

SIMULTANEOUS SEISMIC COMPRESSION AND DENOISING USING A LAPPED TRANSFORM CODER

Laurent C. Duval

Institut Français du Pétrole
Technology Department
1 et 4 av. de Bois-Préau
F-92500 Rueil-Malmaison, France
laurent.duval@ifp.fr

ABSTRACT

Compression and denoising are two of the most successful applications of wavelets to signals and natural images. Both techniques have also been successfully applied to seismic signals, but compression is not widely accepted yet, since it is often believed to harm seismic information.

Trying to look at compression and denoising in another direction, the present work stresses on the idea that they could be viewed as two sides of the same coin. As a result, in the case of naturally noisy seismic data, compression could be seen as a denoising tool, instead of a mere noise source. We substantiate this statement on a noise-free seismic data model and actual seismic field data.

We show that, depending on the amount of initial ambient noise in the data, a lapped transform coder with embedded zerotree coding may be able to effectively denoise seismic data, over a wide range of compression ratios.

1. INTRODUCTION ON SEISMIC DATA COMPRESSION

Modern large scale 3-D seismic surveys may generate hundreds of Terabytes of data. Seismic data compression has already been experienced to reduce complexity in the management of these ever increasing amounts of data, for instance in vessel-to-base satellite transmission or to reduce memory requirements or storage costs [1, 2].

Unfortunately, seismic compression needs usually exceed 10 to 1 compression ratios (CR). Moreover, higher compression ratios (up to 100 to 1) are highly desirable, but they can only be achieved through lossy compression.

Audio signal or image lossy compression generally rely on psycho-acoustics or psycho-visual knowledge to hide the compression induced noise, so that the actual information loss is barely perceptible.

Such techniques cannot apply simply to seismic data: "raw" seismic data undergoes complex processing steps be-

fore they can be interpreted as a "cross-section" of the underlying ground. Seismic data properties change – sometimes radically – through each of the processing steps, including signal stretching, time of arrival correction, deconvolution, inversion, . . . The ability to extract the maximum possible information from each seismic dataset is often crucial to 3-D seismic processing. As a result, the seismic community is still reluctant to using compression techniques, and no seismic compression standard exists to date.

In order to promote compression, several studies have been carried out to show that, at least at low CRs, the compression noise is almost safe to subsequent processing [1, 2, 3]. One of the most interesting studies has been released in August 2001 to the Seismic Compression Diagnostic Initiative (SCDI), formed in 1997 to address the effects of compression on seismic data interpretation, and to provide (if possible) safe compression diagnostics. Because of the complexity of geophysical requirements, the SCDI Consortium did not reach to a compression standard or universal rules, but instead provided useful guidelines for nearly safe compression [2].

Since seismic data inherently contains ambient noise, and compression (at low CRs) should add nearly random noise to the data, compression might be relatively safe if the compression noise is significantly lower than the ambient noise, and if the noise exhibits the appropriate statistical properties (e. g. gaussianity, whiteness, . . .)

The definition of a safe compression noise threshold for all kind of seismic data was not given, but it was observed to be dependant on the type of data, the level of ambient noise and the processing stage at which one decides to compress the data.

In the present work, we try shed another light on the following canonical point of view, as quoted from [2, p. 28]: "... from the point of view of seismic data processing, compression does nothing useful", which we partially disagree with. The study was motivated by two somewhat related

arguments:

- the idea that noises (ambient ones and from compression) just add might have been slightly overestimated,
- state-of-the-art coders are based on wavelets or filter banks, which are basic tools for denoising, and could thus serve two purposes at a time.

It was also motivated by a preceding joint work with Tage Røsten, in which we investigated the superiority of paraunitary and near-perfect multichannel filter banks on wavelets for seismic data compression and denoising [4].

The paper is organized as follows. In Section 2, we acknowledge previous works discussing the issues of compression and denoising, from the seminal ideas of shrinkage by D. Donoho *et al.* In Section 3 we explain why, with an additive noise assumption, noises from the signal and from compression do not necessarily add to a stronger noise. Sections 4 and 5 state compression/denoising results on model data and actual seismic data respectively. We finally draw conclusions and perspectives in Section 6.

2. ON COMPRESSION, DENOISING AND SHRINKAGE

In the wavelet realm, compression and denoising might be regarded as two sides of the same coin, via the concept of "shrinkage" or "wavelet thresholding". Following the ideas of Donoho and Johnstone [5], numerous works have emerged recently on non-linear techniques for signal denoising, esp. in the case of additive white noise. The basis recipe for shrinkage is:

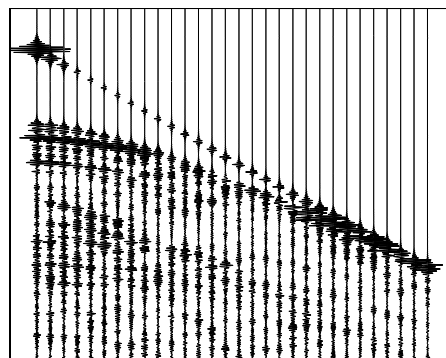
- transform a signal s_n via e. g. a wavelet transform into time-scale domain coefficients w_n ,
- discard some of the w_n , termed as noisy coefficients, according to one or several soft- or hard- thresholds,
- recover s_n by the inverse transform.

After wavelet thresholding, data is generally transformed back to the time domain. But if it is not, assuming the indexes of the remaining wavelet coefficients are known, thresholding could be seen as a means of compression, if one is not concerned by quantization or entropy coding.

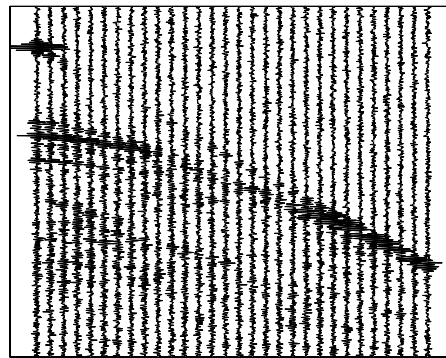
In [6], B. Natarajan initiated the somewhat surprising approach that a good compression algorithm could serve as a good denoising tool. One of his ideas was, when applying lossy compression with an allowed loss set equal to the noise power, the compression loss (noise) and the initial noise tend to cancel rather than add. Others authors addressed the problem in the fields of natural images [7] and geophysics [8]. One could intuitively explain this surprising

approach as follows. Signals of interest traditionally possess strong correlations that efficient coders are able to pack concisely. In contrast, uncoherent noises, such as white noise, exhibit weak correlations or are badly predictable. Good coders could thus be able to tell a redundant signal from a noise, and act implicitly as a noise filter.

As we will see the next sections, compression effectively reduces the ambient noise to some extent. The coder used in this study is based on lapped transform with embedded zerotree coder developed by Tran *et al.* [9] and adapted to seismic signals [10]. The filter banks used here are the 8-channel Walsh-Hadamard transform and a 8-channel 16-tap biorthogonal lapped transform that together yield good compression performance for raw seismic data such as the ones used in this work, in Sect. 4-5. The quantization of coefficients into zero-zones performed by zerotree coding is an approximation of thresholding. It explains the denoising ability of the proposed coder.



(a) Noise-free model data



(b) Noisy model data

Fig. 1. Model data and a noise-corrupted version.

3. ADDITIVE NOISE AND SNR CALCULATIONS

The signal to noise ratio (SNR) is the most common measure of a signal quality. We must be careful of what we call

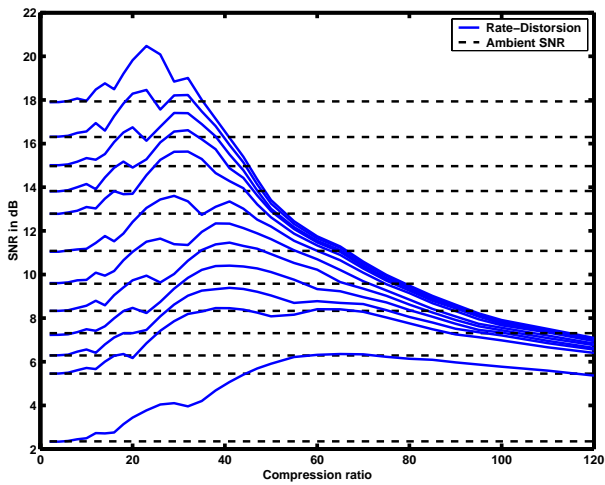


Fig. 2. R-D curve for several ambient noise levels.

signal and what we call noise. Different definitions may lead to different results.

Let $d(t) = s(t) + n_a(t)$ be a typical seismic trace. It is a 1-D signal composed of a coherent seismic signal s and an uncoherent ambient noise $n_a(t)$. If we compress lossily $d(t)$, the recovered signal $\hat{d}(t)$ is generally assumed corrupted by additive compression noise $n_c(t)$. This noise is often believed to be uncorrelated to the data. But to which data exactly? If the coder actually denoises, $n_c(t)$ could be uncorrelated to $s(t)$, but probably not to the original ambient noise $n_a(t)$, at least at low CRs. As a consequence, the recovered signal $\hat{d} = s + (n_a + n_c)$ compared to s may even reach a higher SNR than that of d , compared to s the noise free signal, as long as the energy of $n_a + n_c$ does not exceed that of n_a , which would be difficult if both noises were uncorrelated gaussian noises. We first evaluate the denoising performance of the coder on model seismic data, where we have complete control on the amount of noise (Sect. 4). Actual field data is used in Section 5.

4. RESULTS ON MODEL SEISMIC DATA

Figure 1-(a) represent our model data, resulting from elast-ing modeling based on an actual well log. Each vertical signal represent a seismic trace. When gathered together "side by side", the traces form a shot gather image. The noise-free model is corrupted by an additive gaussian white noise with various levels, playing the role of ambient noise, as in Fig. 1-(b).

Figure 2 represents the rate-distorsion curve (in solid blue) at CRs ranging from 1 to 120, for several ambient noise levels (in straight horizontal black dashed line). The general trend we observe is that SNR first increases with the CR, reaches a maximum and then decreases and finally slips

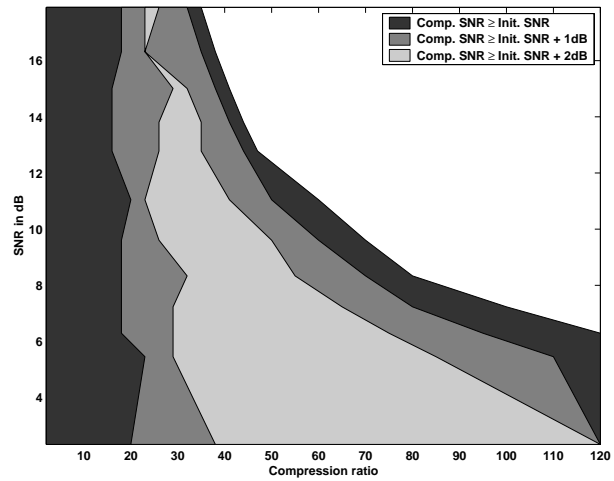


Fig. 3. R-D curve for several ambient noise levels.

under the ambient noise level. The recovered data quality is thus higher (in SNR sense) than the initial data quality for a relatively wide range of CR. SNR improvement may achieve up to 2 dB on the model data.

Figure 3 pictures a schematic view of the domains where the SNR improvement in dB is positive (dark grey), greater than 1 dB (medium grey) and greater than 2 dB (light grey) respectively. For the model data, the bigger the noise, the wider the CR range that yield SNR improvement.

Figure 4 represents one trace signal from the model data, its noisy version, the ambient noise and the results after a 32 : 1 compression, both for the recovered signal and the remaining noise. As we may see, the noise power has been visually reduced, at the expense of a seemingly less random behaviour than that of the original ambient noise. We also remark the results in low signal activity regions are quite poor. This could be explained by a converse of Natarajan's Occam principle: where there is almost no signal, even good coders perform poorly in representing the noise.

5. RESULTS ON ACTUAL FIELD DATA

Figure 5 reflects the effects of compression on one trace signal of actual field data. The signal exhibits more complexity than the model data used in Sect. 4, for instance less low activity regions and no precise knowledge of the ambient noise statistics. We have performed three compression ratios of 20, 30 and 40 respectively, and extracted the noise from the original and the three compressed signals.

It turns out that a higher CR may yield lower noise power (bottom right noise), at the risk of destroying low amplitude seismic information, which if not evident here.

6. CONCLUSIONS AND PERSPECTIVES

We have demonstrated that a lapped transform coder with good compression performance is able to denoise seismic data both in the SNR sense and visually. Results indicate that as the ambient noise increases in the signal (up to a limit), more improvement in SNR could be achieved on a broader range of compression ratios.

However, the compression process could be improved by using soft-thresholding or cleverer thresholding strategies to reduce quantization noise in the remaining noise [7]. Further experiments are needed to interpret the actual effect of the resulting less-white noise on typical seismic processing sequences.

7. REFERENCES

- [1] J. D. Villasenor, R. A. Ergas, and P. L. Donoho, "Seismic data compression using high-dimensional wavelet transforms," in *Proc. 6th Data Compression Conference*. Apr. 1996, pp. 396–405, IEEE computer society press.
- [2] Paul Donoho, "Report on studies to develop diagnostic procedures for safe data compression using lossy compression methods," Tech. Rep., SCDI Steering Committee, Aug. 2001, Draft copy.
- [3] T. Røsten, T. A. Ramstad, and L. Amundsen, "Part I: Subband coding of common offset gathers," *Submitted to Geophysics*, Preprint.
- [4] 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.
- [5] D. L. Donoho and I. M. Johnstone, "Minimax estimation via wavelet shrinkage," Tech. Rep. 402, Stanford University, Department of Statistics, July 1992.
- [6] B. K. Natarajan, "Filtering random noise from deterministic signals via data compression," vol. 43, no. 11, pp. 2595–2605, 1995.
- [7] S. G. Chang, B. Yu, and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression," *IEEE Trans. on Image Proc.*, vol. 9, no. 9, pp. 1532–1546, Feb. 2000.
- [8] N. Saito, "Simultaneous noise suppression and signal compression using a library of orthonormal bases and the minimum description length criterion," in *Wavelets in Geophysics*, E. Foufloula-Georgiou and P. Kumar, Eds., pp. 299–324. Academic Press, 1994.

[9] T. D. Tran, R. L. de Queiroz, and T. Q. Nguyen, "Linear phase perfect reconstruction filter bank: lattice structure, design, and application in image coding," *IEEE Trans. on Signal Proc.*, vol. 48, pp. 133–147, Jan. 2000.

[10] L. C. Duval and T. Nagai, "Seismic data compression using GULLOTS," in *Int. Conf. on Acoust., Speech and Sig. Proc.*, 2001.

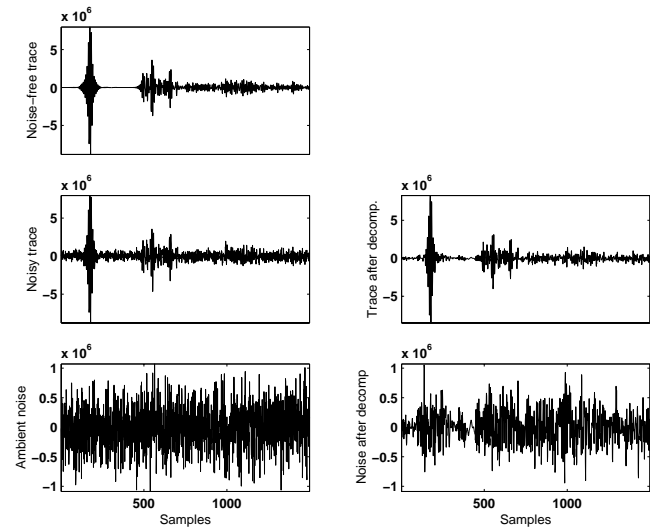


Fig. 4. Model data and noise behaviour before (left, 11.9 dB) and after (right) a 32 : 1 compression.

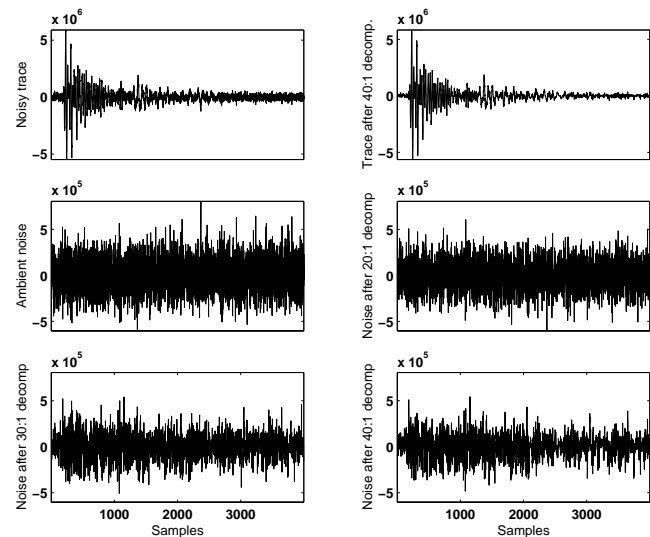


Fig. 5. Actual data and noise behaviour before (left) and after (right) 20, 30 and 40 compression.