Wavelet transforms and compression of seismic data

G. Beylkin [1] and A. Vassiliou [2]

I   Introduction

Seismic data compression (as it exists today) is a version of transform coding which involves three main steps:

1. The transform step, which is accomplished by a fast wavelet, wavelet-packet or local cosine transform;

2. The quantization step, which is typically accomplished by a scalar uniform or non-uniform quantization scheme, and

3. The encoding step, which is accomplished by entropy coding such as Huffman coding or adaptive arithmetic coding.

Let us briefly outline the role of each step. The role of the transform step is to decorrelate the data. Namely, the transform will take a data set with a more or less flat histogram and produce a data set which has just a few large values and a very high number of near zero or zero values. In short, this step prepares the data for quantization. It has been observed that a much better compression is achieved by quantizing the decorrelated data than the original data.

There are several transforms that can be used for decorrelation. For example, the Karhunen-Loeve transform achieves decorrelation but at a very high computational cost. It turns out that the wavelet, wavelet-packet or local cosine transforms can be used instead. These transforms are fast and provide a local time-scale (or time-frequency) data representation, resulting in a relatively few large coefficients and a large number of small coefficients.

At the second step the coefficients of transformed data are quantized, i.e., mapped to a discrete data set. The quantization can take two forms, either scalar quantization, or vector quantization. In the case of scalar quantization every transform coefficent is quantized separately whereas in the case of vector quantization a block of coefficients is quantized simultaneously. Based on practical experience with seismic data it appears that the transformed coefficents tend to be reasonably decorrelated, thus pointing to scalar quantization as a good option.

The most straightforward type of scalar quantization is uniform quantization, where the range of the coefficient values is divided into intervals of equal lehgth (bins), except possibly for a separate bin near zero (zero bin). Since the quantization step is the only lossy step of the process, a certain level of optimization has to take place in order to accommodate the target compression ratio (or the target bit budget).

---

[1]Department of Applied Mathematics, University of Colorado at Boulder, Boulder, CO 80309-0526 and Fast Mathematical Algorithms and Hardware, 1020 Sherman Avenue, Hamden, CT 06514
[2]GeoEnergy Corporation, 2530 East 71 Street, Suite E, Tulsa, OK 74136

A distortion criterion (implied by the size of the seismic gather or ensemble of gathers and the target compression ratio) is minimized subject to the bit budget. In some cases a non-uniform quantization can yield lower distortion level than the uniform quantization.

We note that some new quantization/coding schemes which have been used for image compression may not be directly applicable to seismic data. For example, embedded zero-wavelet tree compression (EZW) scheme [13] does not appear efficient since seismic data violate the basic assumptions of EZW algorithm.

After quantization we are likely to have a number of repeated quantized coefficients and, thus, a significant redundancy. The third step, entropy coding, addresses this issue. Perhaps the easiest analogy to entropy coding comes from the Morse code communication, in which frequently encountered symbols are transmitted with shorter codes, while rarely encountered symbols are transmitted with longer codes. The entropy coding creates a new data set which has the average number of bits/sample minimized. There are two distinct cases of entropy coding. In the case of stationary data one can use Huffman coding. In the case of non-stationary data adaptive arithmetic coding is usually applied.

Now that we have an overall picture, we will desribe individual steps in greater detail. The notions that are considered below are not yet a familiar territory for a geophysicist and, for that reason, our goal will be limited to providing a basic trail map. We will consider the basic steps in the reverse order so that it is clear (at least intuitively) what is desirable to have as an output of the preceding step.

II   Entropy coding

Let us consider a finite set of symbols $S = (x_1, x_2, \ldots, x_N)$. Let p_n denote the propability of occurence of the symbol x_n in some set X (elements of which are from S), $\sum_{n=1}^{N} p_n = 1$. Let us associate a binary word w_n of length l_n with each symbol x_n. By aggregating binary words, we produce a code for X. The average bit rate in order to code X is

$$R_X = \sum_{n=1}^{N} l_n p_n.$$

It makes sense to select short code words w_n for symbols with high propability so to minimize R_X.

It is important to note that we should be able to identify the binary words uniquely from their "aggregated state". This can be achieved by using the so-called prefix codes where short binary words do not occur as the beginning of longer binary words. Such codes are constructed using binary trees where words correspond to the leaves of the tree. The binary digits of the code words are obtained by using 0 and 1 to correspond to left and right branches of the tree along the path to the node. Thus, the number of binary digits corresponds to the depth of the tree for a given node.

Shannon's theorem provides the lower bound for the average bit rate R_X, namely, the entropy

$$H(X) = -\sum_{n=1}^{N} p_n \log_2 p_n \leq R_X.$$

The Huffman algorithm (Huffman coding) [7] constructs an optimal prefix tree and the corresponding code so that

$$H(X) \leq R_X \leq H(X) + 1.$$

The difficulty in obtaining the lower bound via Huffman coding is that it requires $l_n = -\log_2 p_n$ and $\log_2 p_n$ generally is not an integer.

In order to avoid this constrain the algorithm of arithmetic coding [16] codes a sequence of symbols. The average bit rate of arithmetic coding approaches asymptotically the lower bound and in practice performes somewhat better than Huffman coding.

Finally, we note that probabilities of symbols are typically computed from N most recent symbols. This quick overview should make it clear that the role of Steps 1 and 2 is to provide to entropy coder a set with a minimal number of symbols where very few of the symbols carry "most of the probability".

III Quantization

An example of quantization is analog-to-digital converter with a fixed number of bits (say 24 bits). Clearly quantization is a lossy procedure by its nature. In this sense the digital data (e.g., seismogram) is already quantized. We are seeking a more efficient quantization.

The result of quantization is a set X of symbols (e.g., real numbers) that serves as an input to entropy coding. In general, quantization proceeds by taking the interval of variation of the signal and decomposing it into subintervals (called quantization bins). The center of the quantization bin can serve as a symbol (the standard round-off is an example here).

The uniform quantizer is a special case where all bins are of the same length. A non-uniform quantizer uses bins of different size. Another way to visualize quantization is to consider a grid on an interval (e.g., centers of the bins) and think of assigning the closest grid point to the real numbers that are being quantized.

Further distinction is drawn between scalar and vector quantization. Real numbers can be quantized in groups (vectors) and then one should think of a multidimensional grid corresponding to the multidimensional bins.

If one considers quantization and entropy coding together, it is clear that we should try to represent the signal with a minimal number of components, and control the dynamic range and significance of these components. Preparing the data for quantization and entropy coding falls on Step 1. The choice of the transformation is critical for the effective overall compression.

IV The transform step

The role of the transform step is to separate the more significant "coordinates" (coefficients) from the less significant. Initially one can wonder if such goal can be accoplished without the detailed knowledge of the signal. Indeed, if the usual representation of the signals by their samples were optimal, then very little could have been done without additional information. However, sampled representations are not optimal in any sense: they can be locally oversampled and be overly sensitive to the effects of quantization.

Traditionally, the Karhunen-Loeve transform was thought as a method of decorrelating the data. In fact there is a belief that this transform provides the optimal decorrelation. Although there is a theorem with a statement addressing this point (see e.g. [9], Proposition 11.3), it is important to remember that the Karhunen-Loeve transform is optimal for the random Gaussian process. Since this assumption may not be satisfied in practice, one has to be careful in stating optimality of this transformation.

In fact it was shown [1] (under an assumption of "piecewise stationary process" for the undelying signals) that the wavelet transform performs better that the Karhunen-Loeve transform.

Another problem with the Karhunen-Loeve is its cost. It requires diagonalization of covariance matrix and is too expennsive for using on large data sets. This is where the wavelet transform and its cousins play a very important role.

V Wavelet Transform

The Haar basis (Haar 1910) [6] provides a simple illustration of wavelet bases. The Haar basis in $\mathbf{L}^2(\mathbf{R})$, $\psi_{j,k}(x) = 2^{-j/2}\psi(2^{-j}x - k)$, where $j, k \in \mathbf{Z}$, is formed by the dilation and translation of a single function

$$\psi(x) = \begin{cases} 1 & \text{for } 0 < x < 1/2 \\ -1 & \text{for } 1/2 \le x < 1 \\ 0 & \text{elsewhere.} \end{cases} \tag{5.1}$$

The Haar function ψ satisfies the two-scale difference equation,

$$\psi(x) = \varphi(2x) - \varphi(2x - 1), \tag{5.2}$$

where $\varphi(x)$ is the characteristic function of the interval $(0, 1)$. The characteristic function $\varphi(x)$ is a prototype of the scaling function for wavelets and satisfies the two-scale difference equation,

$$\varphi(x) = \varphi(2x) + \varphi(2x - 1). \tag{5.3}$$

Given $N = 2^n$ "samples" of a function, which may for simplicity be thought of as values of scaled averages of f on intervals of length 2^{-n},

$$s_k^0 = 2^{n/2} \int_{2^{-n}k}^{2^{-n}(k+1)} f(x)dx, \tag{5.4}$$

we obtain the Haar coefficients as

$$d_k^{j+1} = \frac{1}{\sqrt{2}}(s_{2k-1}^j - s_{2k}^j) \tag{5.5}$$

and averages as

$$s_k^{j+1} = \frac{1}{\sqrt{2}}(s_{2k-1}^j + s_{2k}^j) \tag{5.6}$$

for $j = 0, \ldots, n-1$ and $k = 0, \ldots, 2^{n-j-1} - 1$. It is easy to see that evaluating the whole set of coefficients d_k^j, s_k^j in (5.5), (5.6) requires $2(N-1)$ additions and $2N$ multiplications.

In two dimensions, there are two natural ways to construct the Haar basis. The first is simply the tensor product $\psi_{j,j',k,k'}(x,y) = \psi_{j,k}(x)\psi_{j',k'}(y)$, so that each basis function $\psi_{j,j',k,k'}(x,y)$ is supported on a rectangle. The second basis is defined by the set of three kinds of basis functions supported on squares: $\psi_{j,k}(x)\psi_{j,k'}(y)$, $\psi_{j,k}(x)\varphi_{j,k'}(y)$, and $\varphi_{j,k}(x)\psi_{j,k'}(y)$, where $\varphi(x)$ is the characteristic function of the interval $(0,1)$ and $\varphi_{j,k}(x) = 2^{-j/2}\varphi(2^{-j}x - k)$.

Let us now turn to compactly supported wavelets with vanishing moments constructed in [4], following [10] and [8]. For a complete account we refer to [5]. We start with the definition of the multiresolution analysis.

Definition V.1 *A multiresolution analysis is a decomposition of the Hilbert space* $\mathbf{L}^2(\mathbf{R^d})$, $\mathbf{d} \geq 1$, *into a chain of closed subspaces*

$$\ldots \subset \mathbf{V}_2 \subset \mathbf{V}_1 \subset \mathbf{V}_0 \subset \mathbf{V}_{-1} \subset \mathbf{V}_{-2} \subset \ldots \tag{5.7}$$

such that

1. *$\bigcap_{j \in \mathbf{Z}} \mathbf{V}_j = \{0\}$ and $\bigcup_{j \in \mathbf{Z}} \mathbf{V}_j$ is dense in $\mathbf{L}^2(\mathbf{R^d})$*

2. *For any $f \in \mathbf{L}^2(\mathbf{R^d})$ and any $j \in \mathbf{Z}$, $f(x) \in \mathbf{V}_j$ if and only if $f(2x) \in \mathbf{V}_{j-1}$*

3. *For any $f \in \mathbf{L}^2(\mathbf{R^d})$ and any $k \in \mathbf{Z^d}$, $f(x) \in \mathbf{V}_0$ if and only if $f(x-k) \in \mathbf{V}_0$*

4. *There exists a function $\varphi \in \mathbf{V}_0$ such that $\{\varphi(x-k)\}_{k \in \mathbf{Z^d}}$ is a Riesz basis of $\mathbf{V}_0$.*

Let us define the subspaces $\mathbf{W}_j$ as an orthogonal complement of $\mathbf{V}_j$ in $\mathbf{V}_{j-1}$,

$$\mathbf{V}_{j-1} = \mathbf{V}_j \oplus \mathbf{W}_j, \tag{5.8}$$

so that

$$\mathbf{L}^2(\mathbf{R^d}) = \bigoplus_{j \in \mathbf{Z}} \mathbf{W}_j. \tag{5.9}$$

Selecting the coarsest scale n, we may replace the chain of the subspaces (5.7) by

$$\mathbf{V}_n \subset \ldots \subset \mathbf{V}_2 \subset \mathbf{V}_1 \subset \mathbf{V}_0 \subset \mathbf{V}_{-1} \subset \mathbf{V}_{-2} \subset \ldots, \tag{5.10}$$

and obtain

$$\mathbf{L}^2(\mathbf{R^d}) = \mathbf{V}_n \bigoplus_{j \leq n} \mathbf{W}_j. \tag{5.11}$$

If there is a finite number of scales then without loss of generality we set $j = 0$ to be the finest scale and consider

$$\mathbf{V}_n \subset \ldots \subset \mathbf{V}_2 \subset \mathbf{V}_1 \subset \mathbf{V}_0, \quad \mathbf{V}_0 \subset \mathbf{L}^2(\mathbf{R^d}) \tag{5.12}$$

instead of (5.10). In numerical realizations the subspace $\mathbf{V}_0$ is finite dimensional.

Let us consider a multiresolution analysis for $\mathbf{L}^2(\mathbf{R})$ and let $\{\psi(x - k)\}_{k\in\mathbf{Z}}$ be an orthonormal basis of $\mathbf{W}_0$. We will require that the function ψ has M vanishing moments,

$$\int_{-\infty}^{+\infty} \psi(x)x^m dx = 0, \qquad m = 0, \ldots, M - 1. \tag{5.13}$$

This property is critical for using wavelet transform as a tool for compression. Property (5.13) implies that the functions which (locally) are well approximated by polynomials have small wavelet coefficients. For the Haar basis $M = 1$ in (5.13).

Are there constructions other than the Haar basis with $M > 1$? The answer is positive and here we will only show how to derive equations (for the orthonormal case) and refer to [5] for numerous important details.

One of the immediate consequences of Definition V.1 is that the function φ may be expressed as a linear combination of the basis functions of $\mathbf{V}_{-1}$. Since the functions $\{\varphi_{j,k}(x) = 2^{-j/2}\varphi(2^{-j}x - k)\}_{k\in\mathbf{Z}}$ form an orthonormal basis of $\mathbf{V}_j$, we have

$$\varphi(x) = \sqrt{2} \sum_{k=0}^{L-1} h_k\varphi(2x - k). \tag{5.14}$$

In general, the sum in (5.14) does not have to be finite and by choosing a finite sum in (5.14) we are selecting compactly supported wavelets. We may rewrite (5.14) as

$$\hat{\varphi}(\xi) = m_0(\xi/2)\hat{\varphi}(\xi/2), \tag{5.15}$$

where

$$\hat{\varphi}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \varphi(x)\, e^{ix\xi}\, dx, \tag{5.16}$$

and the 2π-periodic function m_0 is defined as

$$m_0(\xi) = 2^{-1/2} \sum_{k=0}^{L-1} h_k e^{ik\xi}. \tag{5.17}$$

Assuming orthogonality of $\{\varphi(x - k)\}_{k\in\mathbf{Z}}$, we note that

$$\delta_{k0} = \int_{-\infty}^{+\infty} \varphi(x - k)\varphi(x)\, dx = \int_{-\infty}^{+\infty} |\hat{\varphi}(\xi)|^2\, e^{-ik\xi}\, d\xi, \tag{5.18}$$

and, therefore,

$$\delta_{k0} = \int_0^{2\pi} \sum_{l\in\mathbf{Z}} |\hat{\varphi}(\xi + 2\pi l)|^2\, e^{-ik\xi}\, d\xi, \tag{5.19}$$

and

$$\sum_{l\in\mathbf{Z}} |\hat{\varphi}(\xi + 2\pi l)|^2 = \frac{1}{2\pi}. \tag{5.20}$$

Using (5.15), we obtain

$$\sum_{l\in\mathbf{Z}} |m_0(\xi/2 + \pi l)|^2 |\hat{\varphi}(\xi/2 + \pi l)|^2 = \frac{1}{2\pi}, \tag{5.21}$$

6

and, by taking the sum in (5.21) separately over odd and even indices, we have

$$\sum_{l \in \mathbf{Z}} |m_0(\xi/2 + 2\pi l)|^2 |\hat{\varphi}(\xi/2 + 2\pi l)|^2$$

$$+ \sum_{l \in \mathbf{Z}} |m_0(\xi/2 + 2\pi l + \pi)|^2 |\hat{\varphi}(\xi/2 + 2\pi l + \pi)|^2 = \frac{1}{2\pi}. \tag{5.22}$$

Using the 2π-periodicity of the function m_0 and (5.20), we obtain (after replacing $\xi/2$ by ξ) a necessary condition

$$|m_0(\xi)|^2 + |m_0(\xi + \pi)|^2 = 1, \tag{5.23}$$

for the coefficients h_k in (5.17). On defining the function ψ by

$$\psi(x) = \sqrt{2} \sum_k g_k \varphi(2x - k), \tag{5.24}$$

where

$$g_k = (-1)^k h_{L-k-1}, \qquad k = 0, \dots, L-1, \tag{5.25}$$

or, equivalently, the Fourier transform of ψ by

$$\hat{\psi}(\xi) = m_1(\xi/2)\hat{\varphi}(\xi/2), \tag{5.26}$$

where

$$m_1(\xi) = 2^{-1/2} \sum_{k=0}^{k=L-1} g_k e^{\mathrm{i}k\xi} = -e^{-\mathrm{i}\xi}\overline{m}_0(\xi + \pi), \tag{5.27}$$

it is not difficult to show, that for each fixed scale $j \in \mathbf{Z}$, the wavelets $\{\psi_{j,k}(x) = 2^{-j/2}\psi(2^{-j}x - k)\}_{k \in \mathbf{Z}}$ form an orthonormal basis of $\mathbf{W}_j$.

Equation (5.23) can also be viewed as the condition for exact reconstruction for a pair of the quadrature mirror filters (QMFs) H and G, where $H = \{h_k\}_{k=0}^{k=L-1}$ and $G = \{g_k\}_{k=0}^{k=L-1}$. Such exact QMF filters were first introduced by in [14] for subband coding.

The decomposition of a function into the wavelet basis is an $O(N)$ procedure. Given the coefficients s_k^0, $k = 0, 1, \dots, N$ as "samples" of the function f, the coefficients s_k^j and d_k^j on scales $j \geq 1$ are computed at a cost proportional to N via

$$s_k^j = \sum_{n=0}^{n=L-1} h_n s_{n+2k}^{j-1}, \tag{5.28}$$

and

$$d_k^j = \sum_{n=0}^{n=L-1} g_n s_{n+2k}^{j-1}. \tag{5.29}$$

where s_k^j and d_k^j are viewed as periodic sequences with the period 2^{n-j}. If the function f is not periodic (e.g. defined on an interval), then such straightforward periodization introduces "artificial" singularities. A better alternative is to use wavelets on the interval [3]. In practice, this approach amounts to modifications near the boundary of both decomposition and reconstruction algorithms.

7

Computing via (5.28) and (5.29) is illustrated by the pyramid scheme

$$
\begin{array}{ccccccc}
\{s_k^0\} & \longrightarrow & \{s_k^1\} & \longrightarrow & \{s_k^2\} & \longrightarrow & \{s_k^3\} \quad \cdots \\
& \searrow & & \searrow & & \searrow & \\
& & \{d_k^1\} & & \{d_k^2\} & & \{d_k^3\} \quad \cdots .
\end{array}
\tag{5.30}
$$

The reconstruction of a function from its wavelet representation is also an order N procedure and is described by

$$
\begin{aligned}
s_{2n}^{j-1} &= \sum_{k=1}^{k=M} h_{2k} s_{n-k}^{j} + \sum_{k=1}^{k=M} g_{2k} d_{n-k}^{j}, \\[2mm]
s_{2n-1}^{j-1} &= \sum_{k=1}^{k=M} h_{2k-1} s_{n-k}^{j} + \sum_{k=1}^{k=M} g_{2k-1} d_{n-k}^{j}.
\end{aligned}
\tag{5.31}
$$

Computing via (5.31) is illustrated by the pyramid scheme

$$
\begin{array}{ccccccc}
\{s_k^n\} & \longrightarrow & \{s_k^{n-1}\} & \longrightarrow & \{s_k^{n-2}\} & \longrightarrow & \{s_k^{n-3}\} \quad \cdots \\
& \nearrow & & \nearrow & & \nearrow & \\
\{d_k^n\} & & \{d_k^{n-1}\} & & \{d_k^{n-2}\} & & \{d_k^{n-3}\} \quad \cdots
\end{array}
$$

We observe that once the filter H has been chosen, it completely determines the functions φ and ψ and, therefore, the multiresolution analysis. Moreover, in properly constructed algorithms, the values of the functions φ and ψ are usually never computed. Due to the recursive definition of the wavelet bases, all the manipulations are performed with the quadrature mirror filters H and G, even if they involve quantities associated with φ and ψ.

As an example, let us compute the moments of the scaling function φ,

$$
\mathcal{M}_m = \int x^m \, \varphi(x) \, dx, \quad m = 0, \ldots, M-1,
\tag{5.32}
$$

in terms of the filter coefficients $\{h_k\}_{k=0}^{k=L-1}$. Applying operator $(\frac{1}{i} d/d\xi)^m$ to both sides of (5.15) and setting $\xi = 0$, we obtain

$$
\mathcal{M}_m = 2^{-m} \sum_{j=0}^{j=m} \binom{m}{j} \mathcal{M}_j \, \mathcal{M}_{m-j}^{h},
\tag{5.33}
$$

where

$$
\mathcal{M}_l^h = 2^{-\frac{1}{2}} \sum_{k=0}^{k=L-1} h_k k^l, \quad l = 0, \ldots, M-1.
\tag{5.34}
$$

Thus, we have from (5.33)

$$
\mathcal{M}_m = \frac{1}{2^m - 1} \sum_{j=0}^{j=m-1} \binom{m}{j} \mathcal{M}_j \, \mathcal{M}_{m-j}^{h},
\tag{5.35}
$$

with $\mathcal{M}_0 = 1$ and we can use (5.35) as a recursion to compute the moments.

For the purposes of compression biorthogonal bases are often used [2] but we will not go into details here.

8

VI   What does wavelet transform accomplish?

Although the following consideration deals with wavelet transform, similar points can be made for wavelet packets and local trigonometric bases.

How the coherent portion of the signal in a seismogram appears as a local correlation that our eye easily identifies. Wavelet transform reduces the number of significant coefficients necessary to represent the seismogram locally and, thus, decorrelates the coherent portion of the signal. Vanishing moment are the key to such performance. On the other hand the wavelet transform does not decorelate (or compress) the random Gaussian noise. Thus, the result of application of the wavelet transform is that the coherent portion of the signal will now reside in a relatively few large coefficients whereas the rest of the coefficients will describe a portion of the signal that is like the "random Gaussian noise". Unfortunately, this heuristics is not precise and there is no theorem asserting the result. Yet, a number of mathematically justified statements are available that indicate that the heuristic picture above is basically correct [1]. For other choices of bases with controlled localization in time-frequency domain similar statements can be made and, in practice, their performance can be somewhat better than that with wavelet bases.

How many vanishing moments to use is practice for a wavelet basis? Since the length of the filter is proportional to the number of vanishing moments, there has to be a balance between the number of vanishing moments and the performance of filters. The answer is obtained experimentally and, typically, $2 - 4$ vanishing moments are sufficient.

The histogram for the wavelet coefficients has a sharp peak near zero. This precisely what is needed by quantization and encoding steps.

Let us finish by describing "the state of the art" of seismic data compression. The geophysical community was introduced to compression techniques in [12] and [11]. Further progress was reported in [15]. Currently codes are available [3] for compression of seismic data at about 10 MB/sec, and for its decompression at about 15-18 MB/sec (this measurement is for SGI Origin and roughly corresponds to the speed of reading data from the hard drive). An example of the distortion curve for noisy land prestack data is shown in Figure 1. Typically for marine seismic data compression results are better than for land data (the data contain less noise).

Finally, it is important to point out that seismic procedures such as stacking or migration reduce the distortion introduced by the transform coding. The reason is that averaging tends to separate further the significant vs. non-significant coefficients selected by decorrelating the coherent portion of the signal. An example is shown in Figure 3 for a marine data set.

References

[1] A. Cohen and J-P. D'Ales. Nonlinear approximation of random functions. *SIAM J. Appl. Math.*, 57(2):518–540, 1997.

[2] A. Cohen, I. Daubechies, and J.-C. Feauveau. Biorthogonal bases of compactly supported wavelets. *Comm. Pure and Appl. Math.*, 45(5):485–560, Jun. 1992.

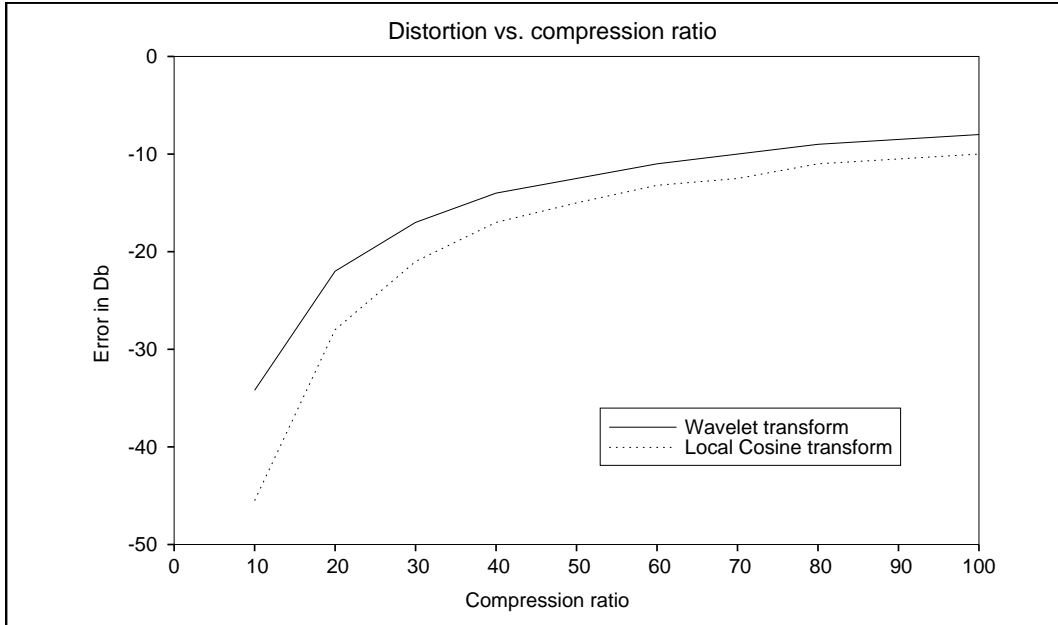[3]GeoEnergy Corporation, 2530 East 71 Street, Suite E, Tulsa, OK 7413

Figure 1: Distortion Curve: wavelet vs. local cosine transform for land data
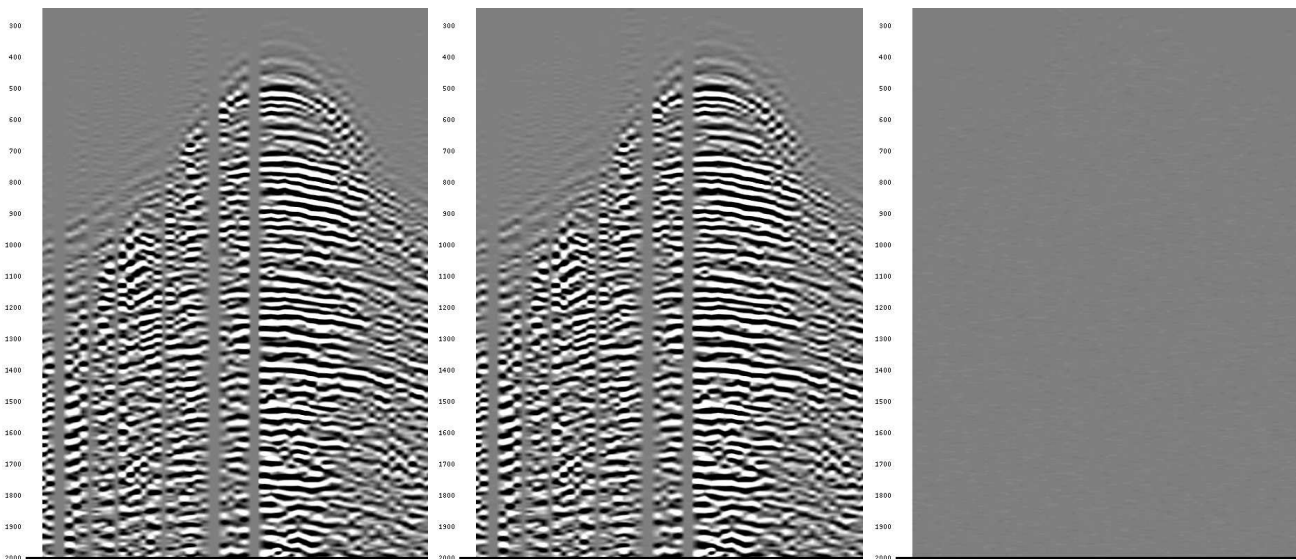

Figure 2: Land data: original, compressed by the factor of 10, and the difference between them
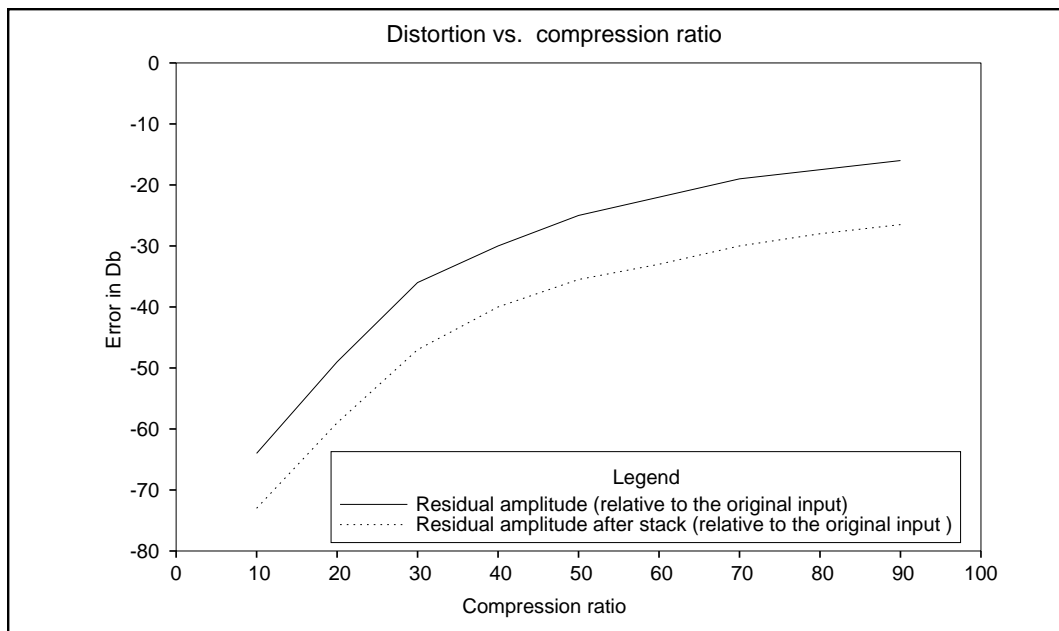
Figure 3: Stacking reduces distortion introduced by the transform coding (LCT in this example).

[3] A. Cohen, I. Daubechies, and P. Vial. Wavelets on the interval and fast wavelet transforms. *Applied and Computational Harmonic Analysis*, 1(1):54–81, 1993.

[4] I. Daubechies. Orthonormal bases of compactly supported wavelets. *Comm. Pure and Appl. Math.*, 41:909–996, 1988.

[5] I. Daubechies. *Ten Lectures on Wavelets*. CBMS-NSF Series in Applied Mathematics. SIAM, 1992.

[6] A. Haar. Zur Theorie der orthogonalen Funktionensysteme. *Mathematische Annalen*, pages 331–371, 1910.

[7] D. Huffman. A method for the construction of minimum redundancy codes. *Proc. of IRE*, 40:1098–1101, 1952.

[8] S. Mallat. Review of multifrequency channel decomposition of images and wavelet models. Technical Report 412, Courant Institute of Mathematical Sciences, New York University, 1988.

[9] S. Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, San Diego, 1998.

[10] Y. Meyer. Wavelets and operators. In N.T. Peck E. Berkson and J. Uhl, editors, *Analysis at Urbana*. London Math. Society, Lecture Notes Series 137, 1989. v.1.

[11] R. Ergas P. Donoho and J.D. Villasenor. High-performance seismic trace compression. Soc. Expl. Geophys, 1995. 65th Annual Intern. Conv. Soc. Expl. Geophys.

[12] E.C. Reiter and P.N. Heller. Wavelet transformation-based compression of nmo-corrected cdp gathers. Soc. Expl. Geophys, 1994. 64th Annual Intern. Conv. Soc. Expl. Geophys.

[13] J.M. Shapiro. Embedded image coding using zerotrees of wavelet coefficients. *IEEE Trans. Signal Processing*, 41(12):3445–3462, 1993.

[14] M. J. Smith and T. P. Barnwell. Exact reconstruction techniques for tree-structured subband coders. *IEEE Transactions on ASSP*, 34:434–441, 1986.

[15] A.A. Vassiliou and M.V. Wickerhauser. Comparison of wavelet image coding schemes for seismic data compression. Soc. Expl. Geophys, 1997. 67th Annual Intern. Conv. Soc. Expl. Geophys.

[16] I. Witten, R.Neal, and J. Cleary. Arithmetic coding for data compression. *Comm. of the ACM*, 30(6):519–540, 1987.