., 1991.

Jawerth, B., and Sweldens, W., 1994, An overview of wavelet based multiresolution analysis: SIAM Review, **36**, no. 3, 377-412.

Kjartannson, E., 1979, Constant Q-wave propagation and attenuation: J. Geophys. Res., **84**, 4737-4748.

Knapp, R. W., 1986, Observations of the air-coupled wave as a function of depth: Geophysics, **51**, No. 9, 1853-1857.

Lemarié, P. G., 1988, Une nouvelle base d'ondelettes de $L^2(R)$: J. Math. Pure Appl., **67**, 227-236.

Li, X., and Haury, J. C., 1995, Characterization of heterogeneities from sonic measurements using the wavelet transform: 65th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 488-491.

Li, X.-G., and Ulrych, T. J., 1995, Tomography via wavelet transform constraints: 65th Ann. Internat. Mtg., Soc. Expl. Geophys., 1070-1073.

Li, X.-G., and Ulrych, T. J., 1996, Multiscale attribute analysis and trace decomposition: 66th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1634-1637.

Lichmann, E., and Northwood, E. J., 1997, High-resolution velocity notch filter without Gibbs effect: Geophysics, **62**, 274-290.

Luo, Y., and Schuster, G. T., 1992, Wavelet Packet Transform and Data Compression: Presented at the 62nd Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1187-1190.

Makarov, V. V., Mushin, I. A., Tarakyn, E. A., and Trolov, B.K., 1996, Frequency-time analysis - a tool for geologic interpretation: 58th Ann. Internat. Mtg., E.A.E.G., Expanded Abstracts.

Mallat, S., 1989, A theory for multiresolution signal decomposition: The wavelet representation: IEEE Trans. Patt. Anal. Machine Intell., **11**, No.7, 674-693.

Mallat, S., and Hwang, W. L., 1992, Singularity detection and processing with wavelets: IEEE Trans. Inform. Theory, **38**, 617-643.

Mallat, S., and Zhang, Z., 1993, Matching pursuit with time-frequency dictionaries: IEEE Trans. Sig. Proc., **41**, 3397-3415.

Mann, S., and Heykin, S., 1992, Adaptive 'chirplet' transform: an adaptive generalization of the wavelet transform: Optical Engineering, **31**, 1243-56.

Mansar, S., and Rodriguez, 1996, J.-M., Analyzing and correcting seismic phases using the wavelet transform: 58th Ann. Internat. Mtg., E.A.E.G., Expanded Abstracts.

Meyer, Y., 1989, Orthonormal Wavelets, *in* Combes, J. M., Grossman, A., Tchamitchian, P., Eds, Wavelets, Time Frequency Methods and Phase space: Springer Verlag.

Meyer, Y., 1990, Ondelettes et opérateurs I: Ondelettes: Hermann.

Mooney, H. M., and Kaasa, R. A., 1962, Air waves in engineering seismology: Geophysical Prospecting, **10**, 84-92.

Morlet, J., Arens, G., Fourgeau, E. and D. G., 1982, Wave propagation and sampling theory: Geophysics, **47**, No. 2, 203-236.

Nawab, S., and Quatieri, T., 1988, Short-time Fourier transform, *in* Lim, J., and Oppenheim, A., Eds., Advanced topics in signal processing: Prentice Hall Signal Processing Series, 289-337.

Newland, D. E., 1994, Wavelet analysis of vibration, part 2: Wavelet maps: J. Vib. Acoustics, **116**, 417-425.

Oppenheim, A. V., and Schafer, R. W., 1975, Digital signal processing: Prentice-Hall, Inc.

Press, W. H., Teulkowsky, S. A., Vetterling, W. T., and Flannery, B. P., 1993, Numerical recipes in C: Cambridge University Press.

Peacock, K., L., 1982, On the practical design of discrete velocity filters for seismic data processing: IEEE Trans. Acoust., Speech, Signal Processing, **30**, 52-60.

Resnick, J. R., 1993, Seismic data processing for AVO and AVA analysis, *in* , Castagna J. P., Backus, M. M., Eds., Offset-dependent reflectivity - Theory and practice of AVO analysis: Soc. Expl. Geophys.

Schneider, W. A., 1971, Developments in seismic data processing and analysis (1968-1970): Geophysics, **36**, 1043-1073.

Schuster, G. T., and Sun, Y., 1993, Wavelet Filtering of Tube and Surface waves: 63rd Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 25-28.

Strang, G., 1989, Wavelets and dilation equations: A brief introduction: SIAM Rev., **31**, 614-627.

Taner, 1980, Long-period sea-floor multiples and their suppression: Geophysical Prospecting, **28**, 30-48.

Taswell, C. and McGill, K. C., 1994, Algorithm 735: wavelet transform algorithms for finite- duration discrete-time signals: ACM Trans. Math. Soft., **20**, No. 3, 398-412.

Verschuur, D. J., and Berkhout, A. J., 1996, Removal of interbed multiples: 58th Ann. Internat. Mtg., Eur. Ass. of Geosc. and Eng., Expanded Abstracts.

Vetterli, M., 1992, Wavelets and filter banks: Theory and design: IEEE Trans. Sig. Proc., **40**, No. 9, 2207-2232.

Vetterli, M., and Kovacevic, J., 1995, Wavelets and subband coding: Prentice Hall.

Wang, B., and Pann, K., 1996, Kirchoff migration of seismic data compressed by matching pursuit decomposition: 66th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1642-1645.

Williams, J. R., and Armaratunga, K., 1994, Introduction to wavelets in engineering: Int. J. Numer. Methods Eng., **37**, 2365-2388.

Wu, Y., and McMechan, G., 1995, Wavefield extrapolation in the wavelet domain with application to poststack migration: 65th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1236-1239.

Yilmaz, O., 1987, Seismic data processing: Soc. Expl. Geophys.