Apex Shifted Radon transform

Daniel Trad*
Veritas DGC Inc., Suite 2200, 715 Fifth Av., Calgary, AB, T2P 5A2
dtrad@veritasdgc.com

ABSTRACT
The Apex Shifted Radon transform (ASRT) is an extension of the standard hyperbolic RT, with hyperbolic basis functions located at every point of a data gather. The mathematical description of such an operator is similar to the kinematic post stack time migration equation, with the horizontal coordinate being not midpoint but offset. This implementation of the transform uses a least squares conjugate gradient algorithm with a sparseness constraint with two different possibilities for the operators, one in the time offset domain and the other in the frequency wavenumber domain (Stolt operator). The frequency wavenumber domain operator is very efficient, not much more expensive in computation time than a sparse parabolic RT and much faster than a standard hyperbolic RT.

The sparseness constraint in the time-offset domain is essential for filling gaps and for interpolation, and it allows the method to interpolate beyond aliasing. For coherent noise attenuation the constraint is not necessary. Examples are shown of filling gaps, interpolation and coherent noise attenuation using the frequency wavenumber domain operator. Near and far offset gaps are filled in synthetics and real shot gathers, with simultaneous interpolation beyond aliasing. Waveforms are well preserved in general except where there is little coherence in the data outside the gaps or when events with very different velocities are located at the same time. Multiples of diffractions are predicted and attenuated by subtraction from the data.

Introduction

In the past years the parabolic and hyperbolic Radon transforms (RT) have been widely used in seismic processing as a tool for attenuating multiples and recently for filling gaps as well. They usually perform well in the separation of hyperbolic events in seismic data, a reasonable difference in curvature between them being the only requirement (Thorson and Claerbout, 1985, Hampson, 1987). However, it is well known that when a hyperbolic event has its apex shifted with respect to zero offset, the standard RTs are unable to properly focus this event and therefore separation of it from the rest of the data can be achieved only partly or not at all. Examples of this kind of event are sideswipe, multiples of diffractions and split peglegs (Ver West, 2002). These can survive the robust stack process because, although the tails of the diffractions stack out, the apex of these diffractions can pass through stack.

As a tool for filling gaps, the parabolic and hyperbolic RTs have been shown to be an interesting alternative to other methods (Kabir and Verschuur, 1995, Sacchi and Ulrych, 1995,Trad et al., 2002). A hyperbolic or parabolic Radon transform applied to a relatively noise free common midpoint gather (CMP) gather with large aperture and good sampling produces a sparse model. If the CMP gather contains gaps, interpolation and/or aperture extension can be achieved by enforcing sparseness on the RT (Thorson and Claerbout 1985, Sacchi and Ulrych 1995).

The quality of the results for these processes depends on how well the RT can focus the different seismic events, which in fact depends in large measure on the similarity of the seismic events to the basis functions. That is the reason why these procedures fail when the hyperbolic events have variable apex locations in the offset dimension. In this case, the only way to focus hyperbolic events to localized areas of the RT space is to shift the basis functions along offset. This goal can be achieved, for example, by using a time variant hyperbolic RT with variable apex locations (van Dedem and Verschuur 2000, Trad, 2002b). This transform represents an expansion of the data in terms of hyperbolic basis functions with apexes located at every time-offset pair of the data space or apex shifted Radon transform (ASRT).

This transformation can be visualized as an RT volume space, whose axes are apex time, velocity (or a function of velocity) and apex offset location. However, because of the high cost of such computation, we have to sacrifice generality by decreasing the number of possible apexes and/or the number of velocities. Usually in the RT we limit the number of possible apexes to one (at zero offset) and scan for as many velocities as necessary (and possible) to obtain a complete representation of the data and avoid aliasing. This can be visualized as extracting a vertical plane from inside the ASRT volume. In the case where the location of the apexes is not fixed at one offset it is less harmful to sacrifice generality in velocity than generality in apex location. Hence, we can extract a surface inside the volume by following a single stacking velocity function, but scanning for different apex locations. Since this velocity function does not necessarily correspond to the stacking velocity for primaries or multiples, it can be better obtained from a localized semblance analysis, for example by using a beam stack method (Biondi, 1992).

The idea of a hyperbolic RT with the apex location as an extra parameter has been applied before by van Dedem and Verschuur (2000, 2001) to predict multiples for cross-lines in 3D surveys. Furthermore, operators that scan for hyperbolas with apexes at different positions in the data are often used in poststack time migration algorithms. In fact, the kinematics of the ASRT is similar to the kinematics of Kirchhoff poststack time migration after replacement of the midpoint coordinate by offset. This similarity is the motivation to implement the ASRT by using efficient operators from poststack migration but fitting the data in a least squares sense.

The procedure described in this paper resembles least squares post stack time migration, but acting on time-offset gathers rather than time-midpoint gathers as in zero offset migration. Therefore, it is also related to some extent to least squares migration (Nemeth et a., 1999, Hu et al., 2001, Kuehl and Sacchi, 2000). Fitting the data in a least squares sense imposes the necessary amplitude and phase characteristics of the model, such that the forward operator acting on this model predicts the data (at least up to the achieved misfit), giving the property of reconstruction necessary for subtraction or interpolation.

In this paper I use a conjugate gradient algorithm with sparseness constraint in the time-offset space to calculate the sparse ASRT. The sparseness constraint is essential to obtain interpolation but not for removing coherent noise. By choosing a sparse solution, the transform supplies the missing information in the data with interpolation and extrapolation of the hyperbolic events in the data. Moreover, given the dispersive nature of aliasing in the time-offset domain, the sparseness constraint attenuates aliasing artifacts leading to interpolation beyond aliasing (Marfurt K. and Nemeth T., unpublished work, Trad et al., 2003).

Although operators both in time-offset and frequency-wavenumber domains are possible, this paper focuses on the frequency-wavenumber version, for several reasons. The time domain operator has been already discussed by van Dedem and Verschuur (2001,2000) and its implementation is not very different from the sparse hyperbolic RT. It is flexible but in my experience also very slow. The frequency wavenumber version, on the other hand, is interesting because of its speed and capability for preserving the spectral nature of the data (waveforms). Its implementation, however, is not trivial, in particular if a time-offset space constraint is applied, because it applies Stolt's mapping (Stolt, 1978), and requires Fourier transformation for different sampling geometries.

Implementation

Given an operator that maps two spaces, let us call them “data” $(\mathbf{d})$ and “model” $(\mathbf{m})$ independently of what these two spaces exactly represent for the problem in hand, we can find the model for a given data using a conjugate gradient methodology. In matrix notation, once $\mathbf{L}$ is defined such that $\mathbf{d} = \mathbf{Lm}$, and its adjoint operator $\mathbf{L}^*$ that performs the reverse mapping $\mathbf{m}_{adj} = \mathbf{L}^*\mathbf{d}$, we can find $\mathbf{m}$ given $\mathbf{d}$. In the context of this paper, $\mathbf{d}$ and $\mathbf{m}$ represent the data gather and transformed image after lexicographic arrangement. $\mathbf{L}$ represents a modeling operator whose impulse response is a hyperbola shifted in time and offset. The adjoint model, $\mathbf{m}_{adj}$, approximates a low resolution transform of $\mathbf{d}$, and mathematically represents a first estimation of the coefficients required in the expansion of the data in terms of a generalized Fourier series. These coefficients produce an error in the Fourier series approximation that is not optimal in a least squares sense. Inversion consists of the process of minimizing this error as much as possible by finding the optimal coefficients. These optimal coefficients are the least squares model $\mathbf{m}$.

The general problem then can be separated into two different sub-problems. One of them is how to compute the coefficients of such decomposition, taking into account that the basis functions are not orthogonal in general, and that only certain types of solution are acceptable for the problem in hand (i.e. certain constraints in the solution are required). The other sub-problem is the design of the operators that perform the mapping between data and model spaces. The first part of the problem in this work has been addressed by following the methodology explained in Trad et al. 2003.

Mapping between the data and model spaces

For the design of the operators $\mathbf{L}$ and $\mathbf{L}^*$ we first need to choose the input and output domains, keeping in mind that the input domain $\mathbf{m}$ is where sparseness (or parsimony) will be enforced, and the output domain $\mathbf{d}$ is where the events will be predicted. Previous experience with the RT (Cary, 1998) suggests that the time-offset domain is the optimal domain in which to enforce sparseness, because in that domain sparseness can be enforced in the temporal and spatial directions simultaneously, leading to clean transformations and aliasing attenuation. Therefore, a good choice for the problem in hand is to use the transformed time-offset gather as the input domain $\mathbf{m}$ and the original time-offset domain as the output domain $\mathbf{d}$.

There are two main possibilities for the hyperbolic operator $\mathbf{L}$: a time-offset operator, similar to a Kirchhoff-type hyperbolic integral, and a frequency-wavenumber operator, similar to a Stolt migration mapping. See van Dedem and Verschuur (2000) or Trad (2002b) for a discussion on the time domain operator.

The frequency wavenumber domain possibility is the poststack Stolt migration operator (Stolt, 1978). This operator is a solution to Helmothz equation that performs hyperbolic decomposition of the data by mapping from the $\omega - k_x$ to the $k_z - k_x$ domains. This mapping is the main action of the adjoint operator $\mathbf{L}^*$ in the CG algorithm. The modeling Stolt operator plays the role of $\mathbf{L}$. These operators act like a forward - inverse hyperbolic moveout operator pair in the frequency wavenumber domain.

Since the data and model spaces are in the time-offset domains, both the forward and adjoint Stolt's operators require the forward and inverse 2D Fourier transform. To enforce sparseness in the time-offset domain, the Fourier transforms have to be applied for every iteration such that the model weights act in the time-offset domain. In this case, the operators $\mathbf{L}$ and $\mathbf{L}^*$ require three main steps applied in sequence: 1) transformation from the time-offset domain to the frequency wavenumber domain, 2) wavenumber forward and inverse mapping and 3) inverse 2D FFT to the time offset domain. If sparseness is not necessary, then the Fourier transforms can be applied only at the beginning and end of the CG algorithm.

The forward and reverse Stolt's mapping in the frequency wavenumber domain are independent of the sampling of the data, but the Fourier transformation in the spatial direction depends on the nature of the offset sampling. Depending on the spatial sampling, there are two main possibilities:

1) If the data are regularly sampled in offset, but undersampled and/or with gaps, an efficient implementation is possible by first inserting zero traces in the gaps and then applying the 2D FFT. Because the inserted traces contain incorrect information they have to be down weighted in the least squares algorithm. If these zero traces were not heavily down weighted then the least squares algorithm would honor their incorrect information and produce zero traces. In spite of the down weighting, these traces are necessary because the energy is moved from the original sparse traces to the new resampled data space, generating in this process interpolation.

2) If the data are irregularly sampled in offset then the FFT in the spatial direction has to be replaced by a discrete Fourier transform (DFT). Although the method with the $\mathbf{DFT}$ can be applied for any sampling the algorithm is much slower than the implementation with $\mathbf{FFT}$ s. For situations where both methods can be implemented the results are very similar.

Examples

Figure 1a shows a shot gather, partially aliased, from a marine data set (Mississippi Canyon data, courtesy of Western Digital). Two gaps, each one of around 200m are created, one in the near zero offsets and another in the middle-far offsets. The offset sampling is extended by reciprocity to the positive offsets. A constant velocity sparse ASRT is calculated (Figure 1b), and the inverse ASRT regenerates the data and missing traces, Figure 1c. The residuals, Figure 1d, are small, less than 5% of maximum amplitude, in most of the gather (a clip=5% of the maximum amplitude is used in this figure). The differences are larger at the gap locations and the sides, where a taper has been applied. Although part of the attenuation due to the taper has been removed, the operator is unable to recover a small part of the aliased data at the high frequencies of steep events.

The difficulty in reconstructing the high frequencies for large dips is associated with aliasing. Nevertheless, the method is able to reconstruct data beyond aliasing. To achieve this goal the final Nyquist frequency has to be larger than the maximum frequency to reconstruct. In other words, if the resampled data have a correct sampling the method will relocate aliased events to their correct spectral locations. For example, Figure 2 shows in detail a part of the shot gather around the large offset gap (the gap is the area inside the rectangle). Both gaps, located at small and large offsets, are filled and at the same time the data are interpolated, reducing the sampling interval from 26.5m to 13.25m. *Figs. 2c and 2d* show the corresponding F-K spectra. It is clear from these spectra that interpolation and gap filling have been achieved beyond aliasing. This capability results from the dispersive nature of the aliased data in the transformed domain, and the action of the sparseness constraint that penalizes them.

In the second example, the ASRT is applied to attenuate coherent noise from a real CMP gather with multiples of diffractions (*Fig. 3a*). These events appear as hyperbolas with apexes very close to zero offset around 6s, but they depart from zero offset as the time increases. After applying the ASRT (*Fig. 3b*), the energy along hyperbolic curves shrinks to short hyperbolas. The closer is the operator velocity to the diffractions stacking velocity the shorter are these hyperbolas. Note how the multiples of diffractions map to areas away from zero offset (area inside polygon in *Fig. 3b*), but the other events collapse mainly to the zero offset region. Hence, it is possible to separate the non-centered hyperbolas in the RT space and, after inverse transformation, to predict them in the data space (*Fig. 3c*). Subtraction of these diffractions from the data yields a cleaner CMP gather (*Fig. 3d*). *Fig. 4* shows a zoom into the data, before (*Fig. 4a*) and after (*Fig. 4b*) subtracting multiples of diffractions.

Conclusions

In this paper an extension of the hyperbolic RT has been presented, that can be applied to interpolation and/or attenuation of hyperbolic-horizontally shifted events, in gathers with undersampling and gaps. All these procedures follow the general methodology of sparse-inversion by conjugate gradient algorithms (Claerbout, 1992).

To summarize:

- The standard hyperbolic RT can be extended to a generalized apex shifted RT that maps hyperbolas in any location of the gather to small areas of the transform domain.

- Apex shifted hyperbolic events in seismic data can be interpolated by enforcing sparseness in the transform domain and mapping back to a regularly sampled offset time domain. Selective prediction and subtraction can be used to attenuate apex shifted multiples.

- The Stolt operator, being a very fast migration operator that collapses hyperbolas to impulses, can be applied in a least squares algorithm to implement a very fast hyperbolic ASRT. In principle the method assumes a flat geology but in practice it performs well in realistic structures.

- The main advantages of the proposed procedure over a conventional hyperbolic Radon transform are that the events to be interpolated may have their apexes located at different offsets, that it is faster and also preserves better the spectral nature of the data.

Acknowledgment

My sincere acknowledgment to Tadeusz Ulrych and Mauricio Sacchi for scientific guidance, to the CDSST sponsors for financial support during this work, to BP, Veritas and Chevron-Texaco for additional support of this work. I greatly appreciate the suggestions and encouragement of Ken Matson and Tamas Nemeth and valuable suggestions by two anonymous reviewers.

References

Biondi, B., 1992, Velocity estimation by beam stack: Geophysics, 57, 8, 1034-1047.

Cary, P. W., 1998, The simplest discrete radon transform: 68th Ann. Internat. Mtg.,

Soc. of Expl. Geophys., Expanded Abstracts, 1999-2002.

Claerbout, J., 1992, Earth sounding analysis, processing versus inversion: Blackwell

Scientific Publications, Inc.

Hampson, D., 1987, The discrete radon transform: A new tool for image enhancement and noise suppression: 57th Ann. Internat. Mtg., Soc. of Expl. Geophys., Expanded

Abstracts, Session: BP3.CSEG.

Hu, J., Schuster, G. T., and Valasek, P., 2001, Poststack migration deconvolution: Geophysics, 66, 939-952.

Kabir, M. M. N., and Verschuur, D. J., 1995, Restoration of missing offsets by parabolic radon transform: Geophysics, 43, 347-368.

Kuehl, H., and Sacchi, M., 2000, Born-wkbj migration/inversion with split-step correction: 70th Ann. Internat. Mtg, Soc. of Expl. Geophys., 1040-1043.

Nemeth, T., Wu, C., and Schuster, G. T., 1999, Least squares migration of incomplete reflection data: Geophysics, 64, 208-221.

Sacchi, M. D., and Ulrych, T. J., 1995, High-resolution velocity gathers and offset

space reconstruction: Geophysics, 60, 1169-1177.

Stolt, R. H., 1978, Migration by Fourier transform: Geophysics, 43, 23-48.

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

Trad, D., Ulrych, T., and Sacchi, 2003, Latest views of the sparse radon transform: Geophysics, 68, 1, 386-399.

Trad, D., 2002b, Interpolation with migration operators: 72th Ann. Internat. Mtg., Soc. of Expl. Geophys., Expanded Abstracts.

Trad, D., Ulrych, T., and Sacchi, 2002, Accurate interpolation with high resolution time variant radon transforms: Geophysics, 67, no. 2, 644-656.

van Dedem, E., and Verschuur, D., 2000, 3-d surface-related multiple prediction, an inversion approach: 70th Ann. Internat. Mtg., Soc. of Expl Geophys., Expanded Abstracts, 1965-1968.

van Dedem, E., and Verschuur, D., 2001, 3-d surface multiple prediction using sparse inversion: 71st Ann. Internat. Mtg, Soc. of Expl. Geophys., 1285-1288.

VerWest, B. 2002, Suppressing Peg-leg Multiples with Parabolic Radon Demultiple, EAGE 64th Conference & Exhibition.

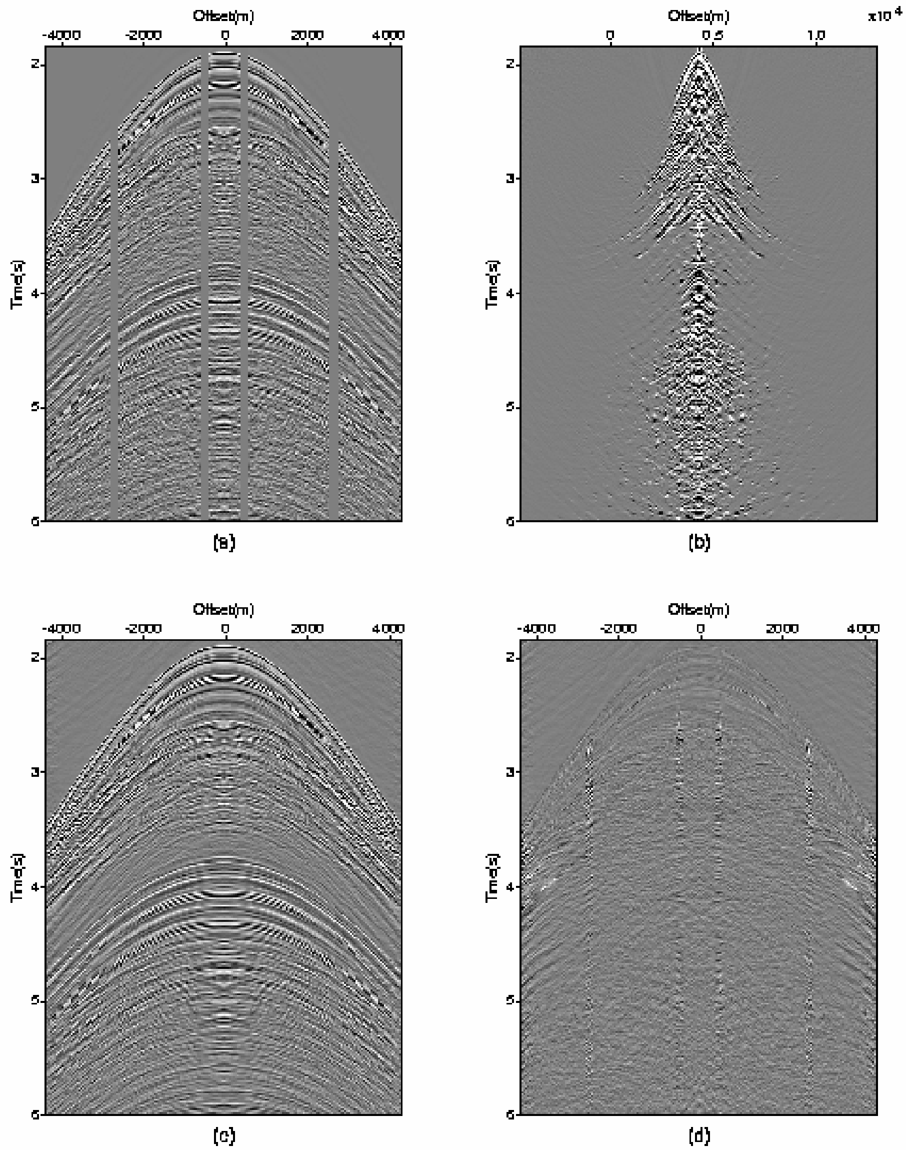

Fig. 1- a) A marine shot gather with two gaps (around 200 meters each) duplicated by reciprocity. (b) Its ASRT transform. (c) Predicted gather (Inverse ASRT). (d) Residuals.

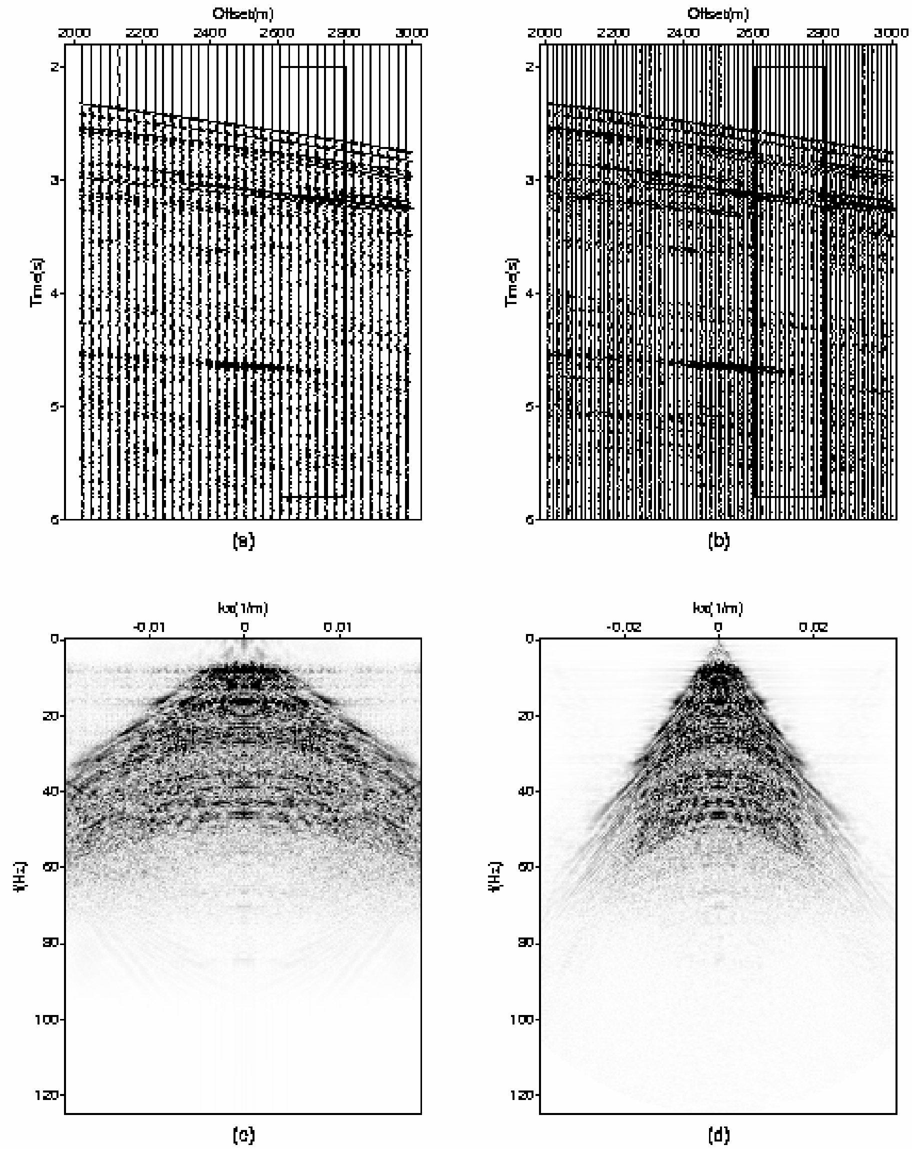

Fig. 2- Gap interpolation and upsampling (a) Shot gather around the far zero offset gap (gap inside de rectangle). (b) Predicted gather with upsampling. (c) FK spectrum of (a), (d) FK spectrum of (b).

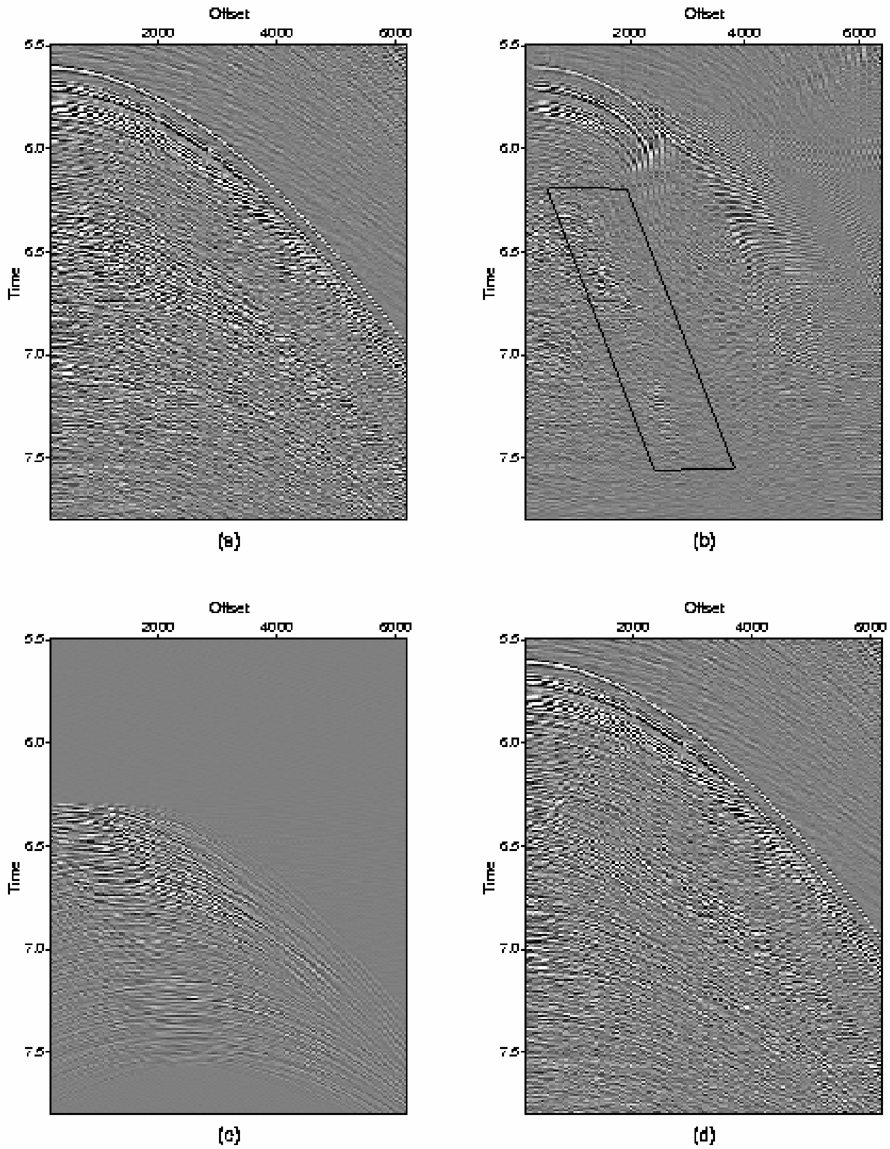

Fig. 3 - Attenuation of multiples from diffractions. (a) Data. (b) ASRT. (c) Predicted multiples of diffractions. (d) Difference between (a) and (c).

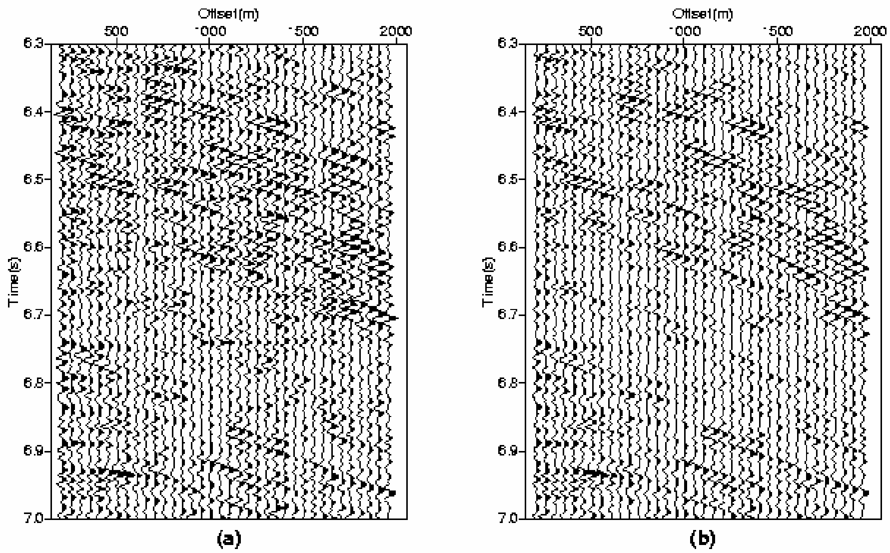

Fig. 4 - Attenuation of multiples from diffractions. (a) data. (b) data after subtracting diffractions.