Seismic data compression and QC using GenLOT

Laurent C. Duval[*] Truong Q. Nguyen [†] Trac D. Tran[‡]

Summary

Modern seismic surveys with higher-precision numerization (24-bits A/D converters) have led to ever increasing amounts of seismic data. Management of these large datasets becomes critical, not only for transmission, but also for storage, processing and interpretation. Compression algorithms have been proposed in the geophysicists' community over the past few years as a way to effectively manage seismic data. For instance, wavelet based compression algorithms can represent seismic data using only a fraction of the original data size. Recently, several works on generalized lapped transform (GenLOT) demonstrate its advantage over wavelets for conventional image compression. Their features are well suited to seismic data properties and have led to better results than for wavelets, in terms of signal-to-noise ratio.

In this study we compare GenLOT with wavelet compression results. We have implemented algorithms in a new embedded scheme allowing exact bit rate compression, and providing straight-forward quality control (QC) checks. Simulations are performed on several types of datasets, including shot, CDP gathers and stack sections. The paper is organized as follows: Section 2 briefly reviews seismic data compression, discusses transform-based coding and the idea of embedded zerotree coding. Here one can easily incorporate quality control issue as part of embedded zerotree coding algorithm. Section 3 reviews the GenLOT theory and motivates why GenLOT transform should give better performance as compared to wavelet transform. Section 4 discusses the simulation results, comparing the performance of the GenLOT-based and wavelet-based algorithms. Section 5 concludes the paper and discusses future work.

1 Introduction and review of previous works

Several works use wavelet-based image compression methods to compress seismic data for high compression ratios. Specifically, the algorithms by Luo and Schuster ([1], 1992) and Ergas *et al.* ([2], 1996) have been shown to be very effective in compressing seismic data and have led to actual field transmission in Stigant *et al.* [3]. Recently, several authors (Tran and Nguyen in [4]) have demonstrated GenLOT's advantage over wavelet compression for conventional images. GenLOT is a block transform with overlap which provides better frequency partitioning than wavelets. It is also applied locally on the non stationary seismic data, without blocking artifact. As a result, it leads to better performance in terms of signal-to-noise ratio and visual interpretation. Figure 1 shows a raw land shot gather and its corresponding decomposition using GenLOT transform and wavelet transform. As we can see, the variances of the high-frequency subbands are smaller in the case of GenLOT transform.

Transform-Based Coding

A typical subband tranform-based coder includes three stages (Fig. 2): *transform*, *quantization* and *entropy coding*. The transform stage (*e.g.* wavelet or GenLOT) reduces redundancy in the dataset, as we can see in Figure 1. *Quantization* is the lossy stage, since it reduces the dynamics of the transform coefficients. *Entropy coding* helps removing the remaining redundancy by exploiting the statistical properties of the quantized coefficients. Our algorithms perform an embedded quantization in the *entropy coding* stage, according to the technique described in the next paragraph.

Embedded Zerotree Wavelet

First introduced by Shapiro in [5], the main ideas behind Embedded Zerotree Wavelet coding (EZW) are

[*]Institut Français du Pétrole

[†]Boston University

[‡]The Johns Hopkins University

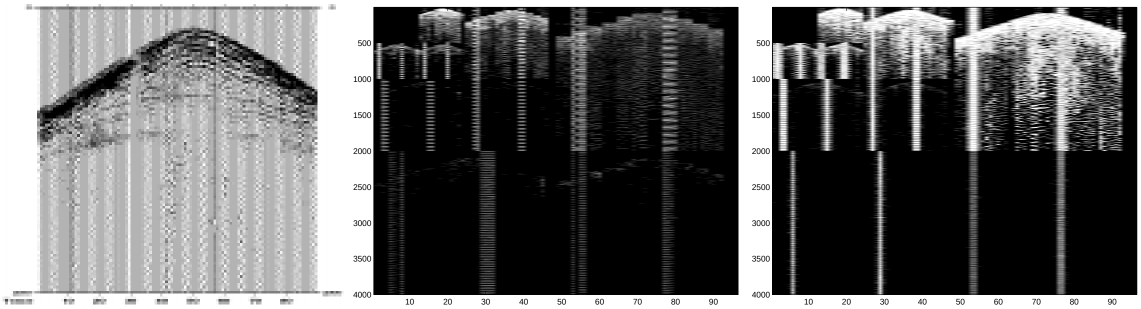

Figure 1: Shot gather, with GenLOT and wavelet decomposition.

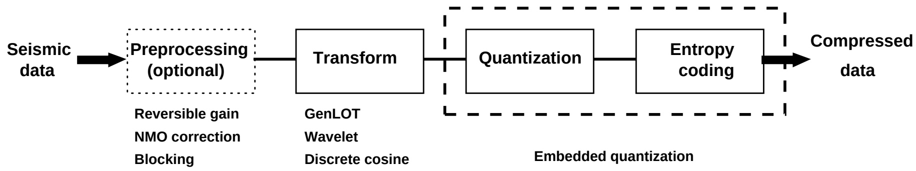

Figure 2: Typical block diagram for image compression.

(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 different subbands should be exploited in a 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 Said and Pearlman [6] for practical implementation.

Quality Control

The EZW encoding scheme possesses the interesting feature of progressive coding. The algorithm is refinable, *i.e.* the bits of the compressed file go from the coarsest level (main features) down to the finest one (remaining details). Figure 3 shows the original data, its $100:1$ compressed quality-control data and its $20:1$ reconstructed data. Consider the case where one would like to send the $20:1$ compressed data with an active quality control mode. One can first send the $100:1$ compressed data as a quality-control data and if the quality control is satisfactory, the remaining 80 % of the compressed file can be sent in order to reconstruct the $20:1$ compressed data. There is no redundancy in the transmission.

2 Review of GenLOT transform

Image compression standard JPEG employs a 8×8 discrete cosine transform (DCT), which is a uniform filter bank with length-8 symmetric basis functions. Figure 1 shows the frequency partitions of a 4-channel uniform filter bank (middle figure) and of a dyadic wavelet (right figure), which can be seen as an iterated 2-channel filter bank. Although DCT can be efficiently computed with simple computations and parallel mode, its reconstructed data suffer from blocking artifacts at medium and high compression ratios. Same blocking artifacts occur when seismic images are reorganized into smaller blocks which have to be compressed independently, due to the limited memory size on seismic equipments. To reduce these artifacts, for instance at higher compresssion ratios, de Queiroz *et al.* [7] have developed GenLOT, which is a more general class of transforms. GenLOT is a M-channel linear phase paraunitary filter bank. Linear phase is crucial in many seismic applications. The paraunitary property allows perfect reconstruction of the data in absence of quantization. GenLOT's filter length is KM, where K is the number of overlapped blocks. Within this scheme, spatially correlated data are compressed according to the statistical properties of the neighboring data. Dependences between correlated blocks remove the blocking artifacts. GenLOT's transform coefficients can be rearranged to fit the EZW structure.

The upper left corner (coarse approximation) in Figure 1 has been zeroed out for interpretation purposes. In our implementation, a one-level wavelet decomposition is applied at the coarse approximation to further decorrelate the data. Since GenLOT is a block transform, its implementation is efficient for application with limited memory cache. Reference [4] discusses design issues and comparisons to wavelet-based image coding.

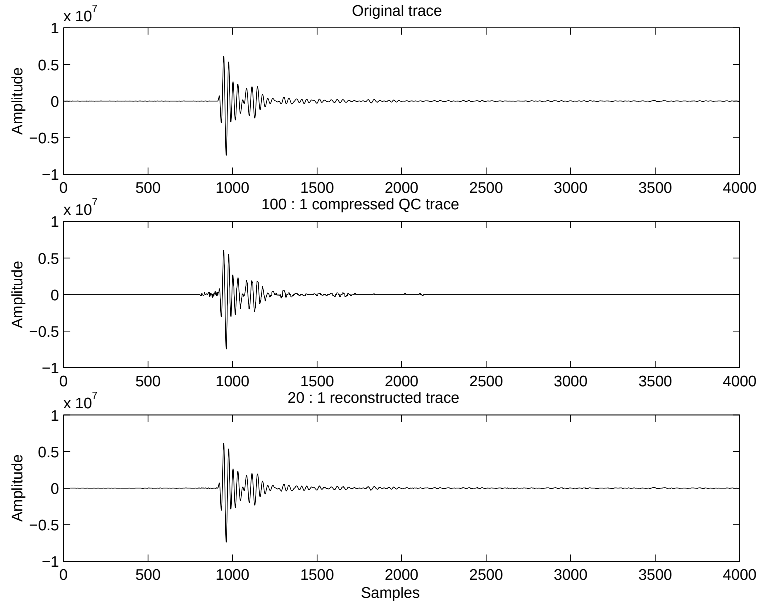

Figure 3: QC transmission and reconstructed data.

3 Comparison and results analysis

Let s_n and b_n be the seismic signal and its reconstruction error after compression. Two different measures are used in the study: the conventional signal-to-noise ratio (SNR), and an *absolute* signal-to-noise ratio (ASNR) following Reiter and Hall ([8], 1996) and Vassiliou and Wickerhauser ([9], 1997). These measures give different insights to the compression induced distortion.

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

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). Since EZW is used on the transformed images, one needs to verify whether the transform image has good zerotrees property 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.

This study uses several choices of GenLOTs, ranging from the DCT with 8 channels and length 8 to a GenLOT with 16 channels, that has good compression performance for conventional images [4]. The DCT is used as the transform for the horizontal dimension because of the limited data size, uncorrelated information and its reasonable performance (as compared to the other GenLOTs). Our experiments confirm that longer, overlapping bases in the vertical direction are necessary to get better results in seismic images. LOT85cg and LOT86cgdc with 8 channels and length 40 and 48 respectively have given the best results. Figure 4 shows that, at most compression ratios (between 7 and 60), the GenLOT-based coder performs better comparing to wavelet-based coder, the difference increases as the compression rate increases (3-4 dB in SNR around 60 : 1). One should note 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 wavelet coder were also observed for marine, synthetic datasets and stack sections (between 2 and 5 dB in SNR).

4 Conclusion and future work

We propose a seismic image coder using GenLOT and an 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 wavelet compression at most bit rates. Preprocessing like NMO correction

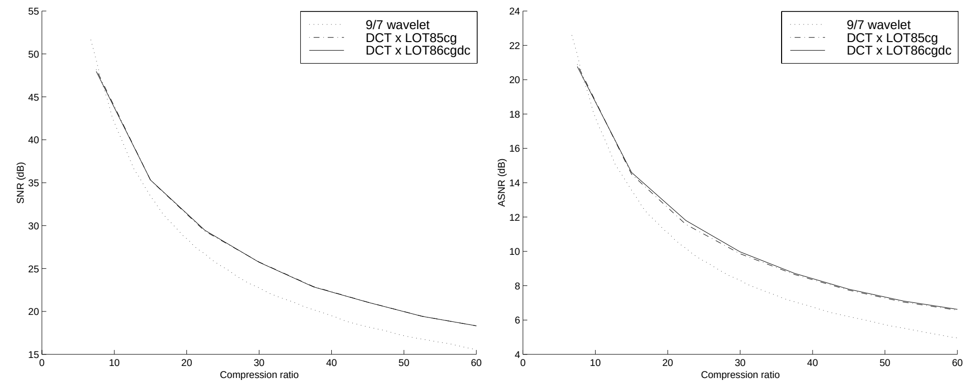

Figure 4: SNR and ASNR distorsion curves.

and reversible gain can be applied on the data as for wavelet compression, in order to improve the quality of the reconstructed data.

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 partitioning in the EZW algorithm.

5 Acknowledgments

We would like to thank Mr. Marc Becquey (IFP) for his useful comments during this work. We are also grateful to Dr. Van Bui Tran (IFP) for many valuable discussions on seismic data compression.

References

[1] Yi Luo and G. T. Schuster. Wave packet transform and data compression. In *Proc. 62nd Annual SEG Int. Meeting*, pages 1187–1190, 1992.

[2] R. A. Ergas, R. S. Polzer, P. L. Donoho, and P. Y. Galibert. Pitfalls in compressing land seismic trace data. In *Proc. 58th EAGE Conference*, 1996.

[3] J. P. Stigant, R. A. Ergas, P. L. Donoho, A. S. Minchella, and P. Y. Galibert. Field trial of seismic compression for real time transmission. In *Proc. 65th Annual SEG Int. Meeting*, pages 960–962, Oct. 1995.

[4] T. D. Tran and T. Q. Nguyen. A progressive transmission image coder using linear phase uniform filter banks as block transforms. *IEEE Trans. on Image Processing*, 1997. To appear.

[5] J. M. Shapiro. Embedded image coding using zerotrees of wavelet coefficients. *IEEE Trans. on Signal Processing*, 41:3445–3462, Dec. 1993.

[6] A. Said and W. A. Pearlman. A new fast/efficient image codec based on Set Partitioning in Hierarchical Trees. *IEEE Trans. on CAS for Video Technology*, 6:243–250, Jun. 1996.

[7] R. L. de Queiroz, T. Q. Nguyen, and K. R. Rao. The GenLOT: generalized linear-phase Lapped Orthogonal Transform. *IEEE Trans. on Signal Processing*, 44(3):497–507, Mar. 1996.

[8] E. Reiter and M. Hall. A comparison of multi-dimensional wavelet compression methods. In *Proc. 58th EAGE Conference*, 1996.

[9] A. Vassiliou and M. V. Wickerhauser. Comparison of wavelet image coding schemes for seismic data compression. In *Proc. 67th Annual SEG Int. Meeting*, volume 2, pages 1334–1337, 1997.