3036 CURVELET DENOISING OF 4D SEISMIC

MORITZ BEYREUTHER, JAMIN CRISTALL, FELIX J. HERRMANN
Department of Earth and Ocean Sciences, University of British Columbia, Canada

Abstract

With burgeoning world demand and a limited rate of discovery of new reserves, there is increasing impetus upon the industry to optimize recovery from already existing fields. 4D, or time-lapse, seismic imaging is an emerging technology that holds great promise to better monitor and optimise reservoir production. The basic idea behind 4D seismic is that when multiple 3D surveys are acquired at separate calendar times over a producing field, the reservoir geology will not change from survey to survey but the state of the reservoir fluids will change. Thus, taking the difference between two 3D surveys should remove the static geologic contribution to the data and isolate the time-varying fluid flow component. However, a major challenge in 4D seismic is that acquisition and processing differences between 3D surveys often overshadow the changes caused by fluid flow. This problem is compounded when 4D effects are sought to be derived from vintage 3D data sets that were not originally acquired with 4D in mind. The goal of this study is to remove the acquisition and imaging artefacts from a 4D seismic difference cube using Curvelet processing techniques.

The denoising problem

In this paper, we argue that computing 4D difference cubes can be recast into the framework of solving a generic denoising problem that estimates the model $\mathbf{m}$ from noisy data [See e.g. 10]

$$\mathbf{d} = \mathbf{m} + \mathbf{n} \tag{1}$$

with white Gaussian noise $\mathbf{n}$. The solution of this inverse problem can be written in terms of the following variational problem [10, 11]

$$\hat{\mathbf{m}}: \quad \min_{\mathbf{m}} \frac{1}{2}\|\mathbf{d}-\mathbf{m}\|_2^2 + \mu J(\mathbf{m}), \tag{2}$$

where $J(\mathbf{m})$ is an additional penalty function that contains *prior* information on the model, such as particular sparseness constraints. The control-parameter μ rules how much emphasis one would like to give to the *prior* information on the model.

The question now is: how can we solve this denoising problem effectively? In other words, how can we construct a diagonal decision operator that minimizes the energy difference between the estimate and the true model given the *prior* information? It appears from the work of [7], [10] and others that, for a certain class of models, one can obtain nearly optimal denoising results, i.e. near optimal SNR for denoised data, by projecting noisy data onto a basis-function representation that is optimal for that particular class of models. In that case, most of the model's energy resides in only a few coefficients, allowing for the definition of a shrinkage estimator that separates noise from the model. For basis functions that are also local, one can show that soft thresholding on the coefficients suffices to approximately solve the above denoising problem, i.e.

$$\hat{\mathbf{m}} = \mathcal{D}(\mathbf{d}) = \mathbf{B}^{-1}\theta_\mu\left(\mathbf{B}\mathbf{d}\right). \tag{3}$$

In this expression, $\mathbf{B^{-1}}$ refers to the (pseudo)-inverse of $\mathbf{B}$, which is the basis-function expansion. θ_μ is a soft/hard thresholding operator with a threshold that for orthonormal basis functions equals [10, 7] $\mu = \sigma\sqrt{2\log_e N}$ with σ the standard deviation of the noise and N the number of data samples. Soft thresholding solves Eq. (3) on the coefficients for a penalty function given by the L^1-norm, i.e. $J(\mathbf{m}) = \|\mathbf{m}\|_1$. Comparison of Eq.'s (2) and (3) establishes the connection between the threshold and the control parameter μ.

Wavelets, and their recent extension to Curvelets, derive their success mainly from their ability to locally and sparsely represent the model, facilitating the definition of simple non-linear thresholding estimators that locally decide (adaptively filter) whether a certain "event" pertains to the model or to the noise. The question now is: can we extend these Wavelet-adaptive filtering ideas to the computation of 4D difference cubes? Before answering this question, let's be

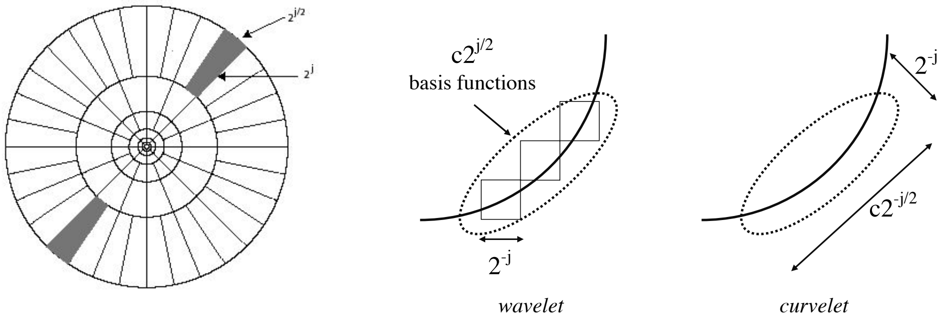

Figure 1: **Left:** Curvelet partitioning of the frequency plane [modified from 3]. **Right:** Comparison of non-linear approximation rates Curvelets and Wavelets [modified from 6].

more specific with respect to the choice of the appropriate basis functions for representing seismic data volumes.

Localized basis-function decomposition

Curvelets as proposed by [2, 1, 4], constitute a relatively new family of non-separable Wavelet bases that are designed to effectively represent seismic data with reflectors that generally tend to lie on piece-wise smooth curves. This property makes Curvelets suitable to represent reflectors within vertical/horizontal slices of migrated volumes- including faults. For these type of signals, Curvelets obtain nearly optimal sparseness, yielding (i) a rapid decay for the reconstruction error as a function of the largest coefficients; (ii) concentration of the signal's energy in a limited number of coefficients; (iii) relatively easy separation of noise *versus* model. So how do Curvelets obtain such a high non-linear approximation rate? Without being all inclusive [see for details 2, 1, 4, 6], the answer to this question lies in the fact that Curvelets are

- *multi-scale*, i.e. they live in different dyadic corona (see Fig. 1(left)) in the (k_x, k_y)-domain (or equivalently (k_x, k_z)-domain).
- *multi-directional*, i.e. they live on wedges within these corona (see Fig. 1(left)).
- *anisotropic*, i.e. they obey the following scaling law: width $\propto$ length^2.
- *directional selective* with $\#$ orientations $\propto \frac{1}{\sqrt{\text{scale}}}$.
- *local* both in $(x, y))$ and (k_x, k_y).
- almost *orthogonal*- they are *tight* frames with a moderate redundancy. Contourlets implement the pseudo-inverse in closed-form while Curvelets provide the transform and its adjoint, yielding a pseudo-inverse computed by iterative Conjugate Gradients.

Clearly, these properties make Curvelets the appropriate basis functions for dealing with seismic data. However, so far no 3-dimensional Curvelets have been developed so we have the choice to apply Curvelet transforms to either depth slices or to the in-line direction, followed by applying a shift-invariant Wavelet transform [5] in the depth or cross-line direction.

Adaptive subtraction

Given the above basis-function decomposition, how can we recast the computation of 4D difference cubes into a denoising problem? The problem is that we would like to reduce (difference) noise, which we define as representing (i) the presence of possible additional incoherent noise on the two volumes; (ii) possible remaining migration misalignments due to differences in processing and (iii) differences in acquisition footprints. During this noise removal, we would like to preserve and emphasize 4D effects. In light of these objectives, we choose not to actually subtract the two datasets but rather filter one with the other. We consider one dataset to be the "noise" $\mathbf{n}$ (not to be mistaken with the noise defined above) of the other, the data $\mathbf{d}$ (cf. Eq.1), and carry out our program by a denoising procedure-yielding an estimate for the difference $\hat{\mathbf{m}}$.

The crux of any filtering operation lies in being able to optimally represent both the model (the difference) and noise (the second dataset). This optimality refers to the basis-function's ability to sparsely represent both the model and the noise. If, in addition, the basis functions are localized, *superior* filtering results are obtained, i.e. better signal-to-noise ratios, when solving linear problems, such adaptive subtraction, non-linearly. This superiority is not only due to a reduction of the dimensionality of the subtraction problem but is also related to the locality of the basis functions, permitting the use of non-linear estimators based on thresholding [see e.g. 10]. These estimators make localized decisions on whether certain events are 4D effects or belong to the (difference) noise, consisting of incoherent noise, processing, imaging, or acquisition artifacts, based on the magnitude of the coefficients in the basis-function decomposition. The better the basis functions approximate the model and, in the case of coherent noise, the noise, the better the local decision operator will be able to discriminate between model (difference) and noise (remaining differences). Since the basis functions we are using are local, the decision operator can be a simple hard- or soft-thresholding operator [10], which shrinks the coefficients towards zero that are below a certain threshold, which depends on the (local) noise-level. We refer to another paper in these proceeding (by the same last author), where the same technique is used to successfully remove predicted multiples and migration noise.

As the example in the next section demonstrates, the optimal denoising capabilities for incoherent noise (cf. Eq. 3) carry over to coherent noise removal [see for successfull removal of coherent migration noise 8], provided we have a reasonable consistency in the processing, imaging, and data acquisition between the two vintages. By choosing a threshold defined by one of the vintages, i.e.

$$\mu \sim 3\eta|\mathbf{B}\mathbf{n}|, \tag{4}$$

with $\mathbf{B}$ representing the combination of Curvelet and Wavelet decompositions, we are able to adaptively decide whether a certain coherent feature belongs to an actual 4D effect or simply to (difference) noise. In this expression, η is a threshold control-parameter, which allows us to (de)-emphasize the subtraction and the 3 is related to the $95\,\%$ confidence interval.

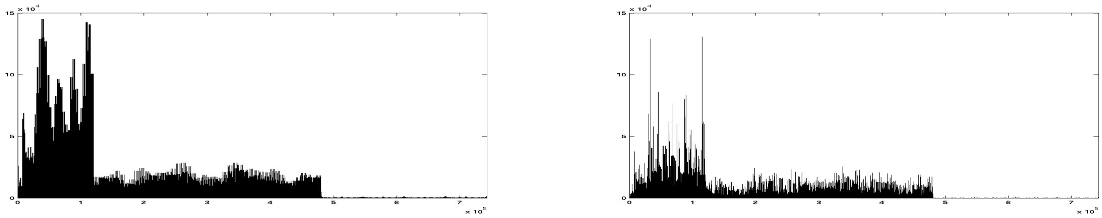

Figure 2: Curvelet coefficient vectors before and after thresholding

Example

Two 3D seismic surveys were acquired at the Ula North Sea oil field in 1984 and 1999 respectively. Even though the datasets have been processed with 4D in mind, the fact that the two surveys were shot in perpendicular directions and with different acquisition systems may complicate the extraction of 4D effects [9].

For the present case, our goal is to suppress the imaging/processing and acquisition artifacts such that the true 4D signature may be better expressed. The Curvelet-estimation technique presented in this paper was used and we proceeded as follows. First, the Curvelet transform was applied along slices in the in-line direction of the 1984 and 1999 vintage datasets. This was followed by a non-decimated Wavelet transform in the cross-line direction. Subsequently, the threshold was calculated according Eq. 4 with $\mathbf{n}$, given by the 1984 dataset and $\eta = 1$. Finally, we applied the thresholding procedure defined in Eq. 3 on $\mathbf{d}$ given by the 1999 dataset. The effects of the thresholding on the vector with the basis-function coefficients are shown in Fig. 2.

To illustrate our results, we compare in Fig. 3 the reconstruction from the thresholded coefficients with the ordinary difference cube along the picked Top Ula horizon which caps the producing reservoir. The superimposed logarithmic color code is representative for averaged amplitude difference magnitudes over the first 5 samples below the picked horizon. Hot colours correspond to large difference while cold colors correspond to small differences.

The first frame of Fig. 3 shows the difference magnitudes for the ordinary difference cube, which suffers from artifacts induced by the colored (difference) noise which seem to overshadow possible smaller differences caused by fluid flow. Notice the particularly large differences at the faulted and steeply dipping part of the horizon near record 10, trace 25. The second frame of Fig. 3 shows difference projected onto the Top Ula horizon after Curvelet processing. The (difference) noise removal is apparent and it is believed that the remaining differences result from true reservoir changes.

Discussion

We presented an alternative method to compute difference cubes. Our method is not based on the actual subtraction of two dataset, risking the introduction of artifacts due to noise, possible misalignments, and differences in processing and acquisition. Rather, it mutes events of one dataset with respect to the other on coefficients of a sparse and local basis function decomposition. In this sense, we do not introduce artifacts by subtracting misaligned events. Instead, events with a strong presence in the other dataset are muted. Even though the example shown is preliminary, it shows promise; as does the application of this method to adaptive subtraction of predicted multiples and migration, on which is reported elsewhere in these proceedings.

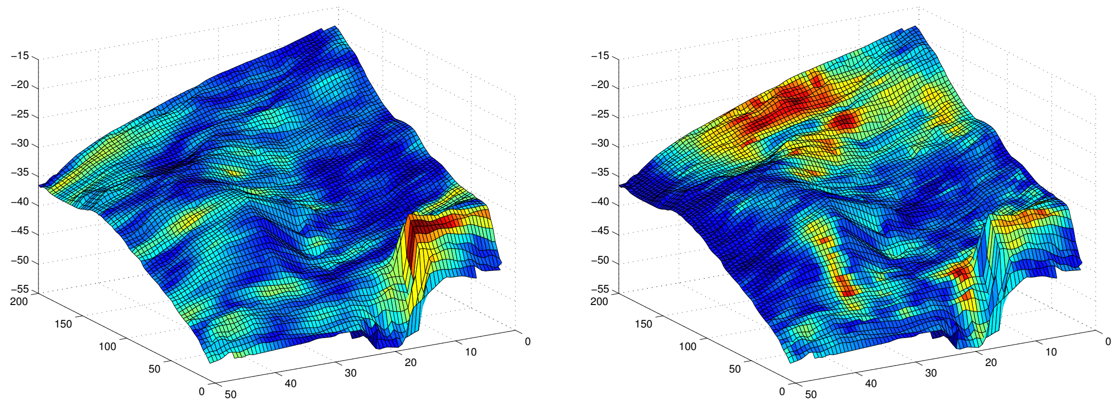

Figure 3: 1984-1999 difference projected onto the Top Ula horizon before and after Curvelet processing

Acknowledgments

The authors would like to thank Paul R. Gutowski and Per Gunnar Folstad of BP for their advice and thank the Ula license partners, BP Norge, Svenska Petroleum, and Dong Norge, for their generous permission to use their data. We would also like to thank Emmanuel Candés and David Donoho for making an early version of their Curvelet code available (Digital Curvelet Transforms via Unequispaced Fourier Transforms, presented at the ONR Meeting). This work was in part financially supported by a NSERC Discovery Grant.

References

[1] E. J. Candès and D. Donoho. New tight frames of curvelets and optimal representations of objects with smooth singularities. Technical report, Stanford, 2002. submitted.

[2] E. J. Candès and D. L. Donoho. Curvelets – a surprisingly effective nonadaptive representation for objects with edges. Curves and Surfaces. Vanderbilt University Press, 2000.

[3] E. J. Candès and D. L. Donoho. Recovering Edges in Ill-posed Problems: Optimality of Curvelet Frames. *Ann. Statist.*, 30:784–842, 2000.

[4] E. J. Candès and F. Guo. New multiscale transforms, minimum total variation synthesis: Applications to edge-preserving image reconstruction. *Signal Processing*, pages 1519–1543, 2002.

[5] R. R. Coifman and D. L. Donoho. Translation-invariant de-noising. Technical report, Stanford, Department of Statistics, 1995.

[6] M. Do and M. Vetterli. *Beyond wavelets*, chapter Contourlets. Academic Press, 2002.

[7] D. L. Donoho and I. M. Johnstone. Minimax estimation via wavelet shrinkage. *Annals of Statistics*, 26(3):879–921, 1998.

[8] F. J. Herrmann. Multifractional splines: application to seismic imaging. In A. F. L. E. Michael A. Unser, Akram Aldroubi, editor, *Proceedings of SPIE Technical Conference on Wavelets: Applications in Signal and Image Processing X*, volume 5207, pages 240–258. SPIE, 2003.

[9] E. K. Ian F. Jones, Per Gunnar Folstad. Continuous high resolution velocity as a 4d attribute. In *CSEG Conference*, 2002.

[10] S. G. Mallat. *A wavelet tour of signal processing*. Academic Press, 1997.

[11] C. Vogel. *Computational Methods for Inverse Problems*. SIAM, 2002.