SEISMIC DATA COMPRESSION USING GULLOTS

Laurent C. Duval

Technology Department
Institut Français du Pétrole
92852, Rueil-Malmaison Cedex, France
E-mail: laurent.duval@ifp.fr

Takayuki Nagai

Course in Electronic Engineering
The University of Electro-Communications
Tokyo, 182–8585 Japan
E-mail: tnagai@ee.uec.ac.jp

ABSTRACT

Recent works have shown that GenLOT coding is a very effective technique for compressing seismic data. The role of a transform in a coder is to concentrate information and reduce statistical redundancy. When used with embedded zerotree coding, GenLOTs often provide superior performance to traditional block oriented algorithms or to wavelets. In this work we investigate the use of Generalized Unequal Length Lapped Orthogonal Transforms (GULLOT). Their shorter bases for high-frequency components are suitable for reducing ringing artifacts in images. While GULLOTs yield comparable performance to GenLOTs on smooth seismic signals like stacked sections, they achieve improved performance on less smooth signals such as shot gathers.

1. INTRODUCTION

Modern seismic marine acquisition surveys may produce 10 to 100 Terabytes of data [1]. Seismic data compression is thus becoming desirable for a variety of geophysical applications, ranging from transmission to storage.

Wavelet coding has long been a favorite technique for seismic data compression [2, 3]. It has led to a real-time field trial satellite transmission in the North Sea in 1995. Stigant *et al.* [2] showed that, for a large 3-D survey, wavelet compression was capable of $60 : 1$ compression ratio (CR) or more without visible degradation in the data. Methods involving local cosine bases [4, 5], non-unitary filter banks [6] or GenLOTs [7] have also been developed more recently. In the later cases, transforms are specifically adapted to the properties of the data, in contrast to the generic use of a wavelet like the 9/7 biorthogonal wavelet.

In the case of filter banks, adaptation to the data can be provided by filter bank optimization driven by objective criteria, such as coding gain or stopband attenuation (see [8] or Section 2.2).

Some of these previous works focus on filter banks with equal length filters. Long overlapping low-frequency filters are desirable to reduce blocking artifacts, which are caused by non overlapping transforms such as the DCT. GenLOT considerably reduces the blocking artifacts, and produces less ringing artifacts than wavelets, as demonstrated by T. Tran *et al.* [9]. Ringing artifacts are nevertheless still present in images, especially around strong edges. A solution to solve this problem is to reduce the length of the high-pass filters, while keeping the same length for the low-pass filters. T. Nagai *et al.* [10, 11] have recently proposed the GULLOT, which produces structurally longer and shorter filters, along with an efficient fast algorithm.

In this work we investigate the use of GULLOTs for seismic data compression. Better coding results for seismic data can be expected using GULLOTs, since they possess a good frequency partitioning like GenLOTs, and a wavelet-like variable basis length property at the same time. While GULLOTs yield comparable performance to GenLOTs on smooth signals like stacked sections, they achieve improved performance on less smooth seismic signals such as shot gathers.

The paper is organized as follows: Section 2 recalls the structure and design of the GULLOTS. We briefly describe the data set used for performance comparison in Section 3. Section 4 details the properties of the filter banks used in this study. Objective SNR results are displayed in Section 5, while Section 6 displays an example of ringing effects on a shot gather. We conclude that GULLOTs are computationally efficient filter banks for seismic data compression.

2. STRUCTURE AND DESIGN OF GULLOTS

2.1. GULLOTs

We briefly describe GULLOTs in this subsection. The reader is referred to [10, 11] for more details.

Now, let $L_i = N_i M$ be the length of i-th basis vector, where M represents the number of channels. GULLOTs are classified into the following four types:

- Type A: All N_i's are odd numbers.
- Type B: All N_i's are even numbers.
- Type C: N_i's consist of both odd and even numbers, and $\max(N_i)$ is odd.
- Type D: N_i's consist of both odd and even numbers, and $\max(N_i)$ is even.

It should be noted that types A and B GULLOT vectors share a common center of symmetry, whereas types C and D have different centers of symmetry. All types of GULLOTs can be factorized as

$$\boldsymbol{E}(z) = \boldsymbol{R}_M^T \hat{\boldsymbol{\Lambda}}(z) \boldsymbol{K}_{N-1}(z) \boldsymbol{K}_{N-2}(z) \cdots \boldsymbol{K}_1(z) \boldsymbol{R}_M \boldsymbol{C}_M^{II} \boldsymbol{J}_M, \quad (1)$$

where

$$\boldsymbol{K}_i(z) = \boldsymbol{\Phi}_i \boldsymbol{A}_{\alpha_i} \boldsymbol{W}_{\alpha_i} \boldsymbol{\Lambda}(z) \boldsymbol{W}_{\alpha_i},$$

$$\boldsymbol{\Phi}_i = \begin{bmatrix} \boldsymbol{U}_i & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_{\alpha_i} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{V}_i & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{I}_{\alpha_i} \end{bmatrix},$$

$$
\boldsymbol{W}_\alpha = \begin{bmatrix} \boldsymbol{I}_{M/2-\alpha} & \mathbf{0} & \boldsymbol{I}_{M/2-\alpha} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{I}_\alpha & \mathbf{0} & \mathbf{0} \\ \boldsymbol{I}_{M/2-\alpha} & \mathbf{0} & -\boldsymbol{I}_{M/2-\alpha} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{I}_\alpha \end{bmatrix},
$$

$$
\boldsymbol{\Lambda}(z) = \begin{bmatrix} \boldsymbol{I}_{M/2} & \mathbf{0} \\ \mathbf{0} & z^{-1}\boldsymbol{I}_{M/2} \end{bmatrix},
$$

$$
\boldsymbol{A}_\alpha = diag[\underbrace{1/2 \ \cdots \ 1/2}_{M/2-\alpha} \ \underbrace{1 \ \cdots \ 1}_{\alpha} \ \underbrace{1/2 \ \cdots \ 1/2}_{M/2-\alpha} \ \underbrace{1 \ \cdots \ 1}_{\alpha}].
$$

The matrices $\boldsymbol{U}_i$ and $\boldsymbol{V}_i$ can be any real orthogonal matrices of size $(M/2-\alpha_i) \times (M/2-\alpha_i)$. In the above equation, α_i represents

$$
\alpha_i = \frac{M - \mathcal{N}(\boldsymbol{m}_i)}{2}, \quad 1 \le i \le N-1 \tag{2}
$$

where $\mathcal{N}(\boldsymbol{a})$ denotes the number of non-negative (zero or positive) elements of $\boldsymbol{a}$ and

$$
\boldsymbol{m}_i = [N_0 - i - 1 \ \ N_1 - i - 1 \ \ \cdots \ \ N_{M-1} - i - 1].
$$

The matrix $\hat{\boldsymbol{\Lambda}}(z)$ is a diagonal matrix, which aligns the center of all filters. $\boldsymbol{C}_M^{II}$ stands for the M-point DCT-II matrix.

2.2. Design strategy

There are some criteria for the design of GenLOTs. In general, following coding gain is used for image coding application:

$$
G_{TC} = \frac{\sum_{k=0}^{M-1} \sigma_{y_k}^2}{M\left(\prod_{k=0}^{M-1} \sigma_{y_k}^2\right)^{\frac{1}{M}}}, \tag{3}
$$

where $\sigma_{y_k}^2$ denotes the variance of the k-th subband signal. However, this direct optimization will never give a good solution for GULLOTs. Hence, the stopband attenuation of the filters $H_k(z)$ is maximized by minimizing the following cost function,

$$
J = \sum_{k=0}^{M-1} \int_{stopband} |H_k(e^{j\omega})|^2 d\omega. \tag{4}
$$

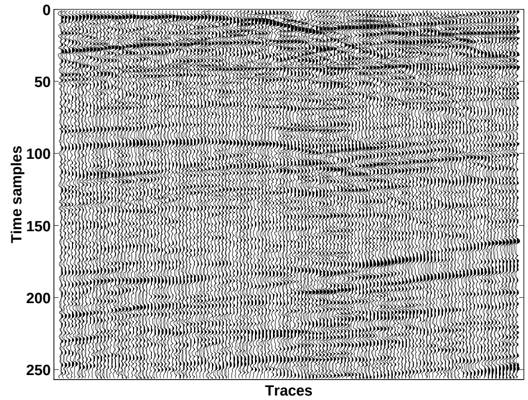
The solution of the above optimization is used as the initial value for maximizing the coding gain. Taking account of the fast implementation, $\boldsymbol{U}_i = \boldsymbol{I}$ is chosen. If the first block is the DCT, the choice also guarantees the GULLOT to have no DC leakage.

3. DATA SET DESCRIPTION

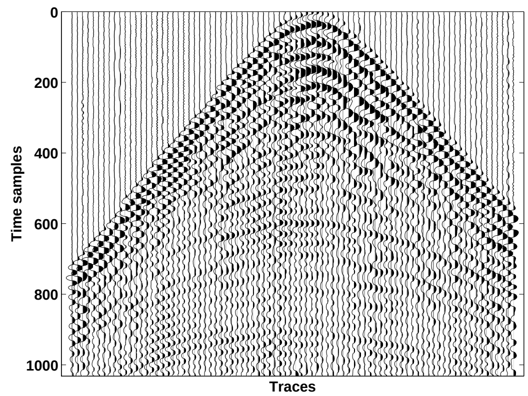
The data set used in this preliminary study consists of a stacked section from the Forêt d'Orléans – France (Fig. 1 (a)) and a raw land shot gather from Louisiana (Fig. 1 (b)). We refer to [12] for a comprehensive survey of seismic processing and the various types of seismic data. The stacked section is obtained after seismic processing. It is a relatively smooth image with structured lateral information representing ground reflectors. The shot gather from Fig. 1 (b) possesses more local amplitude variations. This data could be viewed as an image with strong edges.

4. FILTER BANK AND CODER DESCRIPTION

A set of seven 8-channel filter banks (FB) has been used in this study. It contains the DCT-II, 2 pairs of GenLOTs and one pair of GULLOTs. GENLOTs have two different overlapping factors:



(a) Stacked section.



(b) Raw land shot gather.

Fig. 1. Examples of seismic data.

5 for GenLOT5 (8 × 5 coefficients per filter) and 6 for GenLOT6 (8 × 6 coefficients). The GULLOT46 is a type B GULLOT. It is hence fully symmetric. It has 8×6 coefficients and 8×4 coefficients for the first four low-pass filters and the last four high-pass filters respectively. The choice $\boldsymbol{U}_i = \boldsymbol{I}$ was used for every GenLOT or GULLOT. For fair comparison, all the FBs except the DCT are optimized following exactly the same procedure described in Sec. 2.2. Each FB pair is composed of two FBs:

- one using the reduced angle set, denoted by "r", for which the matrix $\boldsymbol{V}_i$ is characterized by a cascade of $M/2-1$ plane rotations for fast implementation;
- one using the full angle set, denoted by "f", with a cascade of $(M-2)M/8$ plane rotations for the matrix $\boldsymbol{V}_i$.

The FB properties are summarized in Table 1. Table 1 gathers the shortest and the longest filter length of each FB along with the total number of coefficients. The column named "Flops" represents the number of floating point operations required when using the lattice structure in transform calculations. Two numbers are displayed. They correspond to the reduced and the full angle set case, respectively.

We observe that GULLOT46 possesses the same number of coefficients than GenLOT5 and less than GenLOT6. GULLOT46 requires even less floating point operations than both GenLOT5 and GenLOT6 when used in a lattice form. GULLOTs are there-

Name	Short	Long	Coef.	Flops
DCT	8	8	64	42
GenLOT5	40	40	320	178–250
GenLOT6	48	48	384	212–302
GULLOT46	32	48	320	172–226

Table 1. Filter banks properties.

fore altogether more computationally efficient than the GenLOTs used in this study.

The coder used here is described in [13]. It performs an embedded zerotree decomposition on symmetric power of 2-channel separable FBs. The stacked section is compressed with the same FB in the horizontal and vertical direction. The FB pair is chosen amongst the seven 8-channel FBs. In the case of the shot gather, the horizontal FB is fixed to the DCT, because of the poor correlation in the horizontal direction. The vertical FB is varied throughout the seven selected FBs.

5. COMPRESSION RESULTS

We compare the filter banks performance using the traditional SNR measure, expressed in dB. Let s_n and s'_n be original and compressed/decompressed samples, respectively. SNR is defined as follows:

$$\mathrm{SNR_{dB}} \quad = \quad 10\, log_{10} \left(\sum_n s_n^2 / \sum_n (s_n - s'_n)^2 \right).$$

Figure 2 (a–b) represents the rate/distortion curves for the two data sets. It reports the difference between the actual SNR and the SNR obtained with a DCT filter pair. All the curves lie at least 1 dB above the DCT, which confirms the validity of a lapped transform approach to seismic data compression [7].

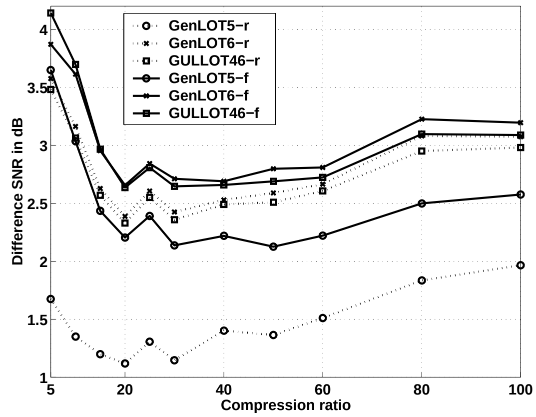
Reduced angle set transforms (denoted by "r") are represented by doted lines, while full angle set transforms (denoted by "f") are represented by solid lines. We observe that f–transforms always outperform their r– counterpart, at the cost of increased computations. In the case of the stacked section (Fig. 2 (a)), GULLOT46–r and –f both outperform GenLOT5–f for CR ≥ 10, with the same number of coefficients and less lattice complexity.

GenLOT6–f and GULLOT46–f obtain the best results, with more than 2.5 dB improvement over the DCT. The two later transforms stay within a range of 0.2 dB to each other. For one interested in more efficiency, the reduced angle set GULLOT46–r generally performs 0.1 dB less than GenLOT6–r.

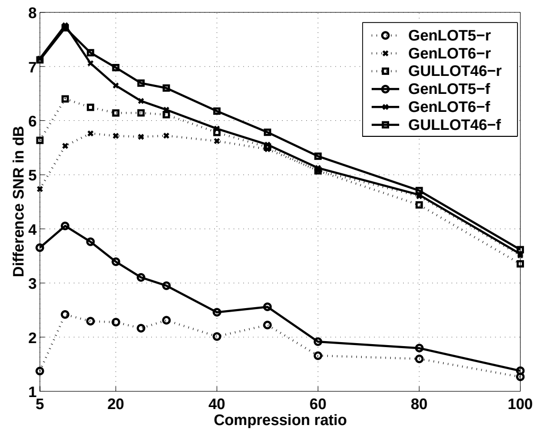
In the case of a less smooth image like a shot gather, one can see from Figure 2 (b) that GULLOT–f now outperforms the other FBs at almost all compression ratios, with up to 0.5 dB over GenLOT6–f. The GenLOT6 and the GULLOT46 pairs have very close performance as the CR increases. Amongst less complex FBs, GULLOT46–r outperforms GenLOT6–r by 0.5 to 1 dB for CR ≤ 30 and GenLOT6–r gradually becomes better above a compression ratio of 60.

6. RINGING ARTIFACTS

Since seismic data are highly oscillatory, ringing artifacts are particularly difficult to spot on seismic images. Moreover, the choice of the non-overlapping DCT in one direction also limits the horizontal extent of the ringing. Figure 3 (a) represents a portion of



(a) Rate/distorsion for the stacked section.



(b) Rate/distorsion for the shot gather.

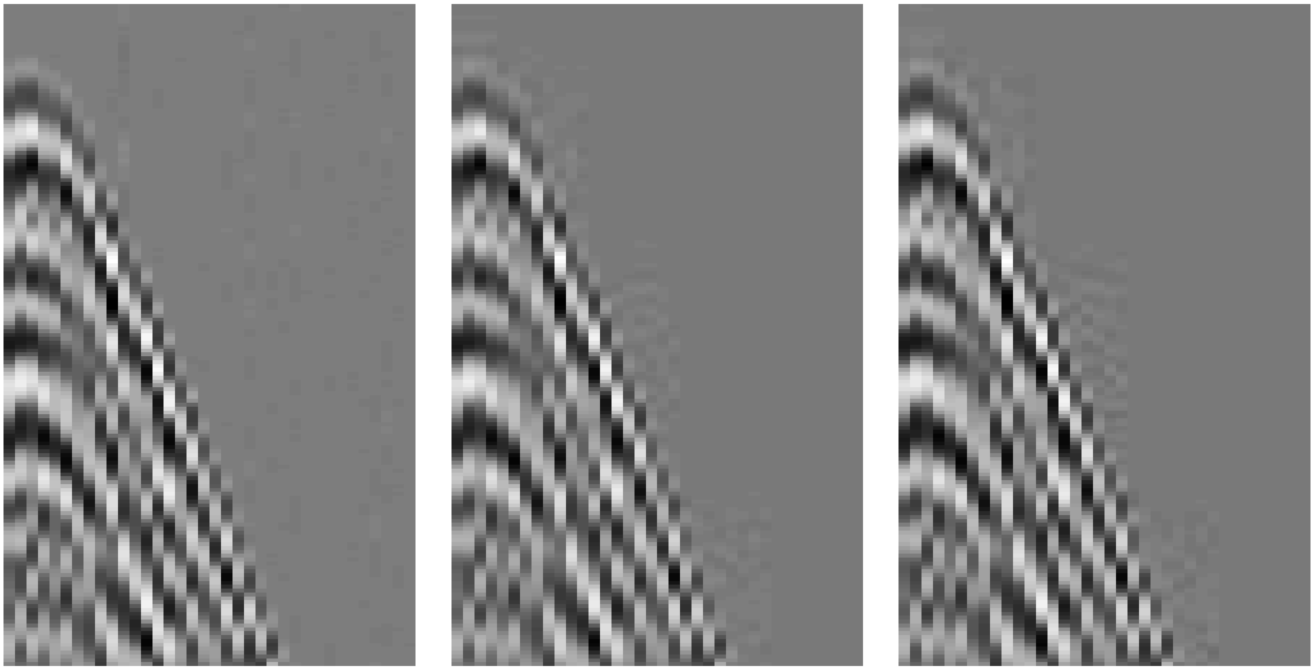
Fig. 2. Compression results against DCT at several compression ratios for a stacked section and a shot gather.

the shot gather from Fig. 1 (b), around the first break waves. This portion could be seen as an equivalent of a strong edge in a natural image. We choose a compression ratio of $100 : 1$. Since ringing artifacts are barely visible even at high compression ratios, small amplitudes have been enhanced to allow ringing evaluation.

Due to shortest high-frequency filters, ringing artifacts tend to spread less for GULLOT46–f than for GenLOT6–f (compare Fig. 3 (b) and (c)). GULLOT46–f achieves consequently a slightly better SNR than GenLOT6–f at $100 : 1$, with much lower complexity. However, GULLOT46–f still exhibits ringing artifacts. As a comment, shortening the four high-frequency filters does not suffice at high compression ratios. Since GULLOTs are available with several other filter length combinations, seismic data compression performance might be further improved by the use of more suitable filter lengths. This issue is left for future research.

7. CONCLUSIONS

We demonstrated the use of Generalized Unequal Length Lapped Orthogonal Transforms (GULLOT) for seismic data compression. Structurally enforced shorter high-pass filters lower the computational burden. GULLOTs generally outperform GenLOTs with



(a) Original shot gather. (b) Compressed w/ GULB46–f. (c) Compressed w/ GenLOT6–f.

Fig. 3. Ringing artifact example on a portion of a shot gather at $100:1$ compression ratio.

comparable complexity. For non-smooth seismic images as produced by shot gathers, GULLOTs may even outperform GenLOTs with higher complexity.

8. REFERENCES

[1] Paul L. Donoho, "Seismic data compression: Improved data management for acquisition, transmission, storage, and processing," in *Proc. Seismic'98*, 1998, pp. 1–7.

[2] 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 *Annual International Meeting*. Oct. 1995, pp. 960–962, Soc. of Expl. Geophysicists, Exp. abstracts.

[3] F. Khène and S. Abdub-Jauwad, "Efficient seismic compression using the lifting scheme," in *Annual International Meeting*. Aug. 2000, pp. 2052–2054, Soc. of Expl. Geophysicists, Exp. abstracts.

[4] Y. Wang and R.-S. Wu, "Improvements on seismic data compression and migration using compressed data with the flexible segmentation scheme for local cosine transform," in *Annual International Meeting*. 2000, pp. 2048–2051, Soc. of Expl. Geophysicists, Exp. abstracts.

[5] F. G. Meyer, "Fast compression of seismic data with local trigonometric bases," in *Wavelet Applications in Signal and Image Processing VII*, M. Unser, A. Aldroubi, and A. F. Laine, Eds. July 1999, vol. 3813, pp. 648–658, SPIE.

[6] T. Røsten, T. A. Ramstad, and L. Amundsen, "Part I: Subband coding of common offset gathers," *Submitted to Geophysics*, 2000, Preprint.

[7] L. C. Duval, V. Bui-Tran, T. Q. Nguyen, and T. D. Tran, "GenLOT optimization techniques for seismic data compression," in *Int. Conf. on Acoust., Speech and Sig. Proc.*, 2000, pp. 2111–2114.

[8] L. C. Duval and T. Røsten, "Filter bank decomposition of seismic data with application to compression and denoising," in *Annual International Meeting*. 2000, pp. 2055–2058, Soc. of Expl. Geophysicists, Exp. abstracts.

[9] 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 Proc.*, vol. 8, pp. 1493–1507, Nov. 1999.

[10] T. Nagai, M. Ikehara, M. Kaneko, and A. Kurematsu, "Generalized Unequal Length Lapped Orthogonal Transform for subband image coding," in *Int. Conf. on Acoust., Speech and Sig. Proc.*, 2000, pp. 520–523.

[11] T. Nagai, M. Ikehara, M. Kaneko, and A. Kurematsu, "Generalized Unequal Length Lapped Orthogonal Transform for subband image coding," *IEEE Trans. on Signal Proc.*, pp. 3365–3378, Dec. 2000.

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

[13] L. C. Duval and T. Q. Nguyen, "Seismic data compression: a comparative study between GenLOT and wavelet compression," in *Wavelet Applications in Signal and Image Processing VII*, M. Unser, A. Aldroubi, and A. F. Laine, Eds. Jul. 1999, vol. 3813, pp. 802–810, SPIE.