

Curvelet imaging and processing: an overview

Felix J. Herrmann

Department of Earth and Ocean Sciences, University of British Columbia, Canada

2004 CSEG National Convention



Abstract

In this paper an overview is given on the application of directional basis functions, known under the name Curvelets/Contourlets, to various aspects of seismic processing and imaging. Key concepts in the approach are the use of (i) directional basis functions that localize in both domains (e.g. space and angle); (ii) non-linear estimation, which corresponds to localized muting on the coefficients, possibly supplemented by constrained optimization (iii) invariance of the basis functions under the imaging operators. We will discuss applications that include multiple and ground roll removal; sparseness-constrained least-squares migration and the computation of 4-D difference cubes.

Introduction

Everybody would agree that a seismologists dream would be the existence of a basis-function decomposition which localizes in space and angle, is very sparse and well behaved under imaging operators. Well, this dream has come true with the recent introduction of directional wavelets, known under the names Curvelets or Contourlets [5,2,1,4]. As their names suggest, these basis functions are designed to represent images with edges on piece-wise smooth curves very sparsely. In our language, this corresponds to a decomposition that is nearly optimal (read as sparse as you can get) for seismic reflection events as well as for reflectors in the earth. Unfortunately, this property by itself is not enough since we are dealing with imaging operators, which may render the basis-functions less effective. For instance, discrete wavelets transforms have failed to really make a impact in migration because their lack of directional selectivity not only makes them less suitable for migration but they also fail to effectively represent the earth's reflectors and seismic data. Curvelets (from now on I will use the name Curvelets to refer to both Curvelets and Contourlets), on the other hand, remain more or less invariant under imaging operators while they obtain nearly optimal sparse representations for seismic data as for well images.

Our efforts are built on the premise that one is much better suited to solve linear imaging/filtering problems if one is able to very sparsely represent both data and denoised image [7,8]. Therefore, an important part of our research is and will be devoted towards the application of Curvelets to (i) coherent denoising problems such as multiple, ground roll elimination and the robust computation of 4D-difference cubes (refer to other contributions of the author to the Proceedings of this conference) and to (ii) imaging, where the sparseness and invariance of Curvelets are used to improve the noise-robustness and computational efficiency of sparseness-constrained least-squares migration.

Our approach differs in several respects from previous attempts to use optimal basis-function decompositions in seismic processing and imaging. First, we do not solely focus on the compression of data or operators. Instead, we concentrate on the design and implementation of basis functions which (i) sparsely represent data and model; and (ii) which are preserved under (de)-migration. This combination of features allows us to reformulate the seismic imaging and inversion problems in terms of non-linear operations on the coefficients of the basis-function decomposition. By replacing linear inversion techniques by non-linear estimation techniques based on thresholding, we limit the smoothing due to regularization and hence we preserve the edges. Therefore, the success of our approach depends primarily on the degree of sparseness we can achieve to represent both data and image. In this abstract, we will present the basic methodology and we mention a number of applications. For more detail, the reader is referred to a series of accompanying abstracts to be presented at this Conference.

Methodology

The inverse problem

Virtually any linear problems in seismic processing and imaging can be seen as a special case of the following generic problem: how to obtain m form data d in the presence of (coherent) noise n :

$$d=Km+n, \quad (1)$$

in which, for imaging, K is the de-migration operator or the identity for noise-elimination by adaptive subtraction problem. The solution of the inverse problem has the following form [see e.g. [9,11]]

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

where $J(m)$ 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. How can we find the appropriate domain to solve this inverse problem? In other words how can exploit the locality of the basis functions in the above setting?

It appears from the work of [6], [9] and others that, for a certain class of data and/or images, one can obtain almost 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, diagonal decision operator (read simple muting of coefficient) can be constructed that minimizes the energy difference between the estimate and the true model, i.e.

$$\text{find}_D: \hat{m}=D(d) \text{ such that } \min_D |\hat{m}-m|_2^2. \quad (3)$$

For basis functions that are local, one can show that simple thresholding on the coefficients [9,6] suffices to solve the above denoising problem, i.e.

$$\hat{m}=D(d)=B^{-1}\theta_\mu(d). \quad (4)$$

In this expression, B, B^{-1} refer to the basis-function expansion and its (pseudo)-inverse. θ_ν is a soft/hard thresholding operator with a threshold that for orthonormal basis functions equals [9,6]

$$\mu=\sigma\sqrt{2\log_e N}, \quad (5)$$

with σ the standard deviation of the noise and N the number of data samples. Comparison of equations (2) and (4) establishes the connection between the threshold and the control parameter ν . Question of course now is: can we extend these results, proven to be valid for $K=I$ and n white Gaussian noise, to include the effects of (i) an operator (e.g. K a de-migration operator) and/or (ii) of colored additive noise (e.g. coherent ground roll and multiples)? In other words, can we decolor predicted coherent noise such that the above thresholding procedure becomes effective again? Before answering this question let us first be more specific with respect to the choice of the appropriate basis functions for seismic data, model and coherent noise. To wet our appetite, I included in Fig. 3 an example of thresholding, which clearly demonstrates the potential of the approach applied to the coefficients of Curvelet transforms performed on constant-angle-sorted Common Angle Gathers in combination with a wavelet transform in the angle direction. Even though, we ignored in this example the influence of coloring of noise by migration we greatly improved the stack.

The basis functions

Curvelets as proposed by [5] and [2,1,4] respectively, constitute a relatively new family of non-separable wavelet bases that are designed to effectively represent functions with singularities (read wavefronts) that lie on piece-wise smooth curves. For these type of functions, non-separable basis functions obtain nearly optimal non-linear approximation rates (NAR). For comparison, the m -term expansion of the sorted 2-D Fourier transform obtains for the object x [3,4,5] the following NAR (measured in the L^2 -norm),

$$|x-\tilde{x}_m|_2^2 \propto m^{-1/2}$$

while separable wavelets (ordinary 2-D discrete wavelet transforms) obtain

$$|x-\tilde{x}_m|_2^2 \propto m^{-1}.$$

The improvement for wavelets has been one of the main reasons of their success in image processing. Clearly, the further improvement in NRA by Curvelets, which obtain

$$|x-\tilde{x}_m|_2^2 \propto m^{-2}(\log m)^3,$$

will likely open up a whole new series of successful applications. Besides the logarithmic term, the above NRA is equivalent to the best possible approximation rate (e.g. by a triangular meshing given the location of the edges/reflection events) attainable for this type functions. In Donoho's words the true miracle of wavelets is that they find the singularities (read wavefronts/reflection events) at virtually no additional cost. While this statement holds for piece-wise smooth 1-D functions, it does not hold for curved events. Curvelets, on the other hand, remedy the wavelets as well the Fourier shortcomings. For example, a simple spike gives rise to a full Fourier spectrum.

So how do Curvelets obtain such a high non-linear approximation rate? without being all inclusive [see for details [2,1,4,5]], the answer to this question lies in the fact that Curvelets are

- *multi-scale*, i.e. they live in different dyadic corona (see Figure 1) in the FK-domain.
- *multi-directional*, i.e. they live on wedges within these corona (see Figure 1).
- *anisotropic*, i.e. they obey the following scaling law $\text{width} \propto \text{length}^2$.
- *directional selective* with $\#\text{orientations} \propto \frac{1}{\sqrt{\text{scale}}}$.
- *local* both in (x,t) and (k,f) .
- 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.

In Figure 1, a number of Curvelet properties are detailed [adapted from [3,4,5]]. Figure 1 describes the Curvelet partitioning of the frequency-wavenumber plane. One Curvelet lives in a wedge and becomes more *directional selective* and *anisotropic* for the higher frequencies (while moving away from the origin). Curvelets are localized in both the space (or (x,t)) and spatial (KF)-domains and have, as consequence of their partitioning, the tendency to align themselves with curves/wavefronts (see Figure 1). As such they are more flexible than a representation yielded by e.g. high-resolution Radon [as described by e.g. [10]], because they are local and able to follow any piece-wise smooth curve. Only Curvelets that align with reflectors yield large inner products and this explains their success in solving the denoising problem for models that contain events on curves. The noise is canceled because the basis functions adapt themselves locally to the dip and hence optimizing their overlap with the data integrating out the incoherent noise. Now what about coherent noise such as predicted multiples?

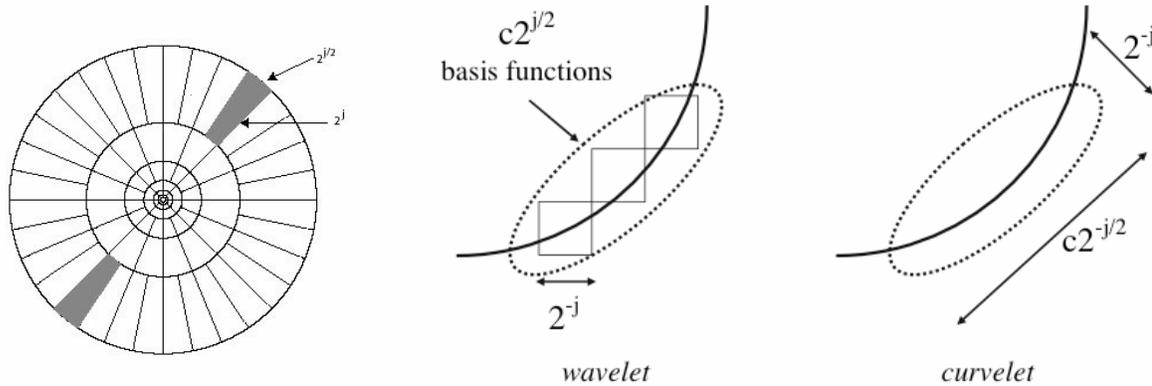


Figure 1: **Left:** Curvelet partitioning of the frequency plane [modified from [3]]. **Right:** Comparison of non-linear approximation rates Curvelets and Wavelets [modified from [5]].

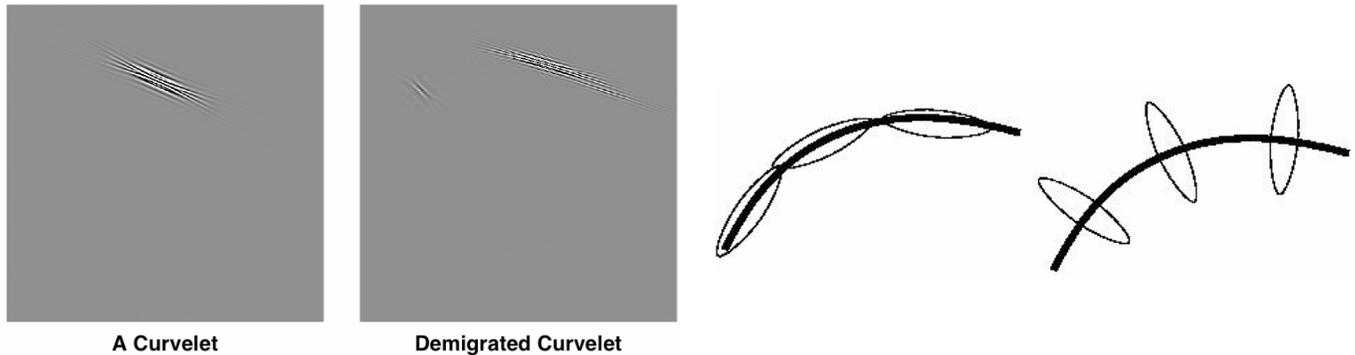


Figure 2: Properties of the Curvelet Transform under imaging. Curvelet and its demigration. Clearly, the Curvelet remains Curvelet-like, which can be explained from its excellent frequency-localization. Far right panel illustrates why Curvelet coefficients are large when they align and small when they are not.

Thresholding, constrained optimization and sparseness constraints

Given the Curvelets and the proposed thresholding procedure how can we operate in an environment where there is coherent additive noise? Well in cases where we have reasonable accurate predictions for the noise, we can simply weight the above threshold (cf. Eq. 5) with the Curvelet coefficients of the predicted noise, i.e.

$$\mu \sim \eta |Bn|. \quad (6)$$

Given the predicted noise, we are now able to adaptively decide whether a certain event belongs to signal or to the noise. The n contains the predicted noise and η represents an additional control parameter which sets the confidence interval (e.g. 95 % for $\eta=3$) (de)-emphasizing the thresholding.

Even though, thresholding takes care of the bulk of the separation of noise and signal, there remains the desire to (i) correct for an operator, if present, to restore the amplitudes; (ii) impose additional *prior* information on the model, e.g. certain sparseness constraints. With the latter, we hope to remove possible estimation and side-band artifacts, enhancing the continuity along the imaged reflectors. To accomplish these goals, we formulate the following constrained optimization problem [4]

$$\hat{m}: \min_m J(m) \quad s.t. \quad |\tilde{A}\tilde{m}-\hat{u}| \leq \mu, \quad (7)$$

where our initial estimate by thresholding, \hat{u} , is updated, subject to constraints, as to minimize the penalty function. The symbol $\tilde{\cdot}$ is used to refer to the coefficients and operators in the Curvelet domain.

Applications

The methodology presented in this paper has been applied to several areas in seismic processing and imaging. The applications can roughly be divided into *adaptive subtraction* ($K=I$ hence $\tilde{A}=I$) for

- **multiple elimination**, where predicted multiples and multiple-free data are represented by the coherent noise n and model m , respectively.
- **ground roll elimination**, where ground roll-free data is the model m and ground roll the noise n .
- **computation 4-D difference cubes**, where one vintage dataset is the noise of the other and *vice versa*.

and *sparseness-constrained least-squares migration* where d is migrated data, $A=K^*K$, the normal operator and n colored noise (by migration). See for details, other contributions by the authors to the Proceedings of this Conference.

Conclusions

Why do Curvelets seem to be the appropriate choice to formulate problems in seismic processing and imaging. The answer to this question is relatively simple. Curvelets provide an almost orthogonal decomposition w.r.t. to basis functions that have far *superior* (compared to e.g. Fourier, Radon, Wavelet) localization and sparseness properties than other transforms. In effect, they can for the high frequencies be seen as local Radon transforms with perfect reconstruction. Their locality and directional selectivity renders local thresholding/muting very effective, while their almost diagonalization of imaging operators potentially makes these basis functions a good choice to solve imaging problems, including the construction of fast migration operators.

Acknowledgments

The authors would like to thank Emmanuel Candès and David Donoho for making their Digital Curvelet Transforms via Unequally-spaced Fourier Transforms available (presented at the ONR Meeting, University of Minnesota, May, 2003). We also would like to thank Mauricio Sacchi for his migration code. This work was in part financially supported by a NSERC Discovery Grant.

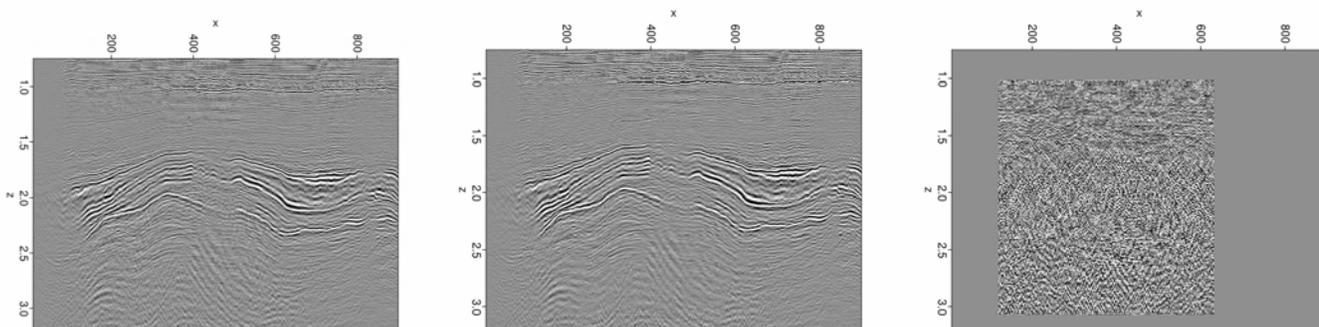


Figure 3: Real data example denoised by a simple thresholding on the coefficients after Contourlet and then Wavelet transforming the image for multiple angles. **Top left:** original stack. **Top right:** denoised stack (within the window). **Bottom middle:** predicted noise. Notice the improved SNR and removal of migration artifacts.

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] M. Do and M. Vetterli. *Beyond wavelets*, chapter Contourlets. Academic Press, 2002.
- [6] D. L. Donoho and I. M. Johnstone. Minimax estimation via wavelet shrinkage. *Annals of Statistics*, 26(3): 879–921, 1998.
- [7] 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.
- [8] F. J. Herrmann. Optimal seismic imaging with curvelets. In *Expanded Abstracts*, Tulsa, 2003. Soc. Expl. Geophys.
- [9] S. G. Mallat. *A wavelet tour of signal processing*. Academic Press, 1997.
- [10] D. Trad, T. Ulrych, and M. Sacchi. Latest views of the sparse radon transform. *Geophysics*, 68(1): 386–399, 2003.
- [11] C. Vogel. *Computational Methods for Inverse Problems*. SIAM, 2002.