ON PROGRESSIVE SEISMIC DATA COMPRESSION USING GENLOT

Laurent C. Duval[1] *Truong Q. Nguyen*[2] *Trac D. Tran*[3]

[1]Institut Français du Pétrole, Rueil-Malmaison, France
[2]Boston University, Boston, MA, USA
[3] The Johns Hopkins University, Baltimore, MD, USA

ABSTRACT

Wavelet and subband coding have been shown effective techniques for seismic data compression, especially when compared to DCT-based algorithms (such as JPEG), which suffer from blocking artifact at low bit-rates. The transforms remove statistical redundancy and permit efficient compression. This paper presents a novel use of the Generalized Lapped Orthogonal Transforms (GenLOTs) for compression of 24-bits seismic data. The proposed implementation provides a better frequency partition than that of the wavelet scheme. It is used with in the Embedded Zerotree Wavelet (EZW) framework, an embedded quantization scheme, allowing exact bit rate compression and incorporating quality control features. The proposed coder has better performance in SNR comparing to the state of the art Set Partitionning in Hierarchical Trees algorithm (SPIHT).

1. INTRODUCTION

Seismic data processing is very crucial in oil and gas exploration. A typical seismic acquisition system consists of several large arrays of seismometers collecting data that have to be transmitted, stored, processed and interpreted. Although storage of huge datasets (requiring terabytes for 3-D surveys for example [1]) is one of the main concerns, their transmission poses a much more challenging problem. One needs efficient data compression algorithms, that do not alter their interpretation after compression and processing. In other words, compression artifacts should preserve the crucial geological information.

Unlike most natural images, where lowpass frequency is dominant, seismic data are inherently bandpass signals with large amplitude discontinuities, short scale unstationarities and non-homogeneities which make them harder to compress.

The first significant work on seismic data compression was written in 1974 by Wood [2], using methods based on Walsh transform, essentially for plotting purposes. Other works used discrete cosine transform in lossy compression and linear predictive coding in lossless compression. Recently, seismic data compression based on wavelet transform is a part of the acquisition process [3, 4]. Compression results are further improved by using more general subband coding schemes, exploiting statistical properties of the signal [1].

High frequency features which are essential in the processing and interpretation of seismic data are attenuated by the ground structure and are further corrupted in the quantization process. Moreover, the poor frequency allocation at high frequency range in the dyadic wavelet transform decreases the effectiveness of the compression algorithm. The wavelet packet can be used to improve the performance but might not be attainable because of the limited computation resources at the seismometers.

In order to improve the frequency partition and still retaining the efficient features of the EZW paradigm, we propose a GenLOT-based coder for seismic images [5, 6, 7]. This coder works on overlapping image blocks, allowing parallel processing mode. Besides better frequency allocation, the proposed coder retains attractive properties such as embedded quantization, exact bit rate control and, most importantly for seismic data acquisition, progressive transmission which can be used as an embedded quality control feature.

Section 2. describes the seismic data and section 3. gives a brief review of wavelet and GenLOT coder. The new results and their analysis are given in section 4., followed by conclusions and scopes of future work in section 5..

2. SEISMIC DATA

There are many forms of seismic data, from 1-D single traces to raw shot point gathers (SP), common-depth point and common-offset gathers, stacked, migrated sections or blocks [8]. Each of them corresponds to a different type of acquisition or processing step. Seismic data can be of one to four dimensions, and exhibit distinct properties, especially for modelling purposes.

We are interested in field transmission and hence focus on unprocessed 24-bit SP images, made of time-sampled, high-precision outputs of spatially distributed sensors when a seismic shot is fired (as shown in Fig. 1). Since modeling these images is difficult and acquisition systems do not have sufficient computation power, we restrict the transforms to a few predesigned sets. Note also that the images used (like in Figure 1) have different dimensions on the two axes. This limits the number of iteration levels in wavelet-based algorithms. In this study, several kinds of SP datasets are used, such as land-explosive, marine-airgun and synthetic data.

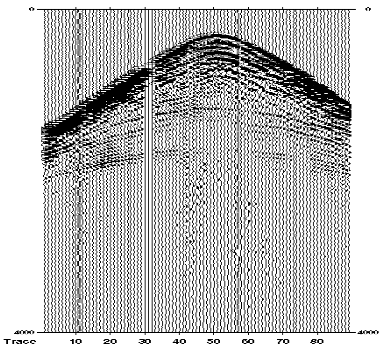

Figure 1. Wiggle plot of a seismic shot gather.

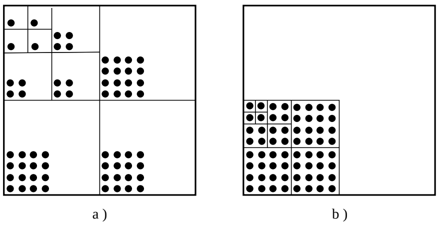
a)
b)

Figure 2. a) 3-level dyadic frequency partition and b) uniform 4-channel frequency partition. Shaded dots represent the DC components.

3. TRANSFORM-BASED CODING

A typical subband tranform-based coder has three stages: subband transform, quantization/thresholding and entropy coding. In this study, the EZW coding scheme is used in both wavelet-based and GenLOT-based coders. Wavelets and GenLOT transforms are briefly reviewed below for 1-D signal processing. 2-D separable transforms are used in this study.

3.1. The Wavelet Transform

From a filter bank's point of view, the discrete dyadic wavelet transform of a signal is obtained by iterating a maximally decimated 2-channel perfect reconstruction filter bank (FB). The dyadic wavelet decomposition iterates on the lowpass outputs, while the wavelet packets may iterate on any output. Figure 2a shows the frequency partition of a 3-level dyadic wavelet decomposition where the DC components are denoted in graycolor.

Wavelets can be either orthogonal or biorthogonal. Orthogonal wavelet transforms are not used in seismic signal compression since they do not have linear phase and very often, the phase information is crucial in geophysics. This is one of the reasons why symmetric biorthogonal wavelets [4] are popular. Wavelets also provide long, implicitly overlapping bases, and lead to easy interpretation in terms of multiresolution, with coarse approximation and details. We refer to [9] for a comprehensive survey on wavelets and filter banks.

3.2. The GenLOT

The 8×8 discrete cosine transform employed by JPEG can be viewed as a filter bank with 8 channels and 8-tap filters. It belongs to a wider class of filter banks (GenLOT) which have linear phase and are paraunitary. Figure 2b shows the frequency partition of a 4-channel GenLOT where the DC components are denoted in dark color. As noticed, the frequency partitioning of a 4-channel GenLOT is comparable to that of a wavelet packet with 2 levels of iteration on every output.

Transforms with non-overlapping filters such as the DCT can be computed with simple computations but the reconstructed images usually suffer from annoying blocking artifacts at low bit rates. Lapped orthogonal transforms (LOT) were developed to overcome this artifact: overlapping basis functions smooth out the block boundary. To reduce blocking artifacts even further, for instance at lower bit-rates, de Queiroz *et al.* [5] developed a more general class of transforms, the generalized lapped orthogonal transforms (GenLOT). GenLOT is a M-channel linear phase paraunitary filter bank with filter length KM where DCT and LOT are special low-order cases. As shown in [6], GenLOT's transform coefficients can be rearranged to fit the conventional zerotree data structure: the DC components of each subband (small black squares, fig. 2b) are gathered to produce a lower resolution version of the whole image, and the AC coefficients can be grouped accordingly to yield the representation shown in the bottom of fig. 3. To decorrelate the lowpass subimage, we apply a 9/7 wavelet decomposition with one-level (because the number of coefficients are fewer on the x-direction) on the DC subband.

GenLOT is designed using lattice structure and weighted optimization where different objective functions, such as *coding gain*, *DC leakage*, *stop-band attenuation*, are used. For instance, higher coding gain leads to more energy compaction and typically yields excellent coding performance. Details on GenLOT design can be found in [6].

3.3. Embedded Zerotree Algorithms

First introduced by Shapiro in [10], the main ideas behind EZW coding are (a) the most important information (here the larger coefficients) should be transmitted first, (b) the values of these coefficients can be progressively refined and (c) spatial correlation between coefficients from differents subbands should be exploited in the tree structure. The coefficients are scanned according to spatial dependencies between the coefficients of the subbands. The three wavelet trees whose roots are on the DC "upper left corner" image are then coded with a four symbol alphabet which is encoded using adaptive arithmetic coding. We refer to [11, 12] for practical implementation.

4. RESULTS

4.1. Measures

Let s_n and b_n be the seismic signal and its reconstruction error after compression. The most widely used distorsion measure is the signal/noise ratio,

$$SNR = 10\, log_{10} \left(\sum_n s_n^2 / \sum_n b_n^2 \right), \tag{1}$$

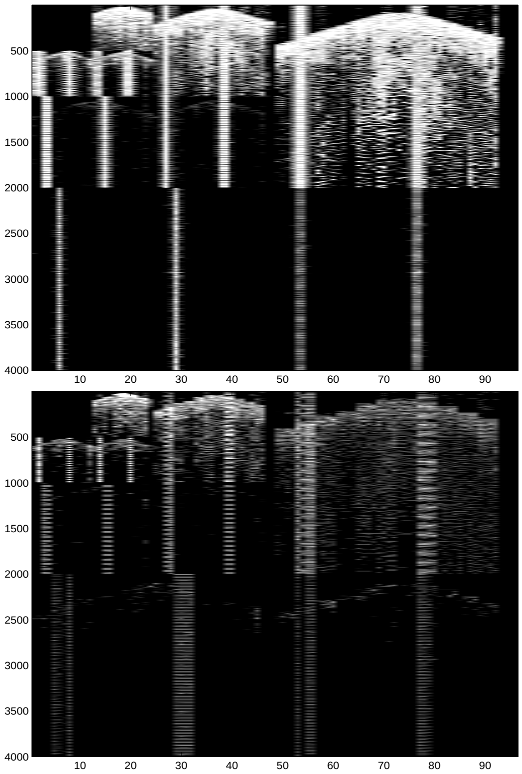

Figure 3. Wavelet (top) and GenLOT (bottom) thresholded transformed images.

where high SNRs typically yield better results. However, we will also use an "absolute" signal/noise ratio (following [13]), defined as:

$$ASNR = 20\, log_{10} \left(\sum_n |s_n| / \sum_n |b_n| \right) \qquad (2)$$

These measures give different insights to the distorsion induced by the compression.

4.2. Comparison and results

As for still images [14], some recent works stated that the 9/7 biorthogonal wavelet provide good bases for seismic signal decomposition [4, 13]. In this study, the performance of the 9/7 biorthogonal wavelet is compared to several choices of GenLOTs (with different length and number of channels). Figure 3 shows the transform image of a SP image using the 9/7 wavelet and using the GenLOT. These images have been thresholded according to a function of their mean pixel values, for interpretation purposes. Since EZW is used on the transformed images, one needs to verify whether the transform image has good zerotrees or not. We noticed that the non-zero coefficients are better concentrated in the GenLOT transform, yielding larger zero-regions after quantization, and generating more zerotrees.

Table 1 shows several choices of GenLOTs used in the study, ranging from the DCT with 8 channels and length 8 to a GenLOT with 16 channels, that has good compression performance for conventional images [6]. The DCT is used as the transform for the horizontal dimension because of the limited data size, uncorrelated information and its decent performance (as compared to the other GenLOTs). Our experiments confirm that longer, overlapping bases in the vertical direction are necessary to get better results with seismic images.

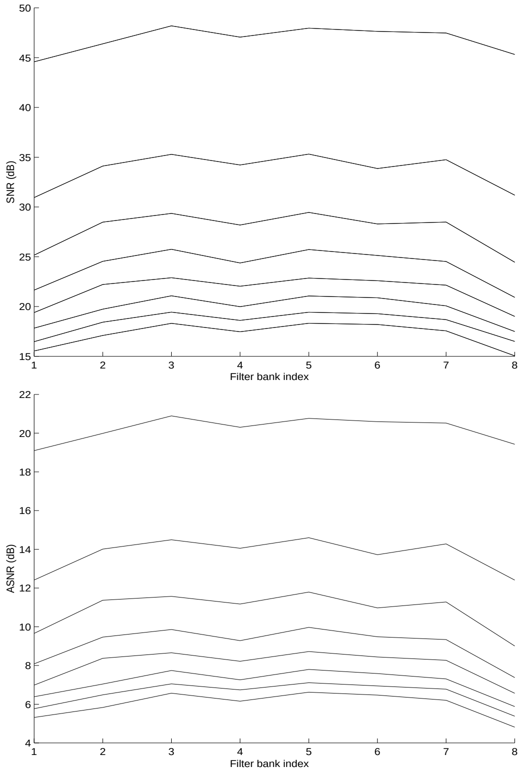

Figure 4. SNR and ASNR comparison for 10:1 (top) to 80:1 (bottom) compression ratios for several GenLOTs.

Name	Channels	Length	Index
DCT	8	8	1
LOT16	16	32	2
LOT85cg	8	40	3
LOT86	8	48	4
LOT86cgdc	8	48	5
LOT86cgmax	8	48	6
LOT86fr	8	48	7
LOT86frmax	8	48	8

Table 1. GenLOTs used in the study.

Figure 5 shows the objective comparison with SPIHT, using the results of the two best GenLOTs: the LOT85cg GenLOT with 8 channels and length 40 and the LOT86cgdc GenLOT with 8 channels and length 48. All computed error measures are obtained from actual compressed files where overheads have been accounted for.

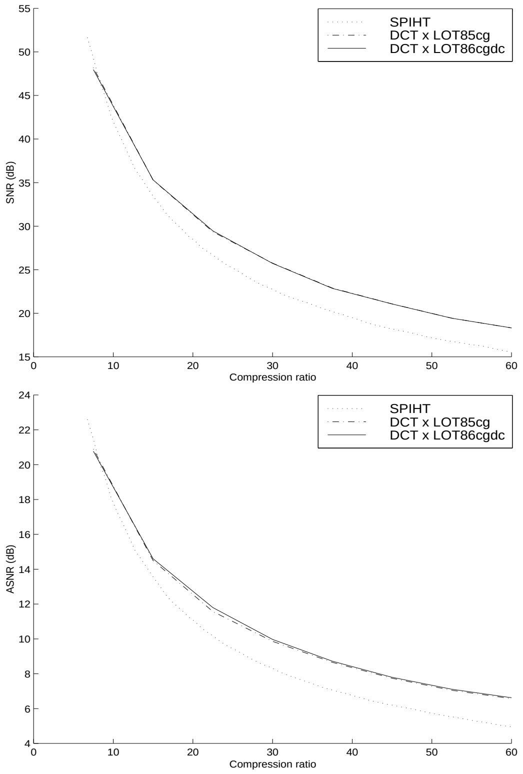

Figure 5. SNR and ASNR distorsion curves.

4.3. Result analysis

Figure 4 displays the SNR and ASNR for the GenLOT-based coders (in table 1) at various compression ratios ranging from 10 to 80. Figure 5 shows that, at most compression ratios (between 7 and 60), the GenLOT coder performs better comparing to SPIHT, where the difference increases as the compression rate increases (3-4 dB in SNR, 2-3 dB in ASNR around 60 : 1). One also observed that the performances of the two best GenLOTs in both SNR and ASNR are similar. One should notice that the same trends occur in both SNR and ASNR plots and since large errors (outliers) contribute more in the SNR measure, the outliers percentage is small. Similar improvements over the SPIHT coder were also observed for marine and synthetic datasets.

5. CONCLUSIONS AND FUTURE WORK

We propose a seismic image coder using GenLOT and embedded zerotree algorithm. The proposed coder provides better frequency partitioning than that of dyadic wavelet, especially for non-smooth signals. It provides exact bit rate control in a zerotree embedded quantization scheme and progressive transmission can be used for quality control. The block transform implementation allows parallel processing. For seismic signal, our coder outperforms the state of the art SPIHT coder at most bit rates.

In future works, we plan to design optimal GenLOTs for various seismic data sets by studying the statistical properties of the seismic data. Moreover, we also plan to adapt the set partitionning in the EZW algorithm.

6. ACKNOWLEDGEMENTS

We are would like to thank Dr. Van Bui-Tran (IFP) for fruitful discussions on seismic data compression.

REFERENCES

[1] J. M. Lervik, T. Røsten, T. A. Ramstad, “Subband Seismic Data Compression: Optimization and Evaluation”, *IEEE Digital Signal Processing Workshop Proceedings*, p. 65–68, June 1996.

[2] L. C. Wood, “Seismic Data Compression Methods”, *Geophysics*, vol. 39, 4, p. 499–525, August 1974.

[3] P. L. Donoho, R. A. Ergas, J. D. Villasenor, “High-performance seismic trace compression”, *65^{th} SEG Annual Meeting, 1995*, p. 160–163.

[4] J. D. Villasenor, R. A. Ergas, P. L. Donoho, “Seismic data compression using high-dimensional wavelet transforms”, *Proceedings DCC 1996*, p. 396–405.

[5] R. L. de Queiroz, T. Q. Nguyen and K. R. Rao, “The GenLOT: Generalized Linear-Phase Lapped Orthogonal Transform”, *IEEE Trans. on Image Processing*, 44–3, March 1996.

[6] T. D. Tran, T. Q. Nguyen, “A progressive transmission image coder using linear phase uniform filter banks as block transforms”, *IEEE Trans. on Image Processing*, to appear.

[7] T. D. Tran, T. Q. Nguyen, “Block Transforms in Progressive Image Coding”, in *Wavelet Image and Video Compression*, edited by P. N. Topiwala, Kluwer Academic, Norwell, MA, June 1998.

[8] Ö. Yilmaz, *Seismic data processing*, Society of Exploration Geophysicists, 1987

[9] T. Q. Nguyen and G. Strang, *Wavelets and Filter Banks*, Wellesley-Cambridge Press, 1996.

[10] J. M. Shapiro, “Embedded image coding using zerotrees of wavelet coefficients”, *IEEE Trans. on Signal Processing*, vol. 41, p. 3445–3462, December 1993.

[11] A. Said, W. Pearlman, “A New Fast and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees”, *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 6, p. 243–250, June 1996.

[12] A. Said, W. Pearlman, “An Image Multiresolution Representation for Lossless and Lossy Image Compression”, *IEEE Transactions on Image Processing*, vol. 5, p. 1303–1310, Sept. 1996.

[13] A. Vassiliou, M. V. Wickerhauser, “Comparison of Wavelet Image Coding Schemes for Seismic Data Compression”, *Proceedings of SPIE*, vol. 3169, 1997.

[14] M. Antonini, M. Barlaud, P. Mathieu and I. Daubechies,“Image coding using the wavelet transform”, *IEEE Transactions on Image Processing*, 1, p. 205–220, 1992.

ON PROGRESSIVE SEISMIC DATA COMPRESSION USING GENLOT

Laurent C. Duval[1] , *Truong Q. Nguyen*[2] *and Trac D. Tran*[3]
[1]Institut Français du Pétrole, Rueil-Malmaison, France
[2]Boston University, Boston, MA, USA
[3] The Johns Hopkins University, Baltimore, MD, USA

Wavelet and subband coding have been shown effective techniques for seismic data compression, especially when compared to DCT-based algorithms (such as JPEG), which suffer from blocking artifact at low bit-rates. The transforms remove statistical redundancy and permit efficient compression. This paper presents a novel use of the Generalized Lapped Orthogonal Transforms (GenLOTs) for compression of 24-bits seismic data. The proposed implementation provides a better frequency partition than that of the wavelet scheme. It is used with in the Embedded Zerotree Wavelet (EZW) framework, an embedded quantization scheme, allowing exact bit rate compression and incorporating quality control features. The proposed coder has better performance in SNR comparing to the state of the art Set Partitionning in Hierarchical Trees algorithm (SPIHT).