Data Reconstruction by Generalized Deconvolution

M. D. Sacchi, Department of Physics, University of Alberta,
D. J. Verschuur, and P. M. Zwartjes, Delft University of Technology

Summary

The Radon transform is a powerful technique that has been primarily used to remove coherent and incoherent noise from seismic records. In addition, it has a long history in image processing as a tool for feature extraction. In exploration seismology, however, a major shortcoming is the requirement of simple integration paths that often do not quite well approximate the spatio-temporal signature of real seismic events.

The purpose of this presentation is twofold. First, we introduce a generalized convolution that allows us to compute Radon transforms with any class of integration path. Secondly, we present a strategy to utilize generalized convolution to represent seismic data via a Local Wavefield Decomposition (LWD). This is a parametric decomposition of the data in terms of local wavefield operators. The latter can be used to filter undesired events and interpolate aliased data.

Introduction

A chief problem in seismic data processing is the filtering of unwanted events like ground roll and multiples. Methods to deal with this problem often exploit move-out or curvature differences between offending events and the events one would like to preserve (primaries). In particular, removal of multiples based on move-out discrimination can be attained via parabolic and hyperbolic Radon transforms. In the parabolic transform, seismic data after normal-moveout correction is assumed to be composed of a superposition of parabolas; in the second case, data are assumed to be a superposition of hyperbolas.

Methods exists to enhance the resolution of both the hyperbolic Radon transform (Thorson and Claerbout, 1985) and the parabolic Radon transform (Sacchi and Ulrych, 1995). In both cases, the operator capable of inverting the Radon transform is constructed in such a way that the Radon panel exhibits minimum entropy or maximum sparseness (synonymous used to describe a distribution of isolated events in the Radon panel). The sparseness assumption might not be optimal when there is a mismatch between the integration path of the Radon operator and the spatio-temporal signature of the seismic event. Amplitude variation with offset can further complicate the problem, as described by Spagnolini (1994). One solution to this problem is to design operators that accurately reproduce both the kinematic and amplitude signature of the data. An example of the latter is the Focal transformation recently introduced by Berkhout et al. (2004). An alternative solution entails adopting local wavefield operators (LWO) and therefore, achieve a match between operator and data for only a small data aperture.

Local Wavefield Decomposition (LWD)

We start by defining a Local Wavefield Operator (LWO) via the following template:

$$\hat{b}(\omega, x, p) = \hat{s}(\omega)\, h(x)\, e^{-i\,\omega \phi(p,x)}\,, \quad -a \leq x \leq a\,. \tag{1}$$

In this expression a defines the operator half-aperture, $h(x)$ is spatial taper and, $\hat{s}(\omega)$ the Fourier transform of the wavelet. Ideally, we choose $\hat{s}(\omega)$ as close as possible to the seismic wavelet embedded in the data. The parametric integration path $\phi(p,x)$ defines the kinematics of the local wavefield. For instance, $\phi(p,x) = p\,x$ defines a local linear Radon operator of ray parameter p. Similarly, we could have chosen a parabolic template $p\,x^2$. We now define the LWO as the inverse Fourier transform of equation (1),

$$b(t, x, p) = \mathcal{F}^{-1}[\hat{b}(\omega, x, p)]\,. \tag{2}$$

It is clear that $b(t,x,p)$ is a small compact operator, a scaled and shifted version of it can be written as

$$A\, b(t - t_0, x - x_0, p)\,. \tag{3}$$

The coefficient A corresponds to the amplitude of a single LWO shifted in time and space. Eq. (3) can be generalized to a superposition of LWOs distributed in the (x_0, t_0) data plane:

$$\sum_{t_0} \sum_{x_0} f(t_0, x_0)\, b(t - t_0, x - x_0, p)\,. \tag{4}$$

Now, $f(t_0, x_0)$ can be interpreted as a 2D "shaping filter" that shapes one LWO into the desired 2D signal (the data). It is important to stress that seismic data, in general, is composed of a superposition of events that cannot be described by a single LWO, therefore, we generalize the convolution sum in equation (4) and obtain the following expression:

$$d(x,t) = \sum_{p} \sum_{t_0} \sum_{x_0} f(t_0, x_0)\, b(t - t_0, x - x_0, p) \tag{5}$$

or, in matrix form: $\mathbf{D} = \sum_p \mathbf{F}(p) \otimes \mathbf{B}(p)$. Our goal, now, is to find a procedure to estimate $\mathbf{F}(p)$ from $\mathbf{D}$. Equation (5) defines a *generalized convolution* (Granlund and Knutsson, 1995) that represents a decomposition of the seismic data in terms of temporal and spatially invariant kernels. The generalized convolution given by

Eq. (5) defines the transformation from the Local Wavefield Decomposition domain ($\mathbf{F}(p)$) to the data space ($\mathbf{D}$). We will denote such a transformation via the expression: $\mathbf{D} = \mathcal{L}[\mathbf{F}(p)]$.

Inversion of $\mathcal{L}$

Given the data $\mathbf{D}$, we now need a procedure to transform the data to the LWD domain; a simple mapping from data space to model space can be implemented via the adjoint transform $\hat{\mathbf{F}}(p) = \mathcal{L}^*[\mathbf{D}]$. From this perspective, $\mathcal{L}^*$ represents a generalized correlation operator. However, like in Radon processing, rather than using the adjoint operator $\mathcal{L}^*$ we prefer to estimate $\mathbf{F}$ from the data via inversion, a process that is equivalent to generalized deconvolution. In other words, we will minimize the following cost function

$$\mathcal{J} = ||\mathbf{W}(\mathbf{D} - \sum_p \mathbf{F}(p) \otimes \mathbf{B}(p))||_2^2 + \mathcal{R}(\mathbf{F})\,. \qquad (6)$$

The first term is the data misfit; the second is the regularization term. The matrix $\mathbf{W}$ is the sampling matrix required to process aliased data. In other words, the spatial sampling of the unaliased operators $\mathbf{B}(p)$ differs from the spatial sampling of the aliased data. This optimization problem can be tackled via the method of conjugate gradients. In our current numerical implementation, generalized convolution and its adjoint are implemented *on the fly* by means of a 2D FFT.

Examples

In Fig. (1) we portray 15 LWOs computed on a grid of 41×279 (space - time) samples. The spatial sampling of the data and operator is $\Delta x = 53$ m. The LWOs were parametrized by linear Radon operators with dips in the range $p_0 = -4.75 \times 10^{-4}$ s/m to $p_{15} = 0$ s/m. These operators are used to compute the LWD ($\mathbf{F}(p)$) of the data portrayed in Fig. 2 (first panel). The operator $\mathbf{F}(p)$ is used to synthesize individual data modes $\mathbf{D}(p_k) = \mathbf{F}(p_k) \otimes \mathbf{B}(p_k)$, $k = 1, 15$. The data modes, in conjunction with the full and partial data reconstructions are also provided in Fig. 2. Notice, that the partial reconstruction is created by summing the modes $k = 13, 14, 15$. The latter is a good representation of the hyperbolic event. A similar decomposition could have been achieved via the hybrid Radon transform (Trad et al, 2001). However, the hybrid Radon transform defines a basis of linear and hyperbolic events that are defined on the complete data aperture rather than a compact basis defined on a sub-aperture of the data like in the LWD.

In Fig. 3 we present an application of the LWD to interpolation beyond aliasing. In this case the true unaliased marine shot gather is decimated and then reconstructed using LWD. In other words, aliased data is used to retrieved unaliased operators $\mathbf{F}(p)$ that are used to reconstruct the data on a properly sampled grid. The spatial sampling interval is $\Delta x = 53.34$ m, whereas the LWO is constructed with a $\Delta x/2$ sampling interval. Fig. 4 displays the $f - k$ spectra of the original data (a), the decimated data (b), the reconstructed data (c) and the interpolation error (d).

Future directions and a few words on *Learning*

At this point it is important to stress that one could have adopted a non-parametric representation of the LWOs. For instance, the LWOs can be estimated from the data as part of a *learning* algorithm (Olshausen and Field, 1996). This concept could lead to a totally data driven noise attenuation process. We first estimate the LWT on individual data sets (e.g., shot records), this stage can be designated as the *local stage*. Then, a *global stage* can be used to update the LWOs by minimizing the cost function J averaged over many shot records. A schematic algorithm will look like:

1) For each individual record estimate the LWD ($\mathbf{F}$(p)) with inital parametric initial

2) Estimate $\Delta\mathbf{B}(p)$ such that $< J >$ (averaged cost) is minimized

3) Update the LWOs, $\mathbf{B}(p) + \Delta\mathbf{B}(p) \leftarrow \mathbf{B}(p)$

4) Re-start the local stage.

The key point is to achieve sparsity in the generalized convolver by finding optimal LWOs directly from the data.

Conclusions

We have presented a generalized convolver that allows us to represent seismic data in terms of a Local Wavefield Decomposition. The ideas presented in this paper have numerous applications: random and coherent (aliased) noise attenuation, interpolation beyond aliasing, and wavefield separation.

References

Berkhout, A. and Verschuur, D., and Romijn, R., 2004, Reconstruction of seismic data using the Focal Transformation, 74th Ann. Internat. Mtg.: Soc. of Exploration Geophys., submitted.

Granlund . G. H. and Knutsson, 1995, Signal Processing for Computer Vision: Kluwer Academic Publishers.

Olshausen, B. A., and Field, D. J., 1996, Emergence of simple-cell receptive field properties by learning a sparse code for natural images: Nature, 381, 607-609.

Sacchi, M. D. and Ulrych, T. J., 1995, High-resolution velocity gathers and offset space reconstruction: Geophysics, Soc. of Expl. Geophys., 60, 1169-1177.

Spagnolini, U., 1994, Compound events decomposition and the interaction between AVO and velocity information: Geophys. Prosp., Eur. Assn. Geosci. Eng., 42, 241-259.

Thorson, J. R. and Claerbout, J. F., 1985, Velocity stack and slant stochastic inversion: Geophysics, Soc. of Expl. Geophys., 50, 2727-2741.

Trad, D., Sacchi, M. D. and Ulrych, T. J., 2001, A hybrid linear-hyperbolic Radon transform: Jornal of Seismic Exploration, 9, 303-318.

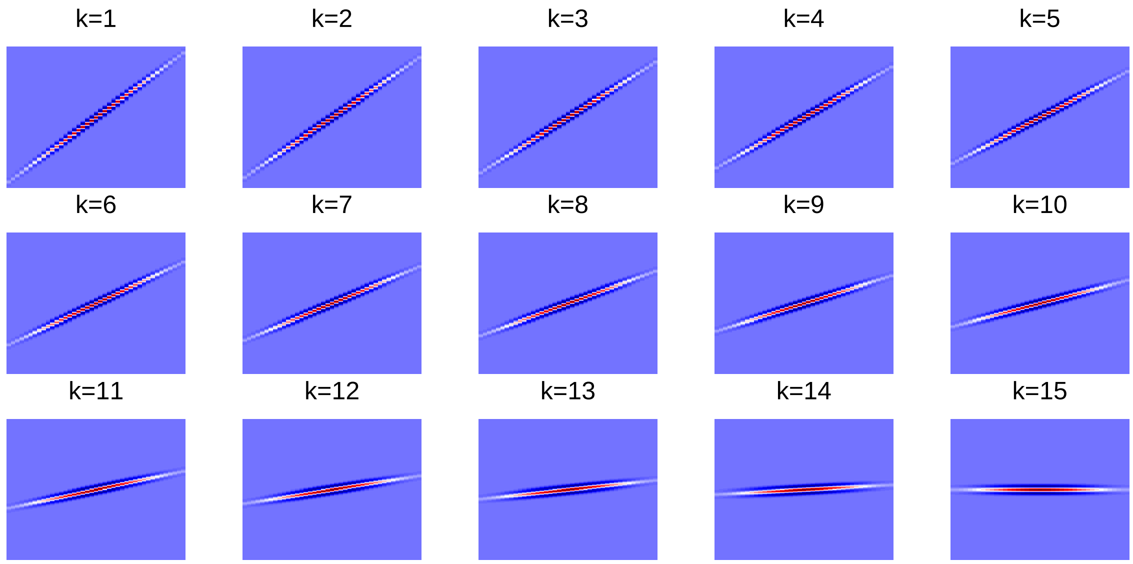

Fig. 1: Local wavefield operators used to decompose the data in Fig. 2. The operator size is 41×278 samples (space-time) ($\Delta t = 4$ msec $\Delta x = 53$ m). Notice that the scale of this figure is different that the scale of Fig. 2.

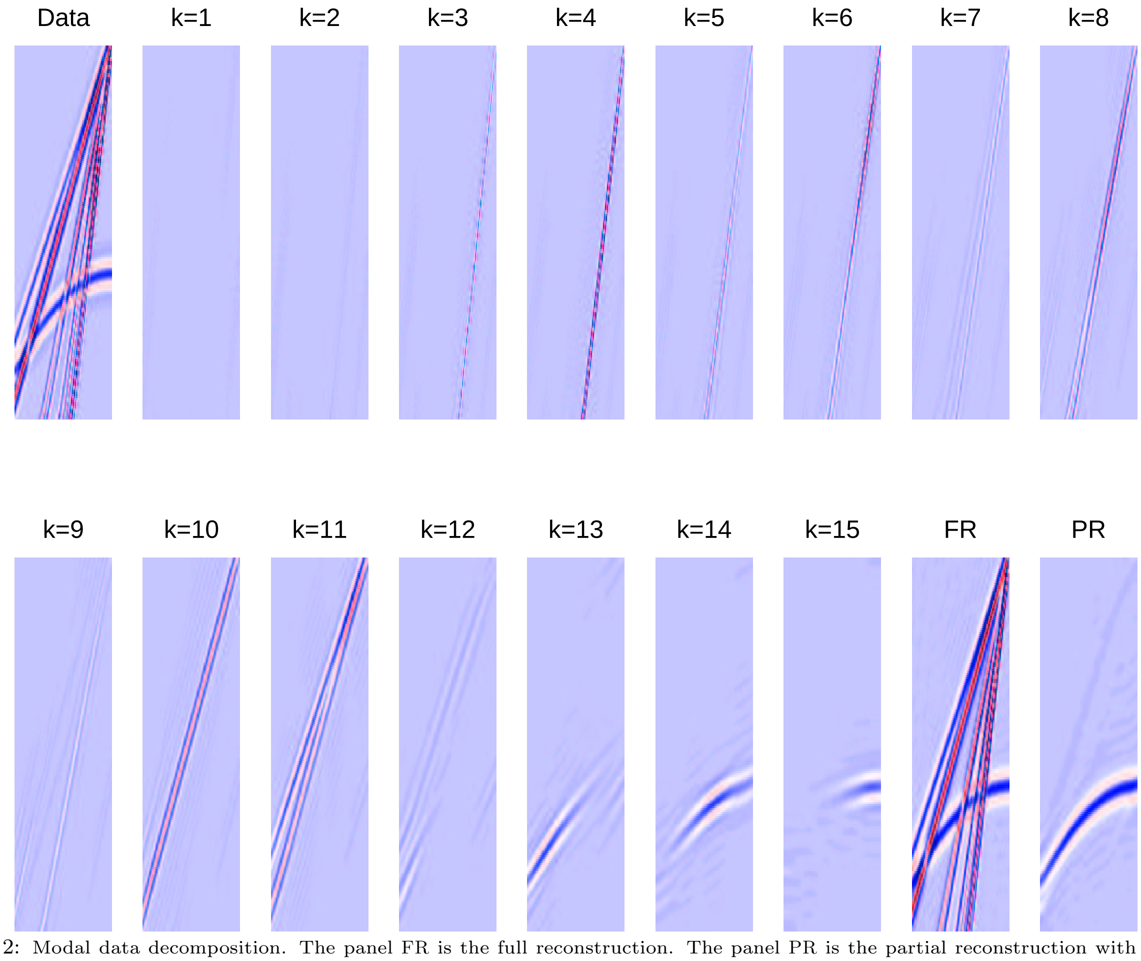

Fig. 2: Modal data decomposition. The panel FR is the full reconstruction. The panel PR is the partial reconstruction with modes $p_k, k = 13, 14, 14$.

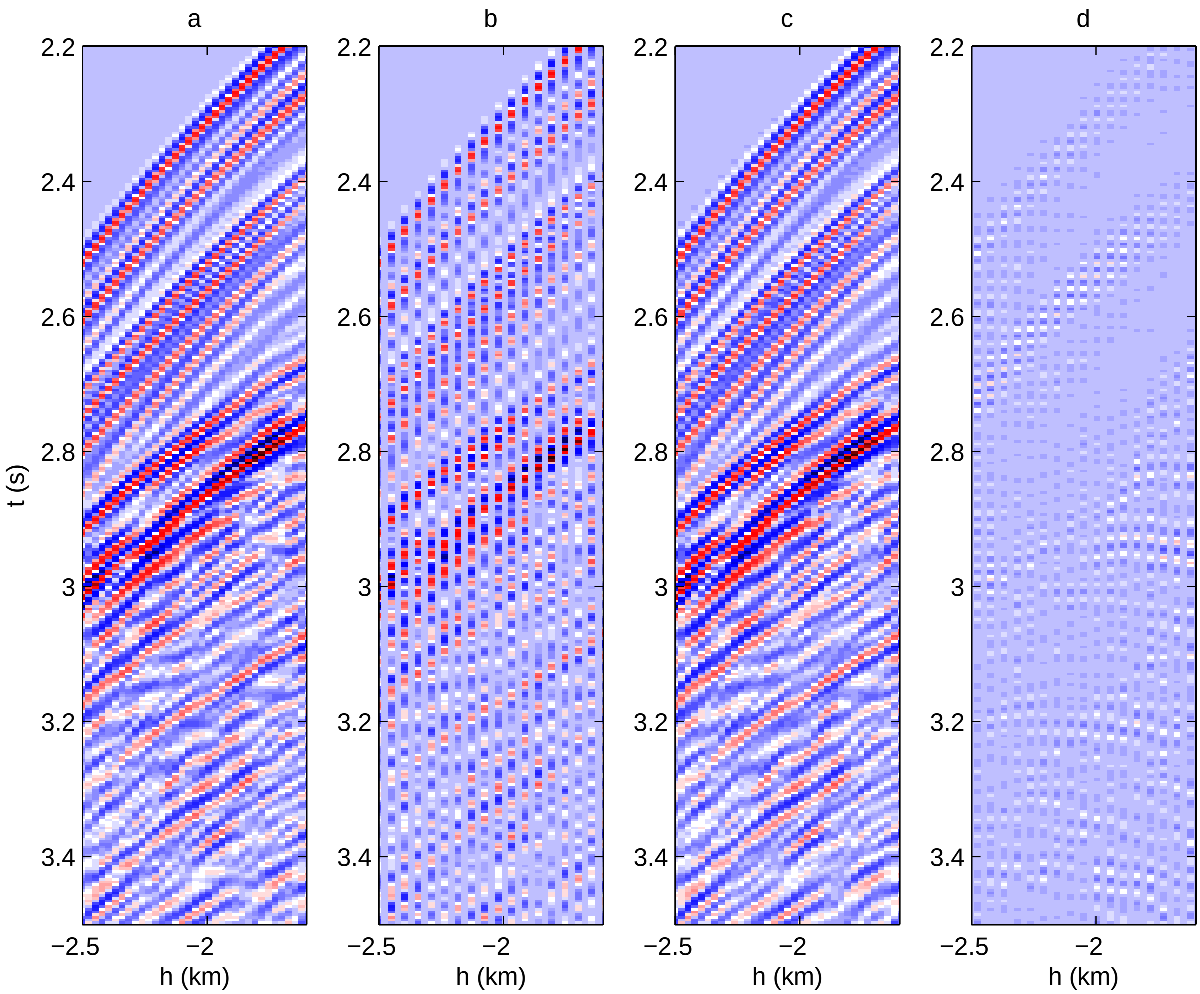

Fig. 3: (a) Gulf of Mexico marine shot record. (b) Decimated shot gather. (c) Reconstructed data. (d) Reconstruction error. Only a portion of the shot record is displayed.

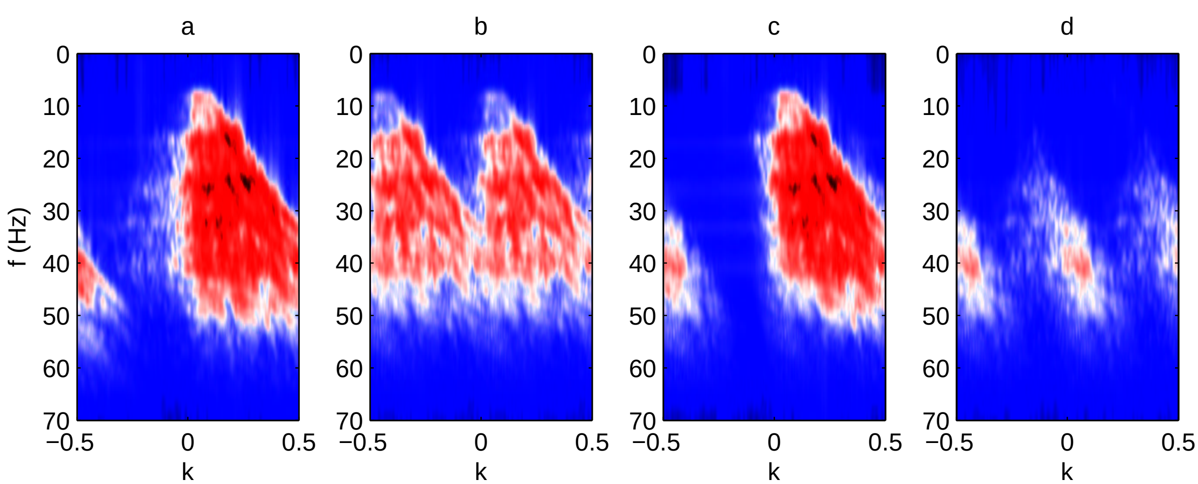

Fig. 4: Amplitude spectra. (a) Original Gulf of Mexico marine shot record. (b) Decimated shot gather. (c) Reconstructed data. (d) Reconstruction error.