LONDON2007

0000

Recent developments in curvelet-based seismic processing

Felix J. Herrmann (fherrmann@eos.ubc.ca)

Seismic Laboratory for Imaging and Modeling
Department of Earth and Ocean Sciences
The University of British Columbia

SUMMARY

In this paper, we present recent developments in nonlinear curvelet-based sparsity-promoting formulations of problems in the seismic data processing flow. We present our latest work on a parallel curvelet transform and recent work on a curvelet-regularized formulation for the focal transform, the prediction of multiples and the computation of the inverse data space. We show that the curvelet's wavefront detection capability and invariance under wave propagation lead to a formulation of these problems that is stable under noise and missing data.

OUTLINE

During this presentation an overview will be given of the latest developments concerning the application of the curvelet transform to seismic data processing. The following topics will be discussed extensively:

- the definition of a windowed parallel curvelet transform, which allows for the application of the curvelet transform to large data volumes (Thomson et al., 2006);
- combination of the curvelet transform with the (de)focal transform (= a data adaptive transform defined by a cross-convolution/correlation with the primaries) for the purpose of improved interpolation and multiple prediction (Berkhout and Verschuur, 2006)
- a curvelet-based formulation for the computation of the data inverse (Berkhout, 2006).

Parallel Windowed FDCT

Seismic data volumes are often too large to fit into the memory of single processor. At the expense of a low-frequency cutoff, a parallel curvelet transform can be constructed by applying a domain decomposition (see Fig. 1), implementing a partitioning of unity (the squares of the windows sum to 1). Consider equally sized windows with uniform overlap ϵ. The overlapping regions are tapered such that the energy of the system remains unity when points in the data set are duplicated at the overlaps. The tapering ensures that edges of the data goes smoothly to zero, eliminating potential edge artifacts.

Once the data is split into overlapping windows and tapered the fast discrete curvelet transform (FDCT, Candes et al., 2006; Hennenfent and Herrmann, 2006), is performed on each window separately. The shape of the taper function and the overlap of neighboring windows are both shown in Fig. 1. The data in this example is split into sixteen overlapping windows. The dashed lines in the image represent the edges of the overlapping windows, with the region between nearby parallel lines being shared amongst the two windows. The taper function is shown, where one can see the value going to zero at the window edges.

This entire process of windowing can be considered as a linear operator acting on a vector containing the reordered to-be-transformed data volumes. For the parallel transform, it is useful to distinguish between global vectors and transforms (=matrix-vector multiplication), which comprise the entire data volume, and local vectors and transforms that exist only on individual nodes. Our notation will reflect this distinction. For instance, the distributed transform and vector $\mathbf{C}$ and $\mathbf{x}$ are distinguished from the local transform and vector $\mathbf{C}_i$ and $\mathbf{x}_i$ that exist on a node indexed by i. Note that there will be cases when we want to apply a local transform $\mathbf{C}_i$ to the data on every node in a global vector $\mathbf{x}$. This involves a block diagonal global operator where each block consists of the particular $\mathbf{C}_i$ corresponding to a given $\mathbf{x}_i$, which is simply a subsection of $\mathbf{x}$. We denote the block diagonal global operator for a local operator as $\mathbf{[C}_i\mathbf{]}$.

An arbitrary global data vector is expanded into overlapping sections using the global windowing operator $\mathbf{W}$. The tapering operator $\mathbf{T}$ is then applied across the system. It is easy to see that the tapering operator is diagonal, and will have repeating patterns in blocks representing each node. These tapering operators $\mathbf{T}_i$ are local on each node. Once $\mathbf{W}$ and $\mathbf{T}$ have been applied, the linear operator $\mathbf{C}_i$ can be applied on each node.

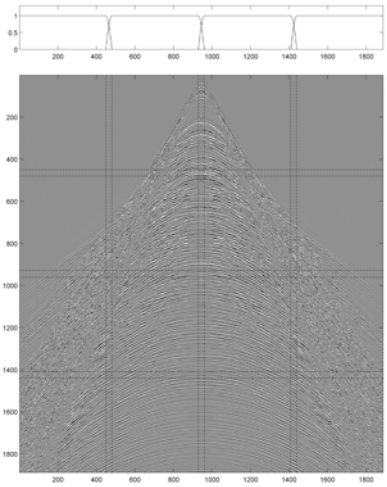


Figure 1: 2D synthetic seismic data with 30% missing vertical traces. Dotted lines represent borders of overlapping windows. Taper function is shown above for clarity.

Since $\mathbf{T}$ is diagonal, it is its own adjoint. For $\mathbf{W}$, the adjoint operation involves summing together overlapping regions. Since the forward operation was a "scatter," the adjoint becomes a "gather," where data that correspond to the same point are summed (Claerbout, 1992). In a parallel computing realization, this means that a given node sends its outer band to the nodes to which they belong, then gathers data related to its

inner band from those same neigbours, and sums them together. The combination of these processes satisfies the relation

$$\mathbf{W}^H\mathbf{T}^H\mathbf{T}\mathbf{W} = \mathbf{I}, \tag{1}$$

which ensures perfect reconstruction of the data after applying the operators followed by their adjoints, i.e. our parallel curvelet transform is a tight frame. This implies that energy is preserved through the entire process. So, a new transform, that we will call the Parallel Windowed FDCT (PWFDCT), can thus be defined. The parallel curvelet transform corresponds to a block-diagonal global operator that is defined by applying the FDCT, C_i, on each node independently after applying the global windowing and tapering operators $\mathbf{T}$ and $\mathbf{W}$. In other words,

$$\mathbf{C} := [\mathrm{C}_i]\mathbf{T}\mathbf{W}. \tag{2}$$

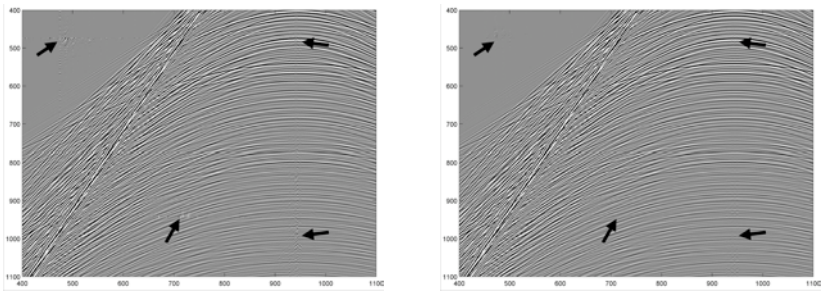
Since the local C is numerically tight, it follows that

$$\mathbf{C}^H\mathbf{C} = \mathbf{I}. \tag{3}$$

Other properties of C are similarly shared by $\mathbf{C}$. It is possible, then, to make use of the PWFDCT in existing algorithms that include the FDCT. It should be noted, though, that curvelets at very large scales (or, equivalently, low frequencies) are not represented in the same way they would be if a single FDCT were performed on the global data. Since the benefits of curvelets are mostly found at the finer scales (higher frequencies), the lack of large scale curvelet representation is not a problem. This makes a wide variety of algorithms capable of handling data sets much larger than otherwise possible. An example of a seismic data interpolation algorithm where the PWFDCT takes the place of the FDCT is included in Fig. 2. From the results presented in that figure one can conclude that algorithms, such as curvelet-based sparsity-promoting data recovery, can readily be scaled up by using our parallel windowed curvelet transform based on a partition of unity.

EXTENSIONS OF THE CURVELET TRANSFORM

So far, curvelets have mainly found their applications in the approximation of migration operators (Douma and de Hoop, 2006; Chauris, 2006), amplitude recovery (Herrmann et al., 2006) and seismic data interpolation (Herrmann and Hennenfent, 2007) and separation (Herrmann



(a) (b)

Figure 2: Interpolation output for **(a)** non-overlapping windows and for **(b)** overlapping, tapered windows with threshold correction. Arrows indicate locations of artifacts in (a) and show the improvement in (b).

et al., 2007). During these applications, the invariance and compressibility of curvelets is used. The curvelet properties can also be used to improve seismic data recovery with the focal transform (Berkhout and Verschuur, 2006), multiple prediction and the computation of data inverse (Berkhout, 2006). In this case, use is made of the property that curvelets remain relatively sparse on data that is focused, defocused or inverted. Focused data in this context corresponds to data stripped of one propagation path interacting with the surface, while defocused data has a single propagation path added, mapping primaries into first-order multiples. The latter corresponds to SRME multiple predicted (Verschuur and Berkhout, 1997), while the former corresponds to the recently introduced focal transform (Berkhout and Verschuur, 2006).

In all above described manifestations of seismic data, we are dealing with data that contains wavefronts whether these are primaries, multiples or the data inverse. Curvelets compress data with wavefronts and therefore serve as a regularization, where the primary operator, adjoint primary operator and data operator can stably be inverted with a curvelet sparsity norm. These different data manifestations can be calculated by solving of the following norm-one nonlinear program:

$$\mathbf{P}_\epsilon : \qquad \begin{cases} \widetilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{s.t.} \quad \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2 \leq \epsilon \\ \widetilde{\mathbf{f}} = \mathbf{S}^T \widetilde{\mathbf{x}} \end{cases} \tag{4}$$

in which $\mathbf{y}$ is the (incomplete) 'data', $\mathbf{A}$ the synthesis matrix and $\mathbf{S}^T$ is the inverse sparsity transform and ϵ, a noise-dependent tolerance level. This constrained optimization problem is solved to within ϵ.

Recovery from incomplete data by focusing

Combination of the non-adaptive curvelet transform with the data-adaptive *focal* transform (Berkhout and Verschuur, 2006) gives rise to an interesting extension, where data is focused by inverting the primary operator (= a multidimensional convolution with the primaries). In the focal domain, data is sparser and hence the curvelet coefficients are sparser, which leads to an improved recovery. Recovery with focusing is possible by defining the synthesis matrix as $\mathbf{A} := \mathbf{R}\mathbf{\Delta}\mathbf{P}\mathbf{C}^T$ and the inverse sparsity transform as $\mathbf{S}^T := \mathbf{\Delta}\mathbf{P}\mathbf{C}^T$. The operator $\mathbf{\Delta}\mathbf{P}$ stands for a multidimensional convolution with the primaries. The data vector $\mathbf{y}$ contains the incomplete measurements, $\mathbf{y} = \mathbf{R}\mathbf{d}$ with $\mathbf{R}$ the restriction matrix that selects the rows from the curvelet matrix that correspond to active traces. The solution yielded by $\mathbf{P}_\epsilon$ corresponds to a curvelet-regularized inverse of the primary operator. The estimated coefficients represent an estimate for the *focused* data. The synthesis operator defocuses the focused data back to data domain. The results as shown Fig 3 show as expected an improvement over interpolation without focusing.

Multiple prediction with defocusing

By defining the synthesis matrix by the adjoint of the primary operator, the solution of $\mathbf{P}_\epsilon$, with the data $\mathbf{y}$ containing the data including multiples, yields an estimate for the predicted multiples, i.e. by setting $\mathbf{S}^T = \mathbf{C}^T$. This estimate corresponds to the inverse of the adjoint of the primary operator, i.e., $\mathbf{A} := \mathbf{\Delta}\mathbf{P}^T\mathbf{C}^T$. As can be observed from Fig. 4, the frequency content of the predicted multiples is increased by this prediction through curvelet-regularized inversion.

Computation of the data inverse

The previous two examples showed that a curvelet regularization for the inversion of the primary operator and its inverse improves the stability and yields better results for the multiple prediction. Similarly, the computation of the data inverse can benefit from the same regularization by solving $\mathbf{P}_\epsilon$ for the synthesis matrix, $\mathbf{A} := \mathbf{P}\mathbf{C}^T$ with $\mathbf{P}$ the multidimensional convolution with the total data, and the 'data' vector, $\mathbf{y} = \boldsymbol{\delta}$, with $\boldsymbol{\delta}$ the discrete delta-function. As the example in Fig. 5 illustrates, the

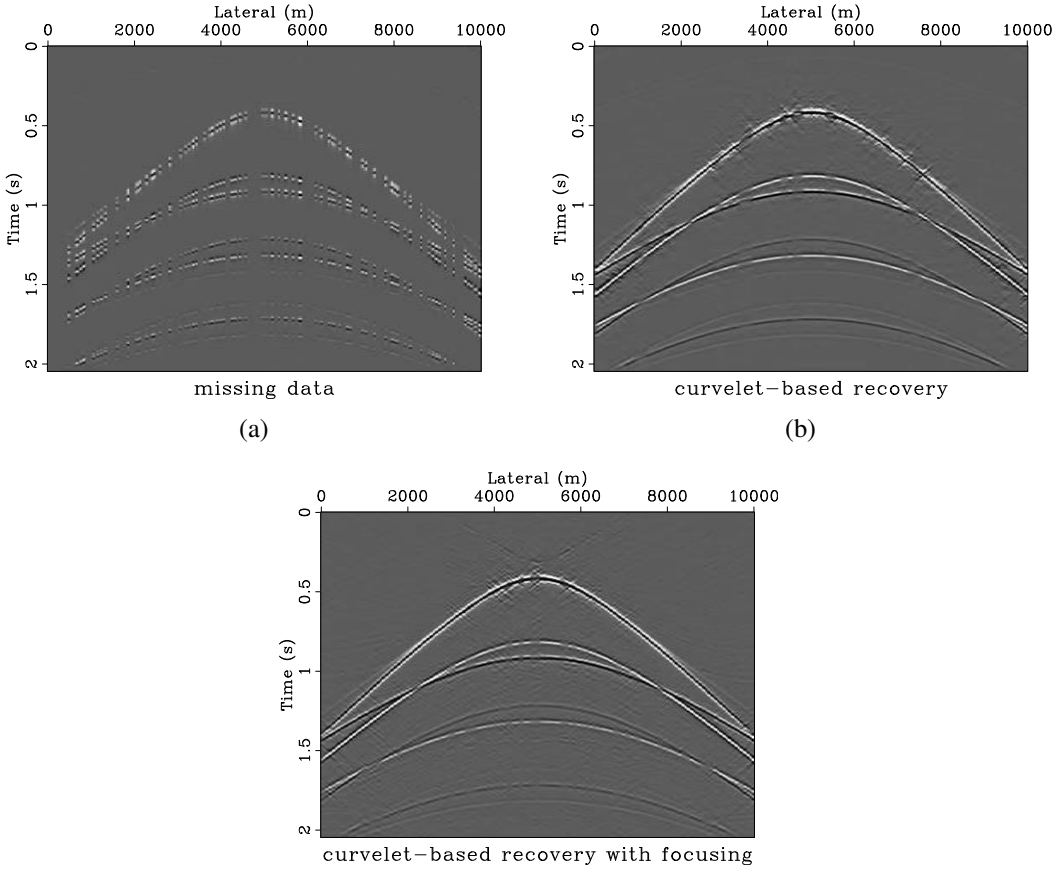


Figure 3: Comparison between curvelet-based recovery by sparsity-promoting inversion with and without focusing. **(a)** Randomly subsampled data with $80\,\%$ of the traces missing. **(b)** Curvelet-based recovery. **(c)** Curvelet-based recovery with focusing. Notice the significant improvement from the focusing with the primary operator.

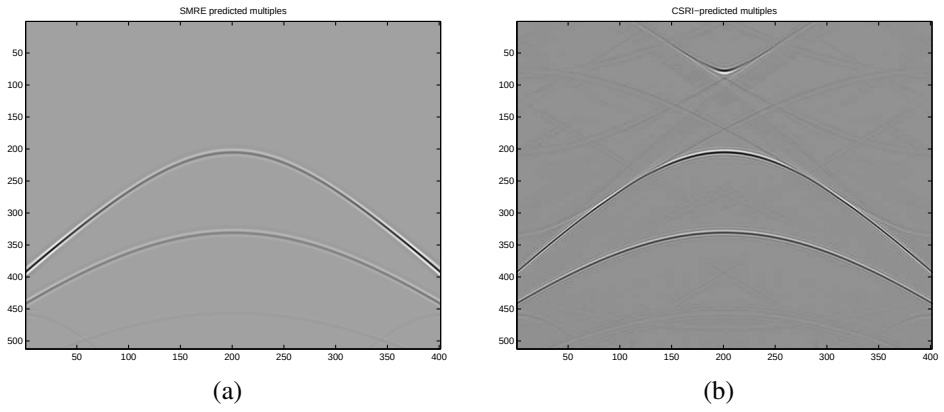


Figure 4: Comparison between convolution-based multiple prediction **(a)** and sparsity-based multiple prediction **(b)**.

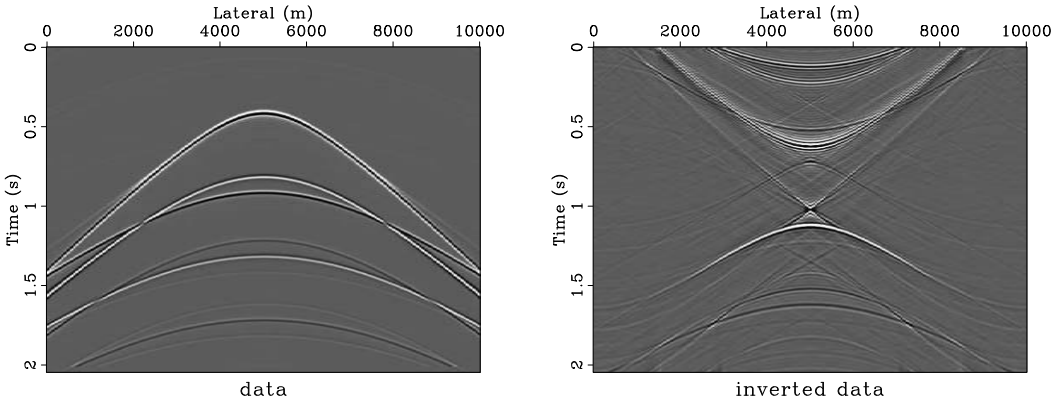


Figure 5: Computation of the data inverse. **(a)** shot record in the data space. **(b)** shot record in the inverse data space computed according to $\mathbf{P}_\epsilon$.

solution of $\mathbf{P}_\epsilon$ yields a stable estimate for the data inverse.

DISCUSSION AND CONCLUSIONS

Applications of the curvelet transform to seismic data processing bank on two favorable properties of curvelets, namely their ability to detect wavefronts and their approximate invariance under wave propagation. In this paper, we showed some exciting recent developments of this transform towards a large-scale parallelization and a combination with the physics of wave propagation. The successful application of curvelets, juxtaposed by sparsity-promoting inversion, opens a range of new perspectives on seismic data processing, wavefield extrapolation and imaging. Because of their singular wavefront detection capability, curvelets represent in our vision an ideal domain for future developments in exploration seismology.

ACKNOWLEDGMENTS:

The author would like to thank the authors of CurveLab for making their codes available, Darren Thomson for helping to implement the parallel curvelet transform and Eric Verschuur for his input. The examples presented in this paper were prepared with Madagascar (`rsf.sourceforge.net/`). SAGA and ExxonMobil are thanked for making test datasets available. This work was in part financially supported by the Natural Sciences and Engineering Research Council of Canada Discovery

Grant (22R81254) and Collaborative Research and Development Grant DNOISE (334810-05) of Felix J. Herrmann and was carried out as part of the SINBAD project with support, secured through ITF (the Industry Technology Facilitator), from the following organizations: BG Group, BP, Chevron, ExxonMobil and Shell.

REFERENCES

Berkhout, A. J., 2006, Seismic processing in the inverse data space: Geophysics, **71**.

Berkhout, A. J. and D. J. Verschuur, 2006, Focal transformation, an imaging concept for signal restoration and noise removal: Geophysics, **71**.

Candes, E. J., L. Demanet, D. L. Donoho, and L. Ying, 2006, Fast discrete curvelet transforms: SIAM Multiscale Model. Simul., **5**, 861–899.

Chauris, H., 2006, Seismic imaging in the curvelet domain and its implications for the curvelet design: Presented at the 76th Ann. Internat. Mtg., SEG, Soc. Expl. Geophys., Expanded abstracts.

Claerbout, J. F., 1992, Earth soundings analysis: Processing versus inversion: Blackwell Scientific publishing.

Douma, H. and M. de Hoop, 2006, Leading-order seismic imaging using curvelets: Presented at the 76th Ann. Internat. Mtg., SEG, Soc. Expl. Geophys., Expanded abstracts.

Hennenfent, G. and F. J. Herrmann, 2006, Seismic denoising with nonuniformly sampled curvelets: IEEE Comp. in Sci. and Eng., **8**, 16–25.

Herrmann, F. J., U. Boeniger, and D.-J. E. Verschuur, 2007, Nonlinear primary-multiple separation with directional curvelet frames: Geoph. J. Int. To appear.

Herrmann, F. J. and G. Hennenfent, 2007, Non-parametric seismic data recovery with curvelet frames. Submitted for publication.

Herrmann, F. J., P. P. Moghaddam, and C. Stolk, 2006, Sparsety- and continuity-promoting seismic imaging with curvelet frames. In revision.

Thomson, D., G. Hennenfent, H. Modzelewski, and F. Herrmann, 2006, A parallel windowed fast discrete curvelet transform applied to seismic processing: Presented at the SEG International Exposition and 76th Annual Meeting.

Verschuur, D. J. and A. J. Berkhout, 1997, Estimation of multiple scat-

tering by iterative inversion, part II: practical aspects and examples: Geophysics, **62**, 1596–1611.