Seismic data compression: a comparative study between GenLOT and wavelet compression

Laurent C. Duval[a] and Truong Q. Nguyen[b]

[a]Institut Français du Pétrole, 1 et 4 avenue de Bois-Préau
92852 Rueil-Malmaison Cedex, FRANCE
[b]Boston University, ECE Department, 8 St. Mary's street
Boston MA 02215, USA

ABSTRACT

Generalized Lapped Orthogonal Transform (GenLOT) based image coder is used to compress 2-D seismic data sets. Its performance is compared to the results using wavelet-based image coder. Both algorithms use the same state-of-the-art zerotree coding for consistency and fair comparison. Several parameters such as filter length and objective cost function are varied to find the best suited filter banks. It is found that for raw data, filter bank with long overlapping filters should be used for processing signals along the time direction whereas filter bank with short filters should be used for processing signal along the distance direction. This combination yields the best results.

Keywords: seismic, compression, GenLOT, wavelet, zerotree coding

1. INTRODUCTION

The amount of data collected in a modern seismic survey may exceed Terabytes, depending on the high resolution representation, dense sensor arrays and long time observation. Management of these large datasets, despite recent increases in mass storage capacity, still presents challenging problems, especially in transmission, but also for storage, processing and interpretation.

Wavelet coding methods has been shown effective for compressing natural image as well as seismic data. They have generated a lot of exciting development, including software and hardware implementation for real-time field transmission.

Recently, works of several authors on GenLOT and its application to natural image coding demonstrate its advantage over wavelet-based coding. Since GenLOTs provide a better frequency partitioning scheme than wavelets, and are applied locally on non-stationary seismic signals, they lead to better performance in terms of SNR and visual interpretation (less ringing artifacts) than state-of-the-art wavelet coders.

In this paper, we compare the performance of GenLOT-based and wavelet-based image coder on seismic data. Both compression algorithms use the same zerotree coding for consistency and fair comparison, on several sets of seismic data including land, marine and synthetic data. Complete comparison on the algorithm performance from acquisition to interpretation, and from raw data sets to stacked sections are proposed. We compress the data using several compression ratios from 10 : 1 to 150 : 1. Preliminary results show that GenLOT-based coding offers over 2 to 5 dB improvements in distortion for raw land datasets. Impact of compression on seismic compression is tested on natural and synthetic data. We show that GenLOT has better performance than wavelet, both objectively and subjectively. Moreover, the proposed GenLOT coder allows parallel processing, and incorporates a quality control feature for progressive transmission.

Further author information:
E-mail: laurent.duval@ifp.fr, nguyent@bu.edu

2. WAVELET AND GENLOT IMAGE CODING

Traditional lossy image coders consist of three stages:

- a transform or subband decomposition (orthogonal or not) which partially decorrelates the signal,
- a quantization which removes hopefully useless details (in some visual/geophysical sense),
- and entropy coding removing the remaining redundancy.

This section quickly reviews the filter banks theory, including wavelets and GenLOTs. We refer to Strang and Nguyen[1] for a comprehensive survey on wavelets and filter banks. We also describe quantization and encoding schemes, with an emphasis on the *zerotree coding* (ZC) framework which combines the two later stages in an embedded quantization/encoding. This technique allows:

- exact bit rate compression, to match a fixed bandwidth transmission exactly, such as for seismic vessel to satellite,
- straightforward quality checks (QC) by transmitting only a fraction of the compressed file.[2]

In this study, we work with 2-D separable transforms on 2-D seismic images. We therefore concentrate on 1-D filter banks for wavelets and GenLOTs as well. One reason is that seismic data is often modeled as 2-D separable processes, as demonstrated by Røsten *et al.*[3] It is shown useful to use filter banks with different properties on the non-isotropic seismic data (they possess different statistical properties in the time and in the space direction).

2.1. Review of previous works

Subband coding is widely used for seismic data compression. Among these methods are the wavelet transform, as used by Luo and Schuster[4] and Donoho *et al.*[5,6] Wavelet compression has been shown to be very effective and has led to actual field transmission (Stigant *et al.*[7]). Local cosine transform (Vermeer *et al.*[8]) has also been used: its overlapping windows are well suited to capture the oscillatory nature of seismic signals. Most general filter banks present promising alternative, as in Røsten *et al.*[9] We design here filter banks based on the Generalized Lapped Orthogonal Transform.

2.2. Wavelet transform

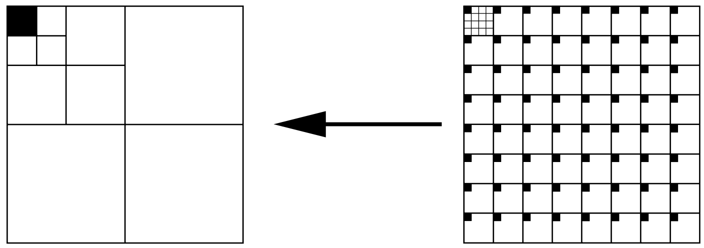
From a filter bank perspective, the wavelet transform is an octave-band decomposition of signals: the dyadic wavelet transform can be obtained by iterating on the low-pass output of a PR 2-channel filter bank.

There are two families of wavelets, orthogonal and biorthogonal. The later is the most often used in data compression: it can have compact support with symmetric and finite filter impulse response, leading to linear phase filters. Orthogonal wavelets (except Haar's) are non-linear phase, which leads to annoying artifacts after decompression. We use in this study short, medium and long wavelet filter banks (FB), respectively with 5-3, 8-4 and 9-7 taps for the low-pass and the high-pass analysis filters. The later is the one used in the FBI fingerprint compression algorithm. It is one of the most used FB for wavelet data compression so far, for natural image[10] and seismic data[11] as well.

2.3. From M-band filter banks to GenLOT

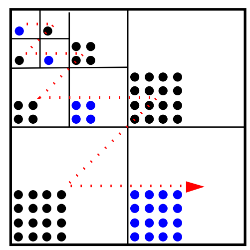
The wavelet transform can achieve a partition of the signal's spectrum in M dyadic bands. However, for medium or high-frequency signals, M-channel uniform filter banks often offer better results. They also provide more degrees of freedom for proper filter design.

Image coding already uses M-channel filter banks: the JPEG standard uses a 8×8 discrete cosine transform (DCT), which is a 8-channel 8-tap paraunitary filter bank. Artifacts such as *blocking effect* arise in JPEG compression at higher bit rates. Moreover, independent compression of 8×8 blocks results in a loss of compression (there still exists correlation between blocks). Lapped Orthogonal Transform (LOT) is an extension of the DCT. It is a M-channel (M even) paraunitary filter bank with $2M$ taps, *i. e.* with overlapping on the two neighboring M pixel blocks. LOT partly smoothes out the block boundaries and reduces blocking artifacts.



Dyadic and regular partition

Figure 1. Dyadic wavelet (left) and regular GenLOT (right) frequency partition.



Embedded coding

Figure 2. Parent-descendent relationship in the dyadic segmentation

But longer overlaps are needed for further reduction of blocking effect, motivating the development of GenLOT by De Queiroz *et al.*[12] Assuming the number of channels M is even and all the filter length L is a multiple of M ($L = NM$), all linear phase paraunitary FBs can be factorized into modular building blocks, where DCT and LOT are special cases (with $N = 1$ or $N = 2$ respectively). The input signal is divided into sequences of L samples, and overlapped sequences have $M(N-1)$ samples in common. GenLOT has regular frequency partition (uniform-band decomposition), whereas wavelet's is dyadic, as seen in Fig. 1. The DC subband still presents correlation. It is further decorrelated by a dyadic wavelet transform.

2.4. Quantization and encoding techniques

Traditional coders perform a thresholding/quantization on the transformed coefficients, followed by entropy coding (Run Length Encoding, Huffman or arithmetic coding). The later stage reduces the remaining coefficients' entropy.

Shapiro[13] has defined an novel approach for progressive coding of images. It relies on the idea that the most important information should be transmitted first. With mean squared error (MSE) as a distortion measure and a paraunitary transform, it can been shown that when the transform coefficients $c_{i,j}$ are sent one by one, the MSE decreases by $c_{i,j}^2/N_{pixels}$ (see *e.g.* Said and Pearlman[10]). The most important information is the largest coefficients, which can still be transmitted in bit plane order, in a refinable process, from most to least valued bit weight. Further redundancy removal between subbands is obtained by a hierarchical tree structure. This structure parses across subband coefficients having the same spatial location. Natural and seismic image possess the feature that, after a proper transform stage, null coefficients are likely to have zero or small descendents (coefficients below one node called parent, sharing the same spatial location). Figure 2 shows the parent-descendent relationship in the classical dyadic segmentation of the transformed image. We refer to Said and Pearlman[10] for practical implementation. Tran and Nguyen[14] have shown that a block transform can be rearranged in the same tree-like structure, which is demonstrated in fig. 1. The same zerotree coding algorithm also applies to GenLOT. The performance of zerotree coding then relies on the transform's ability to decorrelate coefficients.

2.5. Filter bank design

Recent works in image processing have shown than GenLOT with proper design outperform wavelet compression for conventional images.[14] Several criteria are used for transformation optimization: for instance, *Coding gain* (CG) optimization usually correlates with higher SNRs (objective measure). Other objective measures include *DC leakage* (DC). Though not essential, it often improves the visual quality of the reconstructed data (subjective measure). The article[14] discusses issues on filter bank optimization, and comparison to wavelet coders. Optimization for the filters used in this study are to be found in Table 1.

Filter	Channels (M)	Length (N)	Optimization	Index
DCT	8	8	none	1
LOT	8	16	none	2
LOT16	16	32	none	3
LOT42cgmax	4	8	cgmax	4
LOT84cg	8	32	cg	5
LOT84cgdc	8	32	cgdc	6
LOT85cg	8	40	cg	7
LOT86	8	48	none	8
LOT86cgdc	8	48	cgdc	9
LOT86cgmax	8	48	cgmax	10
LOT86fr	8	48	fr	11
LOT86frmax	8	48	frmax	12
LOT89	8	72	none	13
ULLOT	8	16	none	14
ULLOT834cgfr	8	24	cgfr	15

Table 1. Parameters for the chosen GenLOT filter banks

Symbol	Explanation
M	number of channels
cgdc	maximal coding gain and no DC leakage
cgmax	maximal coding gain
dcfr	no DC leakage and maximal frequency attenuation
frmax	maximal frequency attenuation
none	no DC leakage

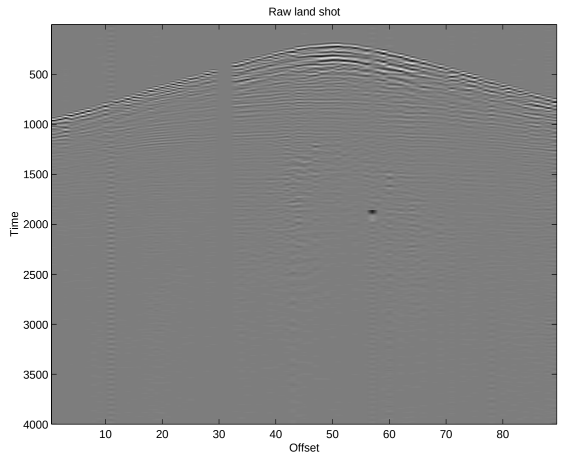
Table 2. Explanation for symbols in Table 1

3. RESULTS AND COMPARISONS

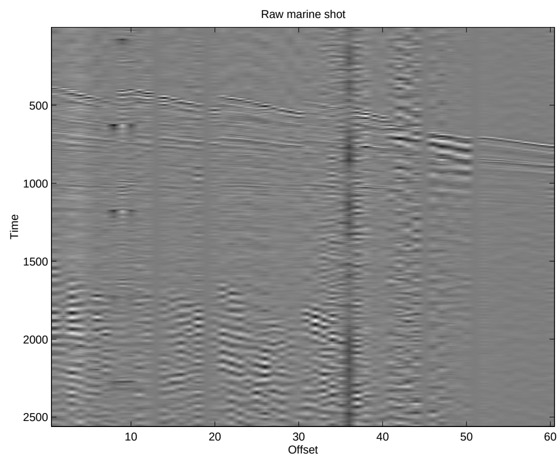
We demonstrate the effectiveness of GenLOT-based compression on two sets of raw seismic data, and one stacked section, as shown in Fig. 3.

It is well known that traditional error measures such as signal/noise ratio (SNR) do not really measure the data quality. It nevertheless can be used for comparison between two compression methods at the same compression ratio. Following Vassiliou and Wickerhauser,[11] we compute the following two types of signal/noise ratios, the conventional SNR and the absolute SNR (ASNR):

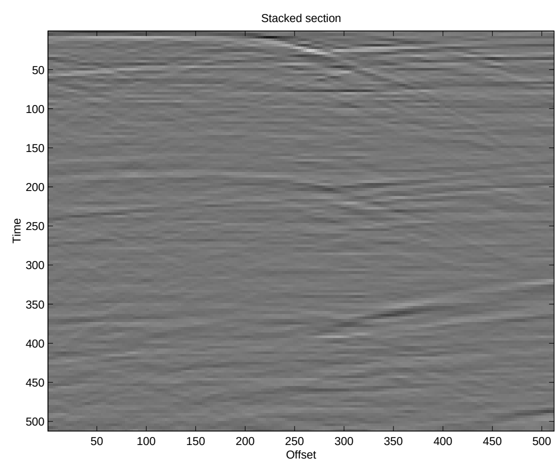
$$\begin{aligned} \text{SNR} &= 10\,log_{10}\left(\textstyle\sum_n s_n^2 / \sum_n \Delta s_n^2\right) \\ \text{ASNR} &= 20\,log_{10}\left(\textstyle\sum_n |s_n| / \sum_n |\Delta s_n|\right). \end{aligned}$$



Land data



Marine data



Stacked section

Figure 3. Seismic datasets used in this study

3.1. Results

We first performed a comparison between biorthogonal wavelets for both error measures for the three datasets (Fig. 4 and 5). These figures confirm previously reported results[11] about the relatively better performance of the 9-7 wavelet compared to other wavelets. We therefore use the 9-7 performance as a benchmark for the performance of the GenLOT filters. One should note that the same trends occur in both conventional SNR and ASNR plots. We therefore use the sole SNR for the remaining comparisons. Finally, wavelet behaviour seems to be asymptotically the same for raw data, land and marine, as the compression ratio increases.

The performance of the 15×15 GenLOT pairs (15 in both the x and y-direction) is evaluated on the land data at compression ratio 10 : 1. This performance is indicated in Fig. 6, where good SNRs are denoted in white. We choose the three brighter horizontal stripes, with indexes 1, 14, 4 (namely DCT, ULLOT and LOT42cgmax) and 7, 13, 11 (LOT85cg, LOT89 and LOT86fr) in the vertical direction. One can see that we obtain better results (for raw data sets) with relatively short bases in the horizontal direction, indicating less correlated information. Our experiments also confirm that longer, overlapping bases in the vertical direction give higher SNRs.

For the land raw data set, the best filters over all compression ratios are:

Horizontal direction	Vertical direction
DCT	LOT85cg
DCT	LOT89
LOT42cgmax	LOT85cg

Table 3. Best filters for the land raw data set

For marine raw data set, the best filters over all the compression ratios are:

Horizontal direction	Vertical direction
DCT	LOT85cg
ULLOT	LOT85cg
DCT	LOT89

Table 4. Best filters for the marine raw data set

Plots from Fig. 7 demonstrate the importance of choosing different filter banks for vertical and horizontal directions, especially for the land data. We can therefore select the DCT which performs better in both cases. We obtain in average between 2 and 5 dB with the best filter bank, comparing to the 9-7 wavelet.

Experiments on the stacked section (Fig. 8) show that short filters only achieve good results at lower compression ratios. Decomposing data with the same longer filter banks (40 or 48 taps) gives the best results, 3 to 5 dB above the wavelet distorsion curve. The difference stack plots (difference between the original and the compressed data) in Fig. 9 show that, at higher compression ratios, coherent seismic structures appear with wavelet based compression, while the GenLOT-based plot exhibits mostly uncoherent noise, and thus does not harm final seismic interpretation. More detailled results of compression on synthetic data are exposed in Duval *et al.*[15]

4. CONCLUSIONS

We propose a compression algorithm for seismic data, for raw seismic shots or processed stack sections as well. This algorithm is based on GenLOT, a class of filter banks with a regular frequency partition, well suited to seismic data. The use of well designed filter banks, matching the properties of the data leads to 2 to 5 dB improvement over wavelet filter banks. Future studies will encompass more specific designs for the GenLOT filter banks.

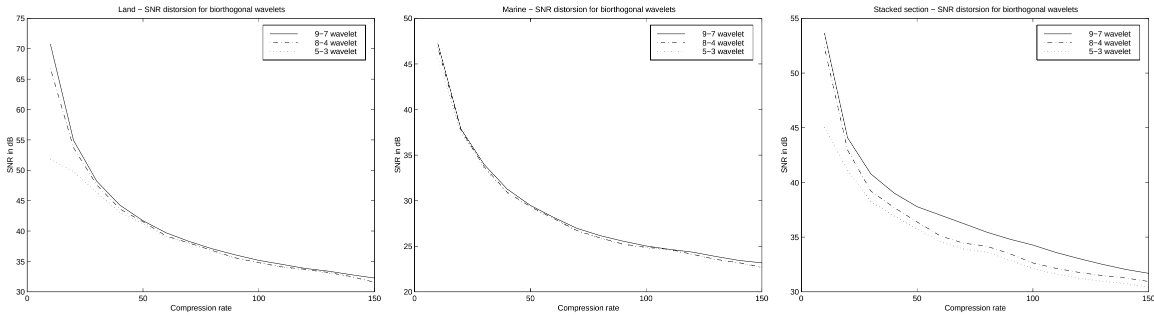


Figure 4. Wavelet compression SNR comparison for the three datasets

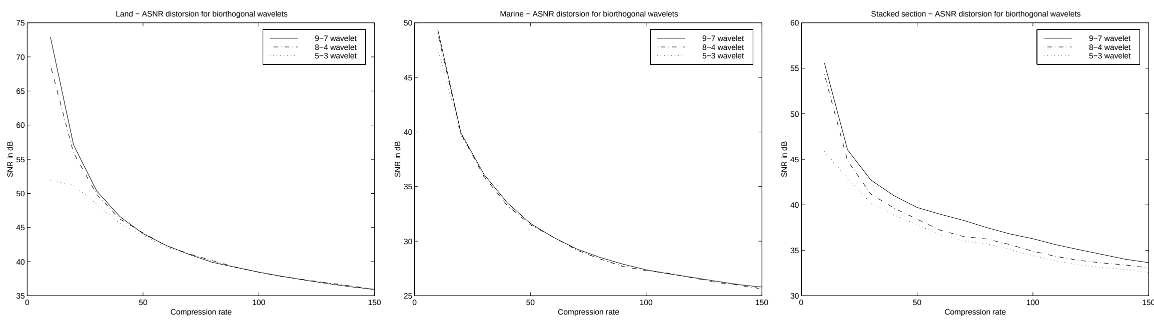


Figure 5. Wavelet compression ASNR comparison for the three datasets

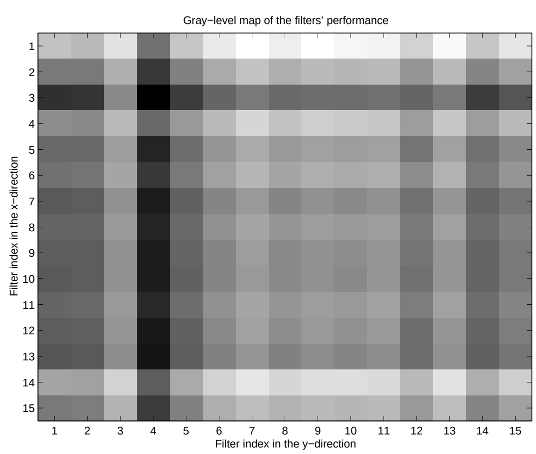


Figure 6. Performance of GenLOT filters at 10:1 compression ratio

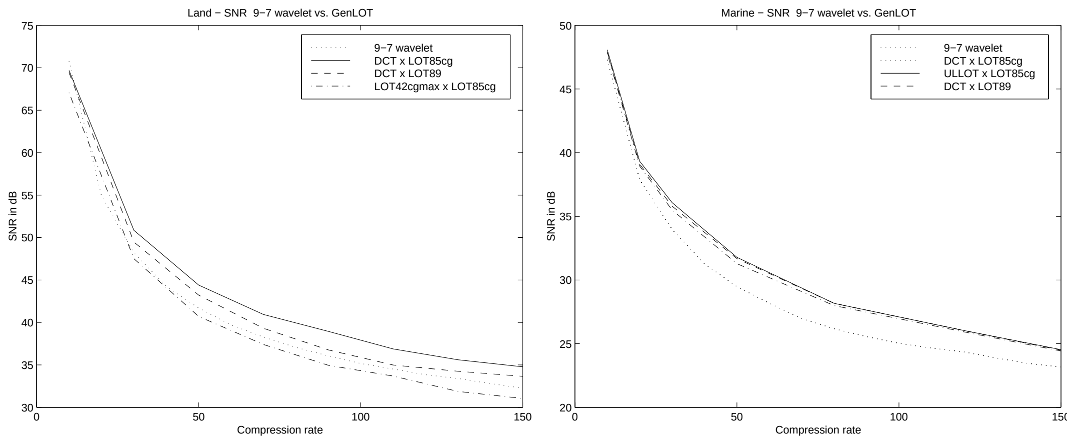


Figure 7. Compression of land and marine sets

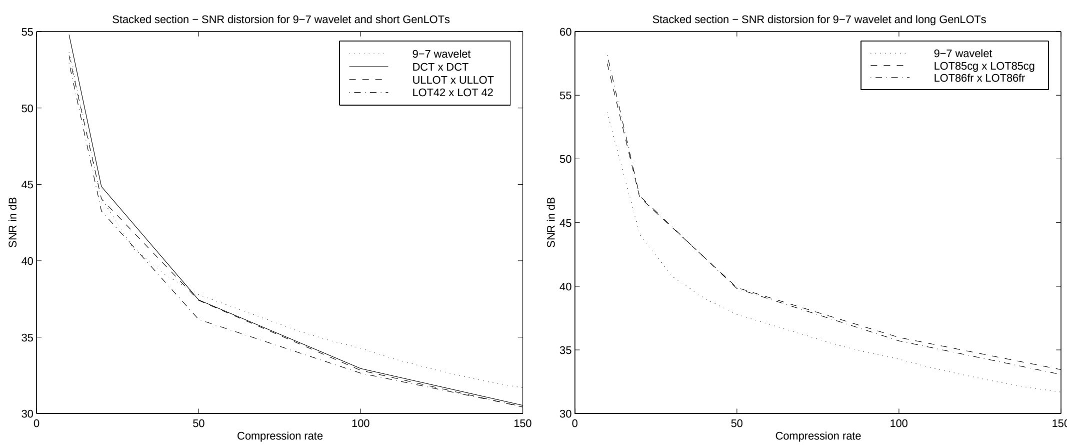
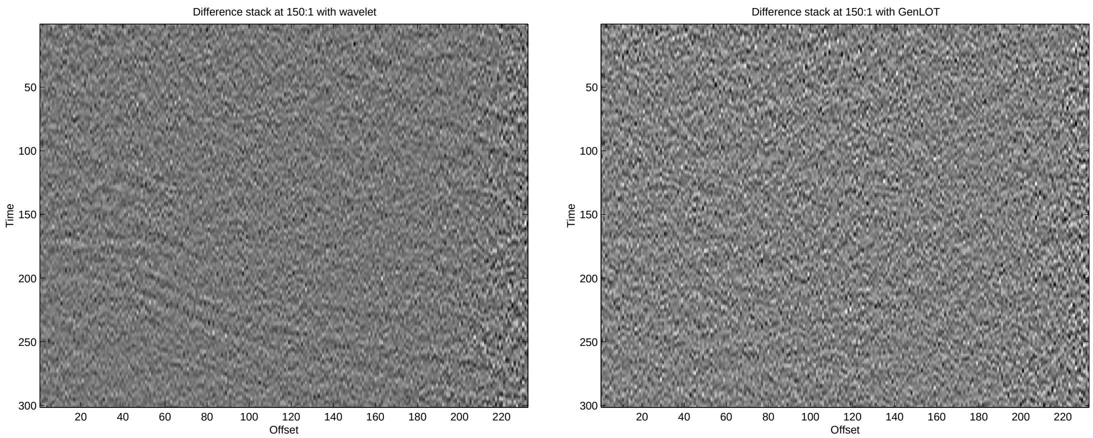


Figure 8. Stacked section: compression with short and long GenLOTs



with 9-7 wavelet   with GenLOT (LOT85cg $\times$ LOT85cg)

Figure 9. Difference stack section after 150 : 1 compression

REFERENCES

1. G. Strang and T. Nguyen, *Wavelets and Filter Banks*, Wellesley-Cambridge Press, Wellesley, MA, 1996.
2. L. C. Duval, T. Q. Nguyen, and T. D. Tran, "Seismic data compression and QC using GenLOT," in *Proc. 61th EAGE Conference*, Jun. 1999.
3. T. Røsten, V. A. Marthinussen, T. A. Ramstad, and A. Perkis, "Filter bank optimization for high-dimensional compression of pre-stack seismic data," in *Proc. Int. Conf. on Acoustics, Speech, and Signal*, March 1999.
4. Y. Luo and G. T. Schuster, "Wave packet transform and data compression," in *Proc. 62nd Annual SEG Int. Meeting*, pp. 1187–1190, 1992.
5. P. L. Donoho, R. A. Ergas, R. S. Polzer, and J. D. Villasenor, "Compression optimization by multidimensional wavelet transforms and data dependent quantization," in *Proc. 66th Annual SEG Int. Meeting*, vol. 2, p. 2042, 1996.
6. P. L. Donoho, R. A. Ergas, and R. S. Polzer, "Improved data management for interpretation systems using compression," in *Proc. 60th EAGE Conference*, 1998.
7. 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*, pp. 960–962, Oct. 1995.
8. P. Vermeer, H. Bragstad, and C. Orr, "Aspects of seismic data compression," in *Proc. 66th Annual SEG Int. Meeting*, pp. 2031–2034, Nov. 1996.
9. T. Røsten, P. Waldemar, I. Balasingham, and T. A. Ramstad, "A review of subband coding of seismic data," in *Proc. Norwegian Signal Processing Symposium and Workshop (NORSIG/IEEE)*, pp. 69–74, May 1997.
10. 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**, pp. 243–250, Jun. 1996.
11. A. Vassiliou and M. V. Wickerhauser, "Comparison of wavelet image coding schemes for seismic data compression," in *Proc. 67th Annual SEG Int. Meeting*, vol. 2, pp. 1334–1337, 1997.
12. 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**, pp. 497–507, Mar. 1996.
13. J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. on Signal Processing* **41**, pp. 3445–3462, Dec. 1993.
14. 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.
15. L. C. Duval, J. Oksman, and T. Q. Nguyen, "A new class of filter banks for seismic data compression," in *Proc. 69th Annual SEG Int. Meeting*, 1999. Submitted.