xxx

SEISMIC DATA FILTERING WITH HIERARCHICAL LAPPED TRANSFORMS AND HIDDEN MARKOV MODELS

LAURENT DUVAL
IFP, 95852 Rueil-Malmaison Cedex, France, laurent.duval@ifp.fr

Abstract

We propose a method for uncoherent noise removal in geophysical data. The Multiple Wavelet Stacking is based on a concurrent use of wavelet-based shrinkage and data and time-scale dependent threshold choice. Since one singular wavelet does not match all the time-varying properties of a signal, the simultaneous use of several wavelets is able to lower some shrinkage shortcomings, such as wavelet dependency, and to further reduce the residual noises.

Motivation and related works

Wavelet transforms have emerged as efficient tools for signal separation and noise filtering in several geophysical applications. In heavy noise conditions, wavelet-based techniques generally owe their robustness to their attractive time-scale properties. Wavelet based denoising, or wavelet shrinkage [Miao and Cheadle, 1998], arises from the work of D. Donoho, based on the energy compaction properties of these time-scale operators [Donoho, 1995]. With the most widely used signal model [Ulrych et al., 1999], let R represent an actual geophysical record. It can be decomposed in an underlying signal S, and two noise components N_c and N_r, respectively coherent and random:

$$R = S + N_c + N_r.$$

In the remaining of this paper, we will focus on the random noise component N_r. One useful property is that the orthogonal transform of a white noise remains a white noise. Meanwhile, a coherent signal is generally efficiently and sparsely represented after a wavelet transform. Donoho [1995] demonstrated that for additive noise, a simple thresholding procedure would discard ”mainly” noise coefficients. This result has been shown to be asymptotically near optimal for a wide class of signals corrupted by gaussian white noise.

As stated in [Ioup and Ioup, 1998], ”the choice of the wavelet and its associated scaling function are very important to obtain the most useful wavelet transform”. As a result, a good wavelet should be matched to the data properties. Unfortunately, discrete wavelets with fast algorithms are relatively rare, and not trivialy related to geophysical signals. Moreover, we cannot expect optimal energy compaction from a single wavelet domain, since most signals are non-stationary and contain a variety of frequency contents. Wavelet packets represent an enrichment of the projecting vectors. This concept is generalized for instance in [Saito, 1994]. The signal S is estimated from a noisy observation using a library of orthonormal bases, consisting in various wavelets, wavelet packets and local trigonometrics bases. A procedure, based the information-theoretic Minimum Length Description criterion, chooses the best base as a compromise between the fidelity (denoising) and the efficiency of the signal estimation (compression). Both goals may be attained simultaneously

[Chang et al., 2000]. Another related option consists in ”physical wavelets” that resemble typical seismic signals or shapes, as proposed in [Wu et al., 1998, Zhang and Ulrych, 2002]. The latter approach is generally considered as computationaly expensive.

In this work, we propose a denoising procedure based on concurrent denoising using several wavelet bases. It yields a multiple wavelet-domain signal estimation with data driven parameters. It is followed by stacking the multiple signal estimations. This algorithm is able to reduce dependencies to the wavelet choice, as well as artifacts (for instance pseudo-Gibbs ringing) due to each single wavelet estimation. Since the estimated denoised traces contain residual uncorrelated quantization noise, the stack mimic a multi-fold estimation succeptible of further noise reduction.

Wavelets as multidimensional signals approximation tools

The discrete wavelet transform (DWT) represents a 1-D signal $s(t)$ as a linear combination of shifted versions of a low-pass scaling function $\phi(t)$ and shifted and dilated copies of a band-pass wavelet function $\psi(t)$ [Mallat, 1998]. For special choices of $\phi(t)$ and $\psi(t)$, j and k being relative integers, the functions $\phi_{j,k}(t) = 2^{j/2}\phi_{j,k}(2^j t - k)$ and $\psi_{j,k}(t) = 2^{j/2}\psi_{j,k}(2^j t - k)$ form an orthonormal basis. The signal $s(t)$ is represented as (time variable omitted):

$$s = \sum_k v_{j_0,k}\psi_{j_0,k} + \sum_{j=j_0}^{\infty}\sum_k w_{j,k}\psi_{j,k}.$$

Wavelets approximate signals localy in time and frequency (roughly an inverse of the scale) in the following way. The scaling coefficient $v_{j_0,k}$ measures the local mean around time $2^{-j_0}k$ and the wavelet coefficient $w_{j,k}$ approximates the details of the signal content around time $2^{-j}k$ and frequency $2^j f_0$. Two-dimensional scaling functions and wavelets on time-space are constructed on the four products of one-dimensional wavelets $\phi(t)\phi(x)$, $\psi(t)\phi(x)$, $\phi(t)\psi(x)$, $\psi(t)\psi(x)$. It is easily generalized to higher dimension.

Wavelet-domain models and threshold selection

Simple wavelet coefficient models state that the wavelet coefficients are independent from each other. The Gaussian distribution generally does not model appropriately the marginal distribution of the wavelet coefficients in each band. More accurate is the zero-mean generalized Gaussian distribution (GGD) defined by its variance σ^2 and a shape parameter ν. The associated funtion p is, up to a normalization constant A:

$$p(x) = A(\sigma,\nu)\exp\left(-|\eta(\nu)|x|/\sigma|^{\nu}\right).$$

All coefficients at one scale, at index j, are often assumed to follow the same GGD model with specific σ_j^2 and ν_j. For a large class of signals, and for practical considerations, the shape parameter is often assumed to be constant across the scales.

Assume that the correct GGD model is known for each subband. The wavelets coefficients $w_{j,k}$ need to be thresholded according to a data and time-scale band dependent threshold λ_j. The two most frequently used methods are:

- *soft-thresholding*: it takes the coefficient w_j and shrinks it toward zero with respect to the threshold λ_j, according to the function

$$\mathrm{ST}(w_j) = \mathrm{sign}(w_j)\cdot\max\left(|w_j| - \lambda_j, 0\right),$$

- *hard-thresholding*: it discards every coefficient smaller than the threshold λ_j,

$$\mathrm{HT}(w_j) = w_j \cdot \mathbb{1}\{|w_j| \leq \lambda_j\},$$

where $\mathbb{1}\{\cdot\}$ is the set characteristic function.

We now need to estimate the noise variance σ^2. In some situations, it may be estimated from low activity portions of the signal (before the first break). It is also possible to perform an estimation in the highest frequency subband, using the robust median estimator

$$\hat{\sigma} = \text{Median}(|w_k|)/0.6745.$$

The shape parameter ν can be retrieved from calculations involving the kurtosis of the coefficients. We refer to Chang et al. [2000] for detailed explanations on the subband threshold selection.

Multiple wavelet stacking and a toy example

The proposed algorithm works as follows: assume we have a set of relatively good orthogonal wavelet filters. In the following, we have exclusively used Daubechies wavelets, with 7 to 14 vanishing moments [see Mallat, 1998, p. 166 sq.]. For each wavelet, data and scale dependent skrinkage is applied to the noisy data, as detailed in the preceding section. Each denoised signal contains residual noise from noise model mismatch, sub-optimal threshold selection or wavelet-domain quantization. This residual noise is rarely uncorrelated to the original noise, and generally exhibits non conventional statistics, or spoors from the wavelet shape. But from one wavelet shrinkage to another, residual noises are generally poorly correlated, while the underlying signal stays exactly in phase. The denoised signals are then stacked with hope to further reduce the uncorrelated residual noises.

Figure 1 demonstrates the MWS (Multiple Wavelet Stacking) effect on a simple chirp segment. On the top left, several shrinkage realizations with differents wavelets are superposed (in colour). The MWS signal on the top right is visibly slightly less noisy that its colleagues. The bottom pictures display the residual noises assuming we knew exactly the original signal. On the left, the residual noises from single wavelet shrinkage are clearly stronger that the MWS estimate on the right.

Real data results and discussion

The MWS has been applied to a portion of a seimic shot gather, displayed in Fig. 2-1. After Multiple Wavelet Stacking, the stacked shot in Fig. 2-2 exhibits low uncoherent noise level, and does not exhibit spurious wavelet-shaped artifacts, nor unsmoothness sometimes associated to hard-thresholding.

Conclusions

We have proposed a denoising algorithm based on concurrent wavelet shrinkage with differents discrete wavelet transforms, using data and scale dependent thresholding. The different denoising realizations are then stacked, in order to reduce the residual noise. As a result, the Multiple Wavelet Stacking algorithm is relatively immune to the wavelet or the decomposition choice. It does not seem to be affected by quantization or spurious wavelet artifacts.

References

S. G. Chang, B. Yu, and M. Vetterli. Adaptive wavelet thresholding for image denoising and compression. *IEEE Trans. on Image Proc.*, 9(9):1532–1546, Sep. 2000.

D. L. Donoho. De-noising by soft-thresholding. *IEEE Trans. on Inform. Theory*, 41(3):613–627, May 1995.

J. W. Ioup and G. E. Ioup. Noise removal and compression using a wavelet transform. In *Annual International Meeting*, pages 1076–1079. Soc. of Expl. Geophysicists, 1998. Exp. abstracts.

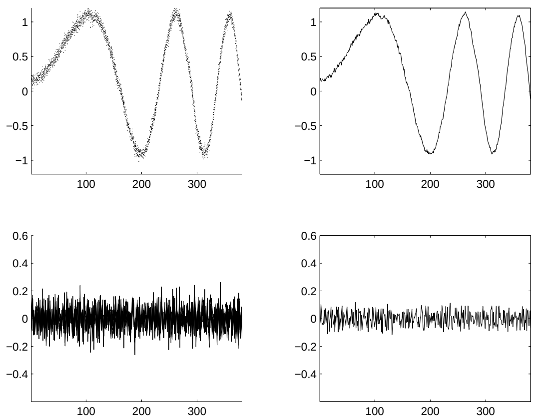
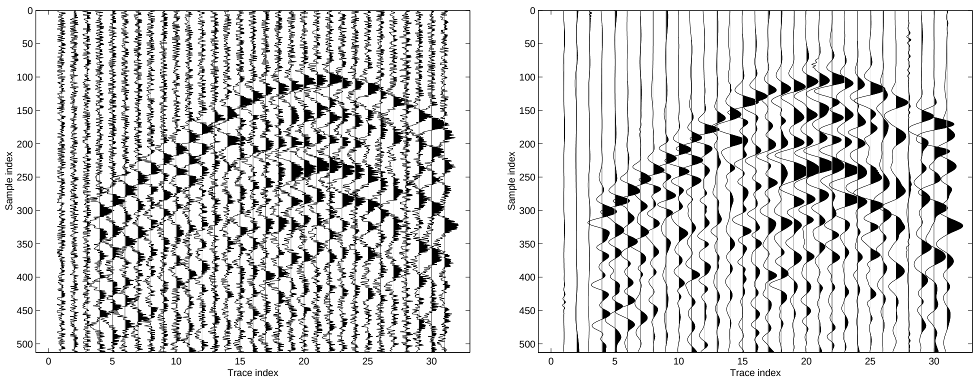


Figure 1: From left to right, and top to bottom: (1) several denoised signal realizations, (2) multiple wavelet stack, (3) residual noise realizations, (4) multiple wavelet stack residual noise.



(1) (2)

Figure 2: (1) Noisy shot gather (2) Denoised by multiple wavelet stacking.

Stéphane Mallat. *A wavelet tour of signal processing*. Academic Press, 1998.

X. G. Miao and S. Cheadle. Noise attenuation with wavelet transforms. In *Annual International Meeting*. Soc. of Expl. Geophysicists, 1998. Exp. abstracts.

N. Saito. Simultaneous noise suppression and signal compression using a library of orthogonal bases and the minimum description length criterion. In E. Foufoula-Georgiu and P. Kumar, editors, *Wavelets in Geophysics*, pages 299–324. Academic Press, Inc., 1994.

T. Ulrych, M. Sacchi, and J. M. Graul. Signal and noise separation: art and science. *Geophysics*, 64(5):1648–1656, 1999.

Ru-Shan Wu, Xiao-Long Dong, and Jin-Huai Gao. Application of acoustic wavelet transform to seismic data processing. In *Annual International Meeting*. Soc. of Expl. Geophysicists, 1998. Exp. abstracts.

R. Zhang and T. Ulrych. Physical wavelet frame denoising of seismic data. In *Conference and Technical Exhibition*. European Assoc. of Geoscientists and Engineers, 2002. Exp. abstracts.