

# Non-linear data continuation with redundant frames

F. Herrmann, and G. Hennenfent  
EOS, University of British Columbia, Canada

2005 CSEG National Convention



## Abstract

We propose an efficient iterative data interpolation method using continuity along reflectors in seismic images via curvelet and discrete cosine transforms. The curvelet transform is a new multiscale transform that provides sparse representations for images that comprise smooth objects separated by piece-wise smooth discontinuities (e.g. seismic images). The advantage of using curvelets is that these frames are sparse for high-frequency caustic-free solutions of the wave-equation. Since we are dealing with less than ideal data (e.g. bandwidth-limited), we complement the curvelet frames with the discrete cosine transform. The latter is motivated by the successful data continuation with the discrete Fourier transform. By choosing generic basis functions we circumvent the necessity to make parametric assumptions (e.g. through linear/parabolic Radon or demigration) regarding the shape of events in seismic data. Synthetic and real data examples demonstrate that our algorithm provides interpolated traces that accurately reproduce the wavelet shape as well as the AVO behavior along events in shot gathers.

## Linear data interpolation

Data continuation has been an important topic in seismic processing, imaging and inversion. For instance, our ability to accurately and artifact-free image or predict multiples, as part of surface related multiple elimination, depends for a large part on the availability of densely and equidistantly sampled data. Unfortunately, constraints on marine and land acquisitions preclude such a dense sampling and several efforts have been made to solve issues related to aliasing and missing traces.

Typically, data continuation approaches involve some sort of a variational problem that aims to jointly minimize the  $L^2$ -mismatch, between measured and interpolated data, and an independent penalty functional (regularization term) on the interpolated data, which we will call the model. For quadratic penalty functionals, the solution for the variational problem [see e.g. 3]

$$\hat{m} = \arg \min_m \frac{1}{2} \|d - Pm\|_2^2 + \lambda \|Dm\|_2^2 \quad (1)$$

can be written explicitly as

$$\hat{m} = (P^*P + \lambda D^*D)^{-1} P^* d. \quad (2)$$

In this expression,  $d$  is the (weighted) acquired vector with the missing data;  $m$  the regularly sampled data and  $P$  the picking matrix, which puts to zero the missing data samples/traces. The symbol  $*$  denotes the adjoint. Since the picking matrix is not invertible, Eq. 1 contains an additional penalty term, which regularizes the inverse problem of data interpolation. The  $\lambda$ -weighted penalty term typically imposes smoothness on the model with  $D$  some type of sharpening operator such as a (fractional) Laplacian, i.e.  $D \bullet = \Delta^\alpha \bullet$  with  $\alpha \geq 0$ . For a specific choice of regularization, Eq. 1 corresponds to kriging or spline interpolation.

## Non-linear sparseness-constrained interpolation

Despite its elegance, Eq. 1 has a number of serious drawbacks amongst which (i) the inherent non-stationarity of seismic data, which makes it difficult to define the appropriate sharpening operator  $D$  defining the covariance operator  $C \bullet = D^*D \bullet$ ; (ii) the notion that wavefronts are typically oscillatory in the normal direction to the wavefront and smooth in the direction along the wavefronts, which is not trivially accounted for; (iii) wavefronts are sparse, i.e. seismic data can be thought of as a superposition of a limited number of events, e.g. plane waves or solutions to some sort of modeling operator such as the adjoint of the migration operator.

Relatively recently these observations have led to a non-linear model/transform-driven reformulation of Eq. 1 given by

$$\hat{m} = \arg \min_m \frac{1}{2} \|P(d - Lm)\|_2^2 + \lambda \|m\|_p^p \text{ with } 1 \leq p < 2 \quad (3)$$

In this expression,  $L$  stands for either a *modeling* operator – such as the adjoint of the (linear/parabolic) Radon, the apex-shifted Radon or the demigration operator [15,16] or for a generic orthogonal basis-function decomposition such as the Fast Fourier Transform [5,11,12]. As long as events in the data roughly behave as the columns in  $L$ , good interpolation results have been obtained by solving the above system (e.g. with the method of Iterated Re-weighted Least Squares IRLS [9]). Unfortunately, events in less-than-ideal data are not likely to behave exactly according to the columns of  $L$ . Moreover, the singular vectors associated with the solution of Eq. 3 tend to spread globally even though the  $l^1$ -norm invokes 'spikyness' in these vectors and the solution. To limit the dependence on assumptions regarding the shape of the events, we propose to use a combination of fast curvelet (FCT) and discrete cosine transforms (DCT) [2,6], which consist of both localized multi-scale & angular elements that can be interpreted as local plane waves which are smooth in one direction and oscillatory in the other as well as Fourier-like monochromatic plane waves. The FCT constitutes a moderate redundant frame while the DCT is an orthogonal basis.

## Non-linear sparseness-constrained interpolation with curvelet frames

The method presented in this paper is similar to the approach of Eq. 3. Major difference, however, is that we replace the modeling or Fourier/Radon transform operators with a generic frame composition defined by the augmentation of the FCT & DCT [6]. Compared to the model-driven approaches, our approach has the advantage that (i) curvelet frames are known to be sparse for seismic data (they obtain close to optimal non-linear approximation rates for seismic data); (ii) discrete Fourier transforms have been applied successfully in data continuation [11,5]; (iii) no assumptions are made regarding the type of events present in the data. For example, events no longer need to be close to parabolic or linear as is the case for parabolic and linear Radon. The events can lie on arbitrary but piece-wise smooth (twice-differentiable) curves [13].

Before going into detail on how to apply the method, let us first provide more details on curvelet properties. We then proceed by showing how to compute the  $l^1$ -minimization by iterative thresholding. We conclude by applying our method to synthetic data.

### Why the FCT & DCT

We combine DCT & FCT to get the best of Fourier-like behavior of waves [11,5] and the high non-linear approximation rate of curvelets for wavefronts. Without being all inclusive (see [2] for details), curvelets obtain such a rate because they are:

- *multi-scale*, i.e. they live in different dyadic corona in the FK-domain
- *multi-directional*, i.e. they live on wedges within these corona.
- *anisotropic*, i.e. they obey the following scaling law  $width \propto length^2$ .
- *directional selective* with  $\#orientations \propto scale^{-1/2}$ .
- *local* both in  $(x,t)$  and  $(k,f)$ .
- a numerically tight frame with moderate redundancy and an explicit construction exists for the adjoint that equals the pseudo inverse:  $C^T \bullet = C^+$ , yielding  $\|f\|_2^2 = c^T c$  with  $c = Cf$  the curvelet coefficients of  $f$  and  $C$  the curvelet transform.

For more detail on curvelet frames and DCT refer to [7,8,10].

### Sparsity-enhancing frame expansions by iterative thresholding

For orthogonal basis functions, soft thresholding on noisy coefficients  $y$  [4,10]

$$\hat{x} = S_\lambda^s(y) \quad (4)$$

solves component-wise the following variational problem

$$\hat{x} = \arg \min_x \frac{1}{2} \|y - x\|_2^2 + \|x\|_1. \quad (5)$$

In our approach, we work with the following augmented system,

$$T = \begin{bmatrix} C^* & F^* \end{bmatrix} \quad \text{with} \quad c = \begin{bmatrix} c_1 & c_2 \end{bmatrix}^*, \quad (6)$$

where  $c$  is the coefficient vector consisting of the FCT and DCT coefficients. Because the system  $T$  is redundant and hence non-orthogonal Eq. 4 is no longer equivalent to Eq. 5. Following recent work by [4,14,6], we replace Eq. 4 by an iterative block thresholding procedure. At each iteration, evaluation of

$$c_j^m = S_{\lambda_m}^s \left( c_j^{m-1} + T_j^* \left( d - \sum_i T_i c_i^{m-1} \right) \right) \quad \text{for} \quad j = 1, 2 \quad (7)$$

yields an approximate estimate for the coefficient vector. For large enough number of iterations, the solution from Eq. 7 converges to the following variational problem on the coefficients

$$\hat{c} = \arg \min_c \frac{1}{2} \|d - Tc\|_2^2 + \lambda \|c\|_1 \quad (8)$$

In Eq. 7,  $d$  is the possibly noisy data vector,  $T^*$  the combined FCT & DCT, its component-wise (pseudo)-inverse and  $c^m$  the combined estimated coefficient vector after  $m$  iterations with possibly decreasing thresholds  $\lambda_m$ . Eq. 7 converges to a sparse set of coefficients that reconstruct the image [4,14,6]. By setting the  $\lambda_m$ 's proportional to the noise-level, the iterations in Eq. 7 yield the denoised coefficients from which the denoised data can be reconstructed by taking the inverse FCT & DCT.

### Data continuation with redundant frames

As shown independently within both the seismic (see e.g. [11,12,16]) and image (see e.g. [6]) processing communities, Eq. 8 can simply be adapted to the situation of missing data by including the picking operator (see also [17])

$$\hat{c} = \arg \min_c \frac{1}{2} \|P(d - Tc)\|_2^2 + \lambda \|c\|_1, \quad (9)$$

which corresponds to replacing Eq. 7 by

$$c_j^m = S_{\lambda_m}^s \left( c_j^{m-1} + T_j^* P \left( d - \sum_i T_i c_i^{m-1} \right) \right) \quad \text{for } j = 1, 2 \quad (10)$$

This expression is used to generate the examples of the next section by setting the maximal number of iterations to  $m=100$  while decreasing  $\lambda_m$ .

## Examples

Above iterative method is applied to synthetic as well as real data. Results for a synthetic shot-record, using the iteration defined in Eq. 10 for 100 iterations with  $\lambda$  decreasing from  $\lambda_1=10^{-1}$  to  $\lambda_{100}=10^{-4}$  (for data with a norm set to unity), are listed in Fig. 1. Our method is clearly able to achieve good results on synthetic data. The shape of the wavelet is accurately recovered as well as the AVO behavior (Fig. 1 b & c). As opposed to the synthetic where the contribution of the DCT is negligible, the DCT plays an important role in the continuation of real data in which there are more oscillations. Including the DCT allows us to capture the fine texture of the oscillations explaining our encouraging results as depicted in Fig. 2.

## Discussion and conclusions

The success of our approach can heuristically be explained by arguing that curvelets act as an “unconditional basis” for ideal seismic data (both primaries and multiples), which we consider to be generated by the (repeated) action of a scattering operator. Whenever a basis is unconditional, one finds that (i) the data’s covariance is near diagonal in the basis and (ii) the norm (say  $\ell^p$ ) shrinks when shrinking the coefficients. The argument that curvelets can be considered to be close to an unconditional basis for seismic data derives from the work on function spaces by [13] and from Theorem 1.1 of [1] which states that Green’s functions are nearly diagonalized by curvelets. Consequently, seismic data can be represented by a sparse coefficient vector for a dictionary rendering our data-continuation method to be (i) effective for caustic-free seismic data with events on arbitrary curves; (ii) robust under additive possibly colored noise (see other contribution by the author to these Proceedings). Results for the synthetic gather confirm the beneficial properties since the contribution of the DCT is minimal in this case. For real data, however, the situation is different because the DCT contributes to the solution. This contribution can be explained by the more oscillatory nature of real data captured by the DCT. From the real data example, we can conclude that the combination of the FCT and DCT provides an adequate frame for seismic data continuation. Our method can relatively easily be generalized to 3-D and speeded up by using more intelligent  $\ell^p$ -solvers. Main challenge lies in extending the method to cases where there are caustics.

## Acknowledgments

The authors would like to thank the authors of WaveLab and the Digital Curvelet Transform (Candes, Donoho, Demanet and Ying). This work was carried out as part of the SINBAD project with financial support, secured through ITF (the Industry Technology Facilitator), from the following organizations: BG Group, BP, ExxonMobil and SHELL. Additional funding came from the NSERC Discovery Grant 22R81254.

## References

- [1] E. J. Candes and L. Demanet. The curvelet representation of wave propagators is optimally sparse. 2004. Preprint: <http://www.acm.caltech.edu/demanet>.
- [2] E. J. Candes and F. Guo. New multiscale transforms, minimum total variation synthesis: Applications to edge-preserving image reconstruction. *Signal Processing*, pages 1519–1543, 2002.
- [3] J. Claerbout. *Earth soundings analysis: Processing versus inversion*. Blackwell Scientific publishing, 1992.
- [4] I. Daubechies, M. Deffrise, and C. de Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraints. *CPAM*, pages 1413–1457, 2005.
- [5] A. Duijndam, A. Volker, and P. Zwartjes. Reconstruction as efficient alternative for least-squares migration, pages 1012– 1015. *Soc. of Expl. Geophys.*, 2000.
- [6] M. Elad, J. Starck, P. Querre, and D. Donoho. Simultaneous Cartoon and Texture Image Inpainting using Morphological Component Analysis (MCA). 2005.
- [7] F.J. Herrmann and P. Moghaddam. Sparseness- and continuity-constrained seismic imaging with curvelet frames. 2005.
- [8] F.J. Herrmann and P. Moghaddam. Curvelet-based non-linear adaptive subtraction with sparseness constraints. In *Expanded Abstracts*, Tulsa, 2004. *Soc. Expl. Geophys.*
- [9] L. Karlovitz. Construction of nearest points in the  $p$ ,  $p$  even and 1 norms. *J. Approx. Theory*, 3, 1970.
- [10] S. G. Mallat. *A wavelet tour of signal processing*. Academic Press, 1997.
- [11] M. Sacchi and T. Ulrych. Estimation of the discrete fourier transform, a linear inversion approach. *Geophysics*, 61(04):1128– 1136, 1996.
- [12] M. Schonewille and A. Duijndam. Parabolic radon transform, sampling and efficiency. volume 66, pages 667–678. *Soc. of Expl. Geophys.*, 2001.
- [13] H. Smith. A Hardy space for fourier integral operators. *J. Geom. Anal.*, 7, 1997.
- [14] J. L. Starck, M. Elad, and D. Donoho. Redundant multiscale transforms and their application to morphological component separation. *Advances in Imaging and Electron Physics*, 132, 2004.
- [15] D. Trad, T. Ulrych, and M. Sacchi. Latest views of the sparse radon transform. *Geophysics*, 68(1):386–399, 2003.
- [16] D. O. Trad. Interpolation and multiple attenuation with migration operators. *Geophysics*, 68(6):2043–2054, 2003.
- [17] F.J. Herrmann. Robust curvelet-domain data continuation with sparseness constraints. 2005. (*submitted to EAGE 2005*)

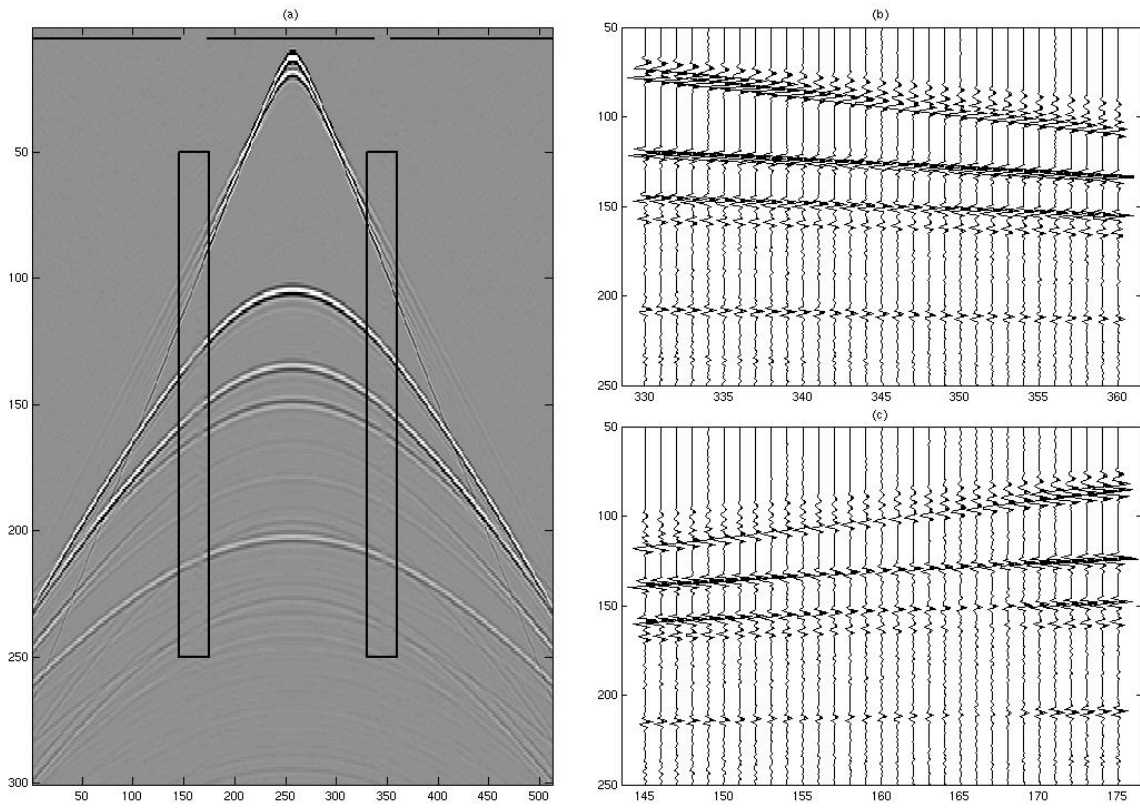


Figure 1: Example of non-linear data continuation with FCT & DCT applied to a synthetic shot record. **(a)** Interpolated data (holes along the horizontal bar on the top show missing traces location); **(b)** Wiggle plot of the right box, and **(c)** wiggle plot of the left box.

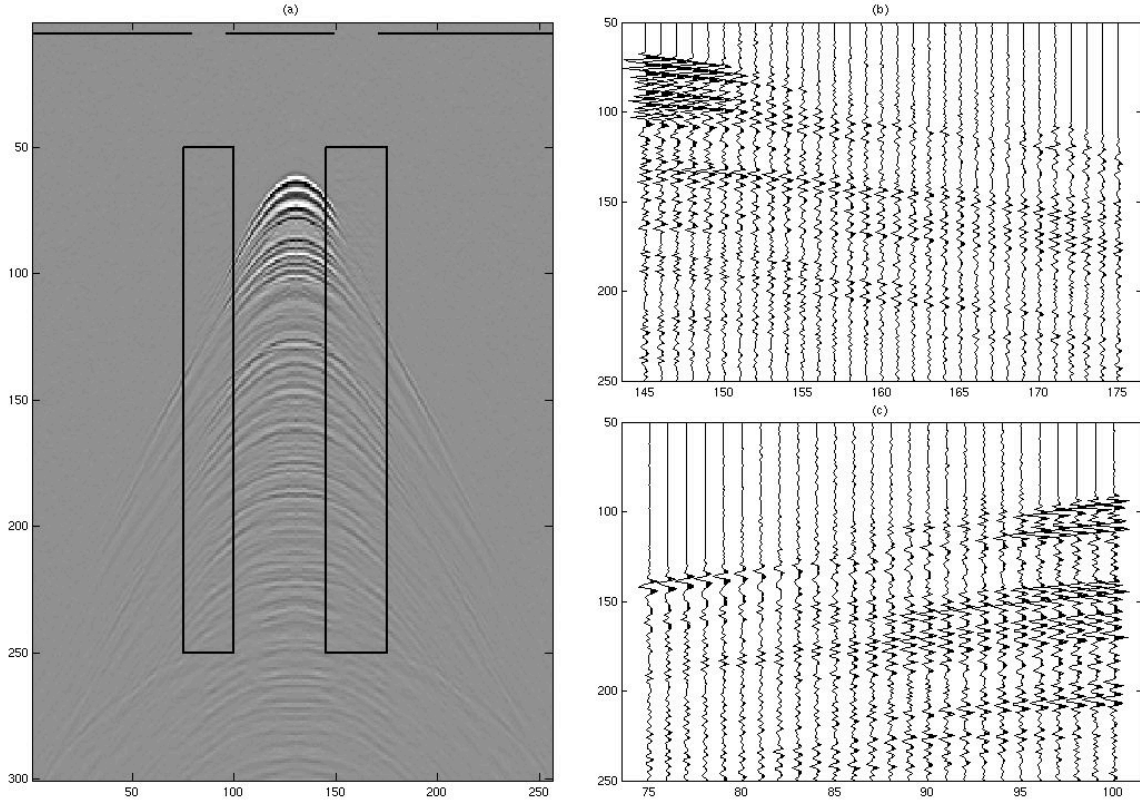


Figure 2: Example of non-linear data continuation with FCT & DCT applied to a real shot record with missing data. **(a)** Interpolated data (holes along the horizontal bar on the top show missing traces location); **(b)** Wiggle plot of the right box, and **(c)** wiggle plot of the left box.