ADAPTIVE SEISMIC COMPRESSION BY WAVELET SHRINKAGE

M.F. Khène, S.H. Abdul-Jauwad

King Fahd University of Petroleum & Minerals
Electrical Engineering Department
Dhahran 31261, Saudi Arabia
{samara, mfkhene@kfupm.edu.sa}

ABSTRACT

In this paper, a sophisticated adaptive seismic compression method is presented based on *wavelet shrinkage*. Our approach combines a time-scale transform with an adaptive non-linear statistical method. First, a discrete 2-D biorthogonal Discrete Wavelet Transform (DWT) is applied to the multi-channel seismic signals to generate a sparse multiresolution (subband) decomposition. Compression is then achieved by shrinking the detail wavelet coefficients using a scale-dependent non-linear soft-thresholding rule. The adaptive scale-dependent thresholds are determined by minimizing the *Stein's Unbiased Risk Estimate (SURE)*. The proposed compression procedure is tested on marine seismic data from the *Midyan* basin (*Red Sea, Saudi Arabia*).

1. INTRODUCTION

Seismic compression is a key technology for managing seismic data in a world of ever increasing data volumes to maintain productivity without compromising interpretation results. By storing data in a format that requires less space than the original data volume, seismic compression provides greater flexibility in managing local or remote server disk space as well as reducing network traffic. Seismic compression not only enables explorationists to maximize the value of the information technology infrastructure, but it encourages innovative interpretation workflow to leverage the vast information content in massive seismic data. Compression thus helps in maintaining or exceeding current productivity levels. Recently, seismic compression has benefited from the advent of wavelets [1], which offer mathematical constructions with a great potential in statistical methodology [2]. Wavelet transforms have been applied extensively in diverse applications including data compression and denoising, image analysis, economics and statistics [3].

In this paper, a sophisticated wavelet-based compression technique is proposed. Both the transform and the compression stages are matched up in view to improving the overall performance. The rationale of our approach is first to generate a (near)-sparse representation of the data in the wavelet domain then to threshold an important part of the coefficients without losing substantial information. In order to exploit the multi-channel seismic data correlation in space and time directions, a 2-D biorthogonal DWT is used. For seismic interpretation, the visual aspect of seismic signals is of utmost importance. This factor is taken into account by judiciously selecting both the wavelets and the thresholding procedure. Thus, a DWT using long wavelet filters from the *Cohen-Daubechies-Feauveau* (CDF) class [4] is intimately associated with a non-linear smooth operator, namely wavelet *shrinkage*. Biorthogonal wavelets offer a good trade-off between the support size, the number of vanishing moments and regularity. In other words, the DWT is computed efficiently while preventing the appearance of artifacts in the reconstructed data. In addition, the sparsity of the multiresolution decomposition is best exploited by wavelet shrinkage. The latter consists of applying a soft-thresholding rule to all the wavelet coefficients but those belonging to the lowest resolution subband. Indeed, the latter is merely a smooth scaled version of the input data and carries the essence of the data. Moreover, it has coefficients of much smaller magnitude than those of the detail subbands do. Thus, this makes its contribution to the compression gain marginal. The values of the scale-dependent thresholds are determined by minimizing the SURE [5][6]. The proposed compression procedure is tested on marine seismic data from the *Midyan* basin *(Red Sea, Saudi Arabia)* [7].

2. MULTICHANNEL SEISMIC SIGNALS

Oil and gas are usually buried deep within the earth, often miles below the surface. Most of the easy or shallow oil has been found. A number of exploration methods are available but in general only the modern seismic reflection method comes close to providing both the ability to see down to great depths and to see the details of the subsurface needed to locate many hydrocarbons [8]. Seismic data stem from a multiscale non-linear distributed parameters remote system, i.e. the earth. In a typical scenario, a spatially distributed acoustical signal is generated by a source (e.g. dynamite) located at the surface of the earth. The generated waves propagate downward, undergo reflection at contacts with different acoustic impedance, and are recorded by an array of seismometers at the surface. This provides the multi-channels discrete seismic signals that are mapped into representations of the earth's interior properties. The underlying complex process is referred to as seismic imaging. The latter is intended to find earth models that explain (or best fit) seismic observations. Seismic signals are commonly displayed using a variable-area and/or a variable-density mode [8].

3. MULTIRESOLUTION DECOMPOSITION

3.1 2-D Wavelet Bases

There are two different ways to build a wavelet basis for a 2-D space, say t and x. The standard dyadic construction consists of

0-7803-5988-7/00/$10.00 © 2000 IEEE

all possible tensor products of 1-D wavelet and scaling basis functions defined respectively as:

$$\psi_{jk}(t) = 2^{\frac{j}{2}}\psi_{jk}(2^j t - k) \quad (1a)$$

$$\varphi_{jk}(t) = 2^{\frac{j}{2}}\varphi_{jk}(2^j t - k) \quad (1b)$$

However, despite its simplicity, the construction that requires different scale indices for each direction, does not benefit from the recursive *Mallat* algorithm [9]. Indeed, for an $m \times m$ matrix data the standard dyadic decomposition requires $4(m^2-m)$ assignment operations against $8/3(m^2-1)$ only for the nonstandard one [10]. Consequently, in the sequel the nonstandard dyadic 2-D decomposition is adopted. It consists of defining a 2-D scaling function, using a unique scale index j as:

$$\varphi_{jkk'}(t,x) = \varphi_{jk}(t)\varphi_{jk'}(x) \quad (2)$$

and three 2-D wavelet functions at each scale given by:

$$\psi^V_{jkk'}(t,x) = 2^{\frac{j}{2}}\varphi\psi(2^j t - k, 2^j x - k') \quad (3a)$$

$$\psi^H_{jkk'}(t,x) = 2^{\frac{j}{2}}\psi\varphi(2^j t - k, 2^j x - k') \quad (3b)$$

$$\psi^D_{jkk'}(t,x) = 2^{\frac{j}{2}}\psi\psi(2^j t - k, 2^j x - k') \quad (3c)$$

These three *anisotropic* wavelets extract matrix data details at different scales and orientations, whereas the scaling function yields a smoothed low-resolution version of the input data. Indeed, starting at scale j, the multiresolution decomposition yields four double-scaled half-resolution panels at scale $j-1$. One of them represents a smoothed version of the data while the remaining ones contain detail wavelet coefficients corresponding to the $\{\psi^V, \psi^H, \psi^D\}$ wavelet functions that are respectively oriented vertically, horizontally and diagonally. The result of the nonstandard dyadic 2-D DWT is usually displayed in four panels as in Fig.1.

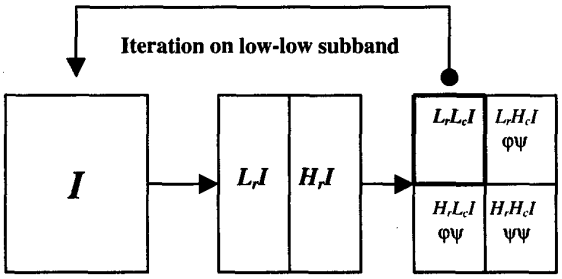

Figure 1. Nonstandard dyadic 2-D wavelet multiresolution decomposition. L and H stand for low- and high-pass wavelet filters, and the subscripts r and c stand for row and column respectively

3.2 Biorthogonal Wavelet Bases

There are three main categories of wavelet bases, namely the orthogonal, the semi-orthogonal and the biorthogonal. Limiting ourselves to orthogonal wavelet bases can be overly restrictive because except for the *Haar* basis, there are no other bases, which are compactly supported and symmetric. The relaxation of orthogonality constraint has many benefits that improve the performance of the wavelet transform while still being implemented with the *Mallat* algorithm. In particular, biorthogonal wavelets offer a good trade-off between the support size, the number of vanishing moments and regularity. In term of digital filters, the biorthogonal transform uses different Finite Impulse Response (FIR) wavelet filters in the decomposition and reconstruction stages. This provides more flexibility in the design of the transform and its inverse [11]. Moreover, FIR filters are preferred because they guarantee a linear phase, which is a very desirable property that prevents from the appearance of artifacts in the reconstructed data. Therefore, the biorthogonal transform uses dual wavelet and dual scaling functions related to the primal ones by:

$$\left.\begin{aligned}\langle \varphi_{jk} | \tilde{\psi}_{jk'} \rangle = 0 \\ \langle \psi_{jk'} | \tilde{\varphi}_{jk'} \rangle = 0\end{aligned}\right\} \text{ for all } j,k,k' \quad (4)$$

In this contribution long biorthogonal wavelets filters of the *CDF* class are used [4].

3.3 Nonstandard Dyadic 2-D Decomposition

The extension to 2-D separable biorthogonal bases is straightforward. Indeed, by alternating the 1-D wavelet filtering operations on rows and columns a 2-D dyadic nonstandard decomposition is generated. This scheme is implemented with a two-channel filter bank [11], where the low-pass and high-pass filters represent the scaling and the wavelet functions respectively. First, a one step low pass (L) and high pass (H) filtering is performed on each row of the matrix I. Next, the same filters are applied to each column of the resulting matrix. The whole process is applied recursively to the quadrant containing averages in both directions, i.e. L_rL_cI panel. The resulting recursive decomposition is illustrated by Fig.1. and the implementation of the one-stage 2-D DWT is depicted by Fig.2.

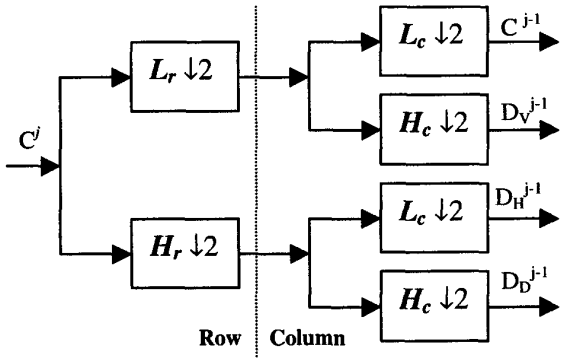

Figure 2. One-stage non-standard 2-D DWT

4. WAVELET SHRINKAGE

4.1 Motivation

Wavelet compression is best understood from an approximation viewpoint. In a wavelet decomposition, each wavelet picks up information about the data at a given location k and at a given resolution or scale j. Thus, the wavelet transform allows us to focus on the most relevant part of the data provided the wavelet bases fit the input data. Consequently, the resulting wavelet coefficients drop off rapidly yielding a (near)-sparse data representation. This is also known as energy compaction property. Wavelet thresholding constitutes thus a natural choice to perform compression. It is a simple yet a very efficient procedure for keeping the most important coefficients that will be used in reconstruction. An intuitive way to achieving thresholding consists of applying a *keep-or-kill* rule referred to as *hard-thresholding*. However, a *soft-thresholding* is preferred because of various advantages. From a visual point of view, the reconstructed data offer a more pleasant aspect, and do not exhibit visible artifacts. This is crucial in the case of seismic data interpretation. From a statistical point of view, soft-thresolding uses a continuous function, leading to simple data driven selection of the thresholds. In fact, the selection of the thresholds is a very delicate and important statistical problem. On one hand, killing too many wavelet coefficients may lead to an important bias in the reconstructed data. On the other hand, small thresholds lead to a poor compression gain. Thus, threshold selection should strike the balance between closeness to fit between the original and the reconstructed data and the degree of sparsity of the wavelet coefficients. We propose to achieve compression through an adaptive soft-thresholding procedure, referred to as wavelet shrinkage. The selection of the optimal threshold for all the scales but the coarsest one is accomplished by minimizing the SURE. The resulting nonlinear thresholding operator is called *SureShrink* [5].

4.2 SURE Principle

SURE has been initiated by *Stein* for mean estimation of a multivariate normal distribution [12] and has been successfully applied for function smoothing by *Donoho* [5]. The foundation of the SURE principle is based on the fact that for nearly arbitrary nonlinear biased estimator, the loss or risk can be estimated unbiasedly. In the sequel, we outline the SURE principle for the general case then in the next paragraph we show how to derive the SureShrink operator.

Consider an empirical data vector $\mathbf{y}$ of dimension N given by

$$y_i = f_i + e_i, \qquad i = 0,1,...,N-1 \tag{5}$$

where f_i are samples of the deterministic function f and e is Gaussian white noise with independent identical distribution *(i.i.d)* $N(0,\sigma)$.

The objective is to find the best estimate of the function f in the mean square sense by minimizing the *Mean Square Error* (MSE) risk defined as,

$$R(\hat{f},f) = \frac{1}{N}\| \hat{\underset{\sim}{f}} - \underset{\sim}{f} \|^2 = \frac{1}{N}\sum_{i=0}^{N-1}(\hat{f}_i - f_i)^2 \tag{6}$$

However, the main drawback of the MSE risk is that in practice, it can never be computed exactly because it relies on the unknown exact value of the function f. Thus in practical situations this MSE has to be estimated. The SURE principle stipulates that if we consider the following estimate for the unknown function f,

$$\hat{f}(y) = y + g(y) \tag{7}$$

where $g(y)$ is a weakly differentiable function from R^N to R^N, then an unbiased estimator for the MSE risk is the SURE defined as [12]:

$$R^{SURE}(\hat{f}(y),f) = N + \left\|\underset{\sim}{g}(y)\right\|^2 + 2\nabla_y.\underset{\sim}{g}(y) \tag{8}$$

where ∇ is the vector differential operator of first partial derivatives , i.e.,

$$\nabla_y.\underset{\sim}{g}(y) \equiv \sum_{i=0}^{N-1}\frac{\partial}{\partial y_i}g_i \tag{9}$$

4.3 Adaptive Wavelet Shrinkage with SureShrink

There are two main classes of wavelet shrinkage regarding whether the threshold is single and global or scale-dependent and adaptive. Our approach consists of deriving a scale-dependent threshold λ_j according to the following soft-thresholding rule:

$$g_\lambda(d) = \mathrm{sgn}(d_k^j)(|d_k^j| - \lambda^j)I(|d_k^j| > \lambda^j) \tag{10}$$

For a given detail subband at resolution j, the shrinkage operator $g_\lambda(d)$ kills all those coefficients below the threshold λ^j and pulls towards the origin the surviving ones by an amount equals to the threshold. The different scale-dependent thresholds stem from the minimization of the SURE, i.e.,

$$\lambda^j = \arg\min_{\lambda \geq 0} R_\lambda^{SURE}(\lambda^j, d_k^j) \tag{11}$$

where the SURE for a soft-threshold estimator is given by [5],

$$R_\lambda^{SURE} = 2^j - 2I\{|d_k^j| \leq \lambda^j\} + \sum_{i=0}^{2^j-1}\min\{|d_k^j|,\lambda^j\}^2 \tag{12}$$

Note that the underlying optimization problem is straightforward and the computational effort is of order $2^j\, log(2^j)$ as a function of the subband size 2^j [5].

5. EXPERIMENTAL RESULTS

A migrated marine seismic profile from the *Midyan* basin in the *Red Sea* is used to demonstrate our adaptive seismic compression by wavelet shrinkage. The discrete 2-D seismic data consist of a collection of 2838 seismic signals (traces) of 2.5 seconds length each sampled at 4 milliseconds. The traces correspond to the *Common MidPoints* (CMP), i.e., successive reflection points midway between the different seismic source locations and the seismometers. The data can be regarded as a (626x2838) matrix of floats entries. The variable density display mode is used to represent the profile, which can be thought of as a transversal section of the prospected area along the seismic line. First we have applied a 2-D DWT with asymmetric long biorthogonal wavelet filters *CDF(6,8)* where the numbers of vanishing moments for the synthesis and analysis wavelets are 6 and 8 respectively. The reader may be wondering why the reconstruction wavelet filters are shorter and have less vanishing moments than the decomposition ones. First note that this is made possible thanks to the flexibility of the biorthogonal wavelet transform. Second, the objectives of the decomposition and reconstruction stages differ. Indeed, the main concern of the wavelet decomposition is to pack the energy of the input data in fewer wavelet coefficients. The higher the number of vanishing moments is, the better the energy compaction would be. From the reconstruction side, we wish to use smooth wavelets to mask the errors introduced by the wavelet shrinkage and to get less annoying visual artifacts. Furthermore, the reconstruction time should be shorter than the decomposition one because, in practice, the data set is compressed once but may be decompressed several times. A three-level multiresolution decomposition has been performed yielding nine detail subbands and one low-resolution subband. The four subbands of the first level are displayed in Fig.3.

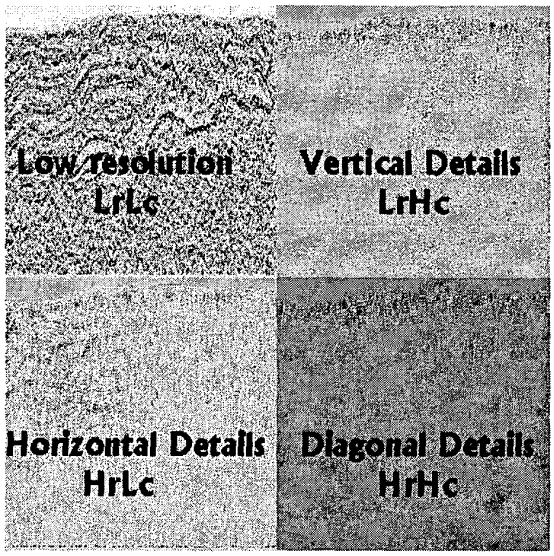

Figure 3. First level nonstandard dyadic 2-D DWT multiresolution decomposition of the Midyan section

Next, we have applied the SureShrink operator to the nine details subbands. The experimental results are displayed in Fig.4. Though, almost 82% of the wavelet coefficients have been killed by shrinkage, 95% of the data energy is recovered in the reconstructed data. Furthermore, the difference section exhibits random noise. Thus, an appreciable filtering effect has also been produced by the compression.

6. CONCLUSIONS

In this work, a sophisticated adaptive seismic compression technique was presented. A time-scale transform was associated with a non-linear statistical method. A pair of different asymmetric biorthogonal wavelet filters was selected to achieve different targets for the compression and decompression processes. Analysis wavelets with more vanishing moments were used to ensure a maximum energy compaction of the input data. This made compression by thresholding a very natural and efficient means. Based on SURE, the scale-dependent thresholds were determined and then the SureShrink operator was applied to each detail subband to kill insignificant coefficients. The experimental results show that the proposed approach does not introduces visible artifacts while achieving a relatively high compression gain.

7. ACKNOWLEDGMENTS

The authors wish to gratefully acknowledge *King Fahd University of Petroleum & Minerals* for its support and *Saudi ARAMCO* for providing the data.

8. REFERENCES

[1] Bosman C. and Reiter, E. "Seismic data compression Using wavelet transform." *63rd SEG Expanded Abstracts,1261-1264, 1993*

[2] Vidakovic B. *Statistical modeling by wavelets.* Wiley Series in Probability and Statistics. John Wiley & Sons, 1999.

[3] Meyer Y. *Wavelets: Algorithms and applications.* SIAM 1993.

[4] Cohen A., Daubechies I., Feauveau J. "Biorthogonal bases of compactly supported wavelets." *Commun. on Pure Appl. Math., 45:485-560, 1992.*

[5] Donoho D. and Johnstone I. "Adapting to unknown smoothness by wavelet shrinkage." *Journ. of the Amer. Statistical assoc., 90:1200-1224, 1995.*

[6] Misitti M., Missiti Y., Oppenheim. G and Poggi J-M. *Matlab wavelet toolbox.* The MathWorks, Inc, 1997.

[7] Mougenot D. and Al-Shakhis A. "Depth imaging a pre-salt faulted block: A case study from the Midyan basin (Red Sea)." *Saudi Aramco Jour. of Tech. Fall 1998.*

[8] Sheriff R. *Geophysical methods.* Prentice Hall, 1989.

[9] Mallat S. *A Wavelet tour of signal processing.* Academic Press, 1998.

[10] Stollinz E., DeRose T. and Salesin D. *Wavelets for computer graphics: Theory and applications.* Morgan Kaufmann Publishers, Inc., San Francisco, 1996.

[11] Vetterli M. and Kovacevic J. *wavelets and subband coding.* Prentice hall PTR, New Jersey, 1995.

[12] Stein C. "Estimation of the mean of a multivariate normal distribution" *The Annals of Statistics, 9(6): 1135-1151,1981.*

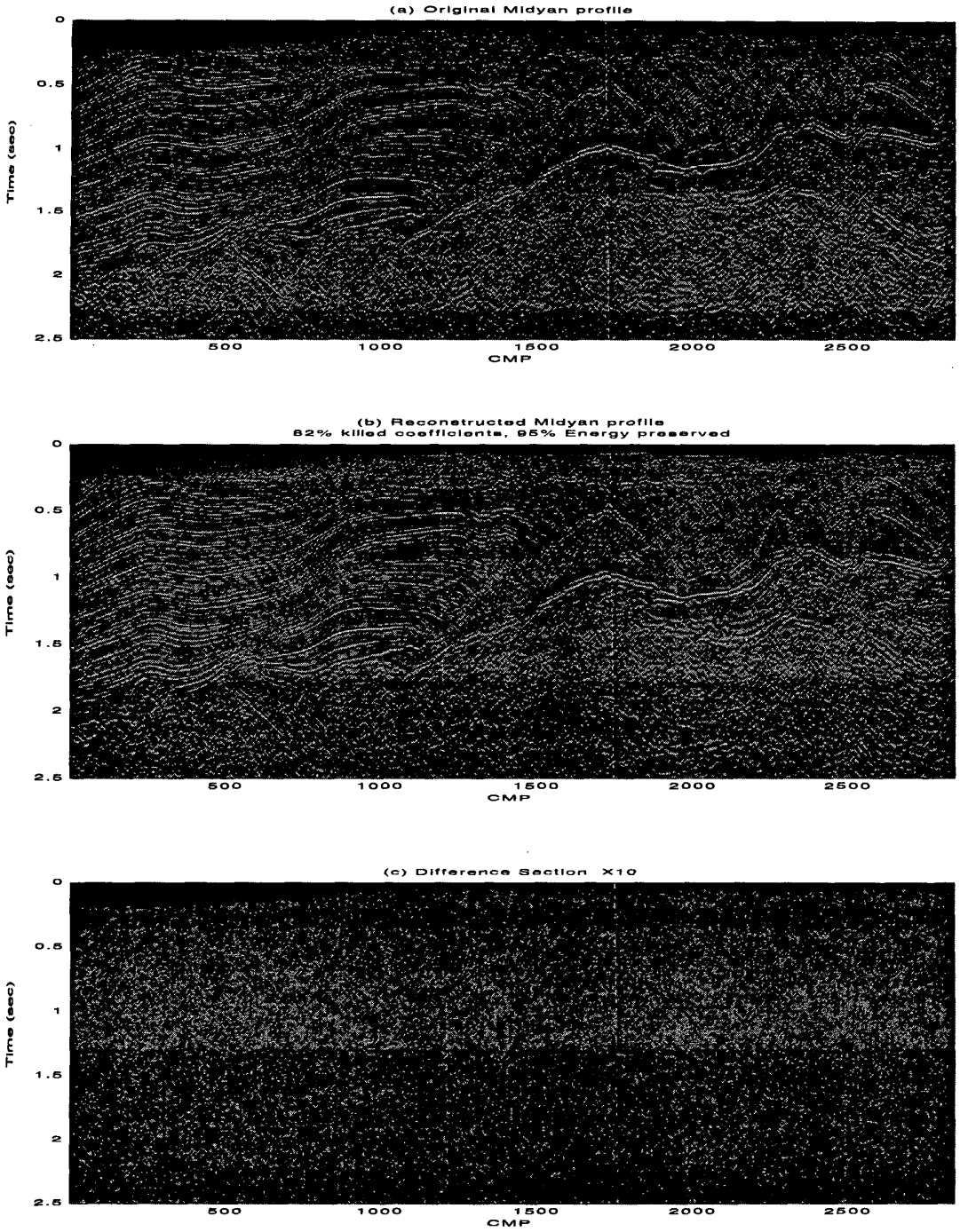

Figure 4. Experimental results