Stanford Exploration Project, Report 129, May 6, 2007, pages 85–94

Interpolation with pseudo-primaries: revisited

William Curry

ABSTRACT

Large gaps may exist in marine data at near offsets. I generate pseudo-primaries by autocorrelating data containing both primary and multiple reflections. These pseudo-primaries are used as training data for a non-stationary prediction-error filter, which is then used to interpolate the missing near offsets. This method yields good results using filters in the t-x domain, whereas f-x domain filters do not fare well at eliminating the crosstalk in the pseudo-primaries.

INTRODUCTION

Interpolation has become more important recently, largely due to increased reliance on algorithms that require dense and regular data sampling, such as wave-equation migration and 3D surface-related multiple elimination (SRME) (Dedem and Verschuur, 2005). Examples of current interpolation methods include Fourier (Duijndam and Schonewille, 1999; Liu and Sacchi, 2004; Xu et al., 2005), Radon transform (Trad, 2003), and prediction-error filter (PEF) based methods (Spitz, 1991). Other methods that rely on the underlying physics (and typically also a velocity model) include migration/demigration (Pica et al., 2005), DMO-based methods (Biondi and Vlad, 2001), and the focal transform (Berkhout et al., 2004), which requires an input focal operator instead of velocity.

In this paper, I further examine a hybrid approach that combines both non-stationary PEFs (Crawley, 2000) and pseudo-primaries generated from surface-related multiples (Shan and Guitton, 2004) in order to interpolate missing near offsets. I generate pseudo-primaries by auto-correlation of the input data, which gives a similar result to the cross-correlation of the input data with a multiple model described in a previous paper (Curry, 2006). Once the pseudo-primaries have been generated, I estimate a non-stationary PEF on the pseudo-primaries by solving a least-squares problem. I then solve a second least-squares problem where the newly found PEF is used to interpolate the missing data (Claerbout, 1999). This is done in both the t-x and f-x domains.

The data used in this example is the Sigsbee2B synthetic dataset where the first 2000 feet of offset were removed. Near-offset data is typically missing from marine data, and large near-offset gaps can exist when undershooting obstacles such as drilling platforms. Estimating a PEF on the pseudo-primaries, which are generated without the near offset data, gives promising results using t-x filters but f-x filters do not eliminate the crosstalk.

GENERATION OF PSEUDO-PRIMARIES

Pseudo-primaries can be generated by computing a slightly modified version of cross-correlation of primaries and a multiple model (Shan and Guitton, 2004)

$$W(x_p, x_m, \omega) = \sum_{x_s} D(x_s, x_m, \omega)\bar{D}(x_s, x_p, \omega), \tag{1}$$

where W is the pseudo-primary data, ω is frequency, x_s is the shot location, x_p and x_m are receiver locations, $\bar{D}(x_s, x_p, \omega)$ is the complex conjugate of the original trace at (x_s, x_p) and $D(x_s, x_p, \omega)$ is that same data at x_m. In this equation, the result of the cross-correlation of primaries, first-order multiples and second-order multiples in D are outlined in the table below, with the first and second columns corresponding to the inputs to the cross-correlation and the third column corresponding to the output.

Input 1	Input 2	Output
first-order multiples	first-order multiples	zero-lag
second-order multiples	second-order multptles	zero-lag
first-order multiples	primaries	pseudo-primaries
second-order multiples	first-order multiples	pseudo-primaries
second-order multiples	primaries	pseudo-first-order multiples

With higher order multiples the trend in this table continues. This produces similar results to cross-correlating primaries with a multiple model (Shan and Guitton, 2004), as the additional correlation of the primaries on one term is already taking place with the identical primaries on the other term of the autocorrelation.

Pseudo-primaries generated in this fashion contain subsurface information that would not be recorded with a non-zero minimum offset. One example of this is a first-order multiple that reflects at the free surface within the recording array, resulting in near offsets being recorded when that wave returns to the surface. This is shown in Figures 1 and 2, where Figure 1 is a cube of the input Sigsbee2B shots (including the negative offsets predicted by reciprocity) but with offsets less than 2000 feet removed, and Figure 2 is the corresponding cube of pseudo-primaries for the same area. Put briefly, the source coverage of the pseudo-primary data is much greater than that of the input data because all receivers in the original data become sources for the pseudo-primaries.

Figure 2 contains a lot of near-offset information present in the pseudo-primaries that is not present in the recorded primaries. However, simply replacing the missing near offsets of the primaries with the corresponding pseudo-primaries would not yield a satisfactory result due to the crosstalk and noise in the pseudo-primaries.

The crosstalk in the pseudoprimaries is largely a function of the number of shots that are summed over in the input data. Figure 3 shows the shot on the right panel of Figure 2, but without the summing over shots in equation 1 where instead of summing over shots each shot is plotted along the front face of the cube. It shows how the stacking procedure greatly increases the signal-to-noise ratio.

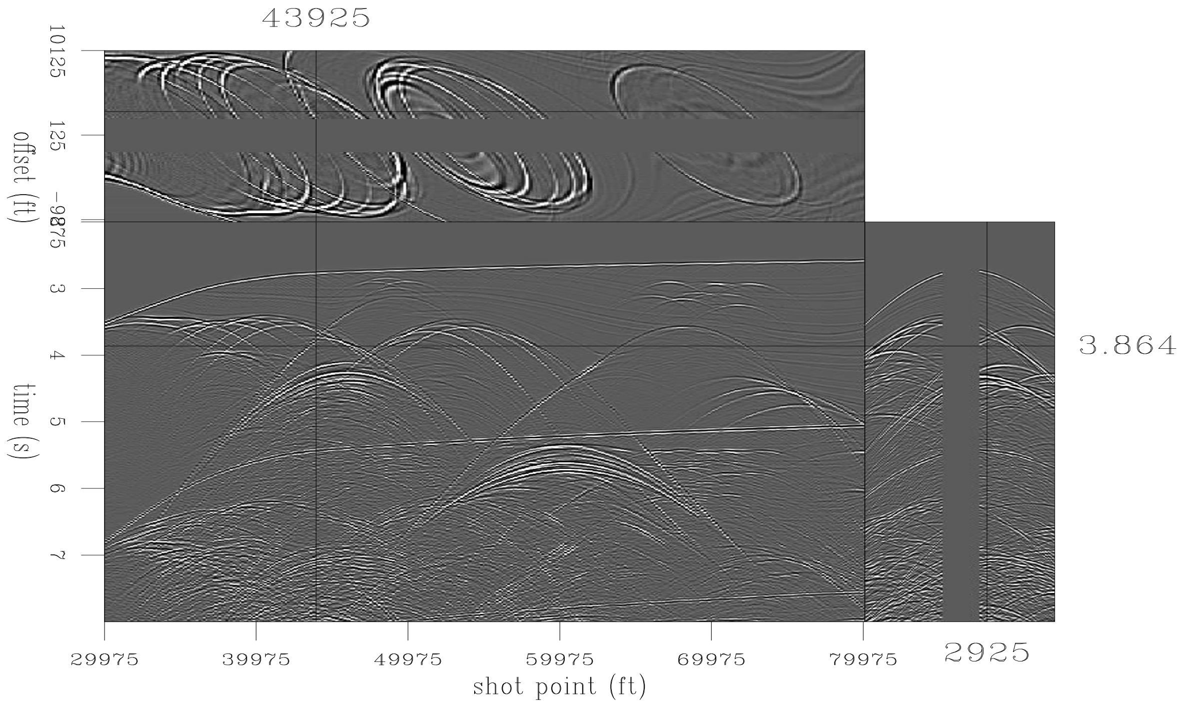

Figure 1: Input dataset missing the nearest 2000' of offsets on either side. bill2-input [ER]

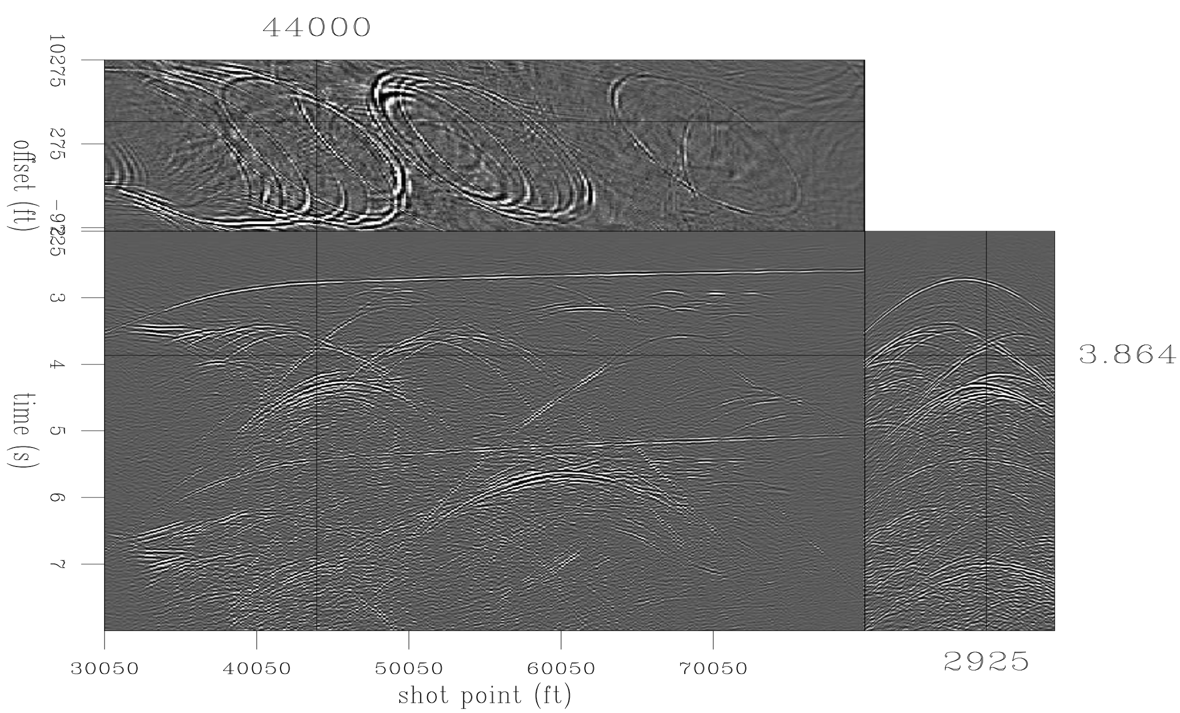

Figure 2: Pseudo-primaries created by autocorrelation and summation of Figure 1. Note that the near offsets have been filled in. bill2-pseudo [CR]

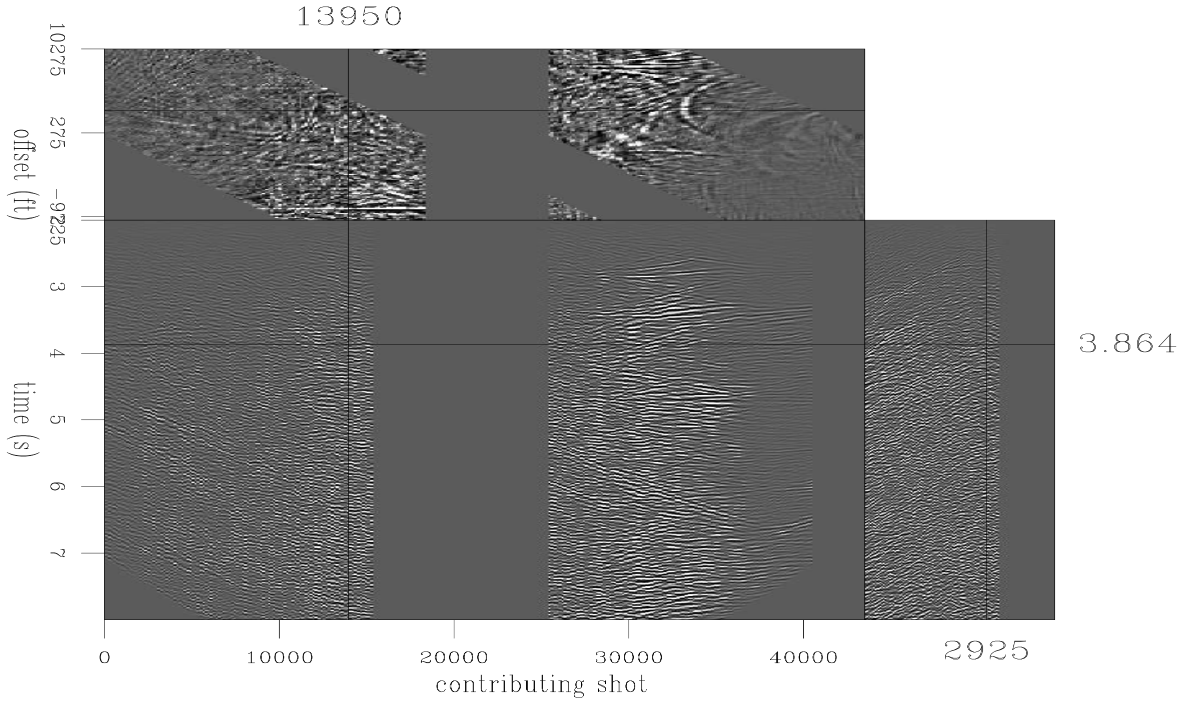

Figure 3: A single virtual shot from figure 2 that has been expanded over the input shot axis. bill2-pseudoex [CR]

INTERPOLATION WITH NON-STATIONARY PEFS

Interpolation can be cast as a series of two inverse problems where a prediction-error filter is estimated on known data and is then used to interpolate missing data. A prediction-error filter (PEF) can be estimated by minimizing the output of convolution of known data with an unknown filter (except for the leading 1), which can be written in matrix form as

$$\mathbf{0} \quad \approx \quad \mathbf{r} = \begin{bmatrix} d_2 & d_1 & d_0 \\ d_3 & d_2 & d_1 \\ d_4 & d_3 & d_2 \\ d_5 & d_4 & d_3 \\ d_6 & d_5 & d_4 \end{bmatrix} \begin{bmatrix} 0 & \cdot & \cdot \\ \cdot & 1 & \cdot \\ \cdot & \cdot & 1 \end{bmatrix} \begin{bmatrix} 1 \\ f_1 \\ f_2 \end{bmatrix} + \begin{bmatrix} d_2 \\ d_3 \\ d_4 \\ d_5 \\ d_6 \end{bmatrix}, \tag{2}$$

where f_i are unknown filter values and d_i are known data values.

The filters used in this paper are all multidimensional, which are computed with the helical coordinate. In the case of a stationary multidimensional PEF, this is an over-determined least-squares problem with a unique solution.

Seismic data is non-stationary in nature, so a single stationary PEF is not adequate for changing dips. I estimate a single spatially-variable non-stationary PEF and solve a global optimization problem (Guitton, 2003). In that case the problem is under-determined, and a regularization term is added to the least-squares problem so that (in matrix notation),

$$\begin{aligned} \mathbf{DKf} + \mathbf{d} &\approx \mathbf{0} \\ \epsilon \mathbf{Af} &\approx \mathbf{0}, \end{aligned} \tag{3}$$

where **D** represents non-stationary convolution with the data, **f** is now a non-stationary PEF, **K** (a selector matrix) and **d** (a copy of the data) both constrain the value of the first filter coefficient to 1, **A** is a regularization operator (a Laplacian operating over space) and ϵ is a trade-off parameter for the regularization. Solving this system will create a smoothly variable PEF.

Once the PEF has been estimated, it can be used in a second least squares problem that matches the output model to the known data while simultaneously regularizing the model with the newly found PEF,

$$\begin{aligned} \mathbf{S}(\mathbf{m} - \mathbf{d}) &= \mathbf{0} \\ \mathbf{Fm} &\approx \mathbf{0}, \end{aligned} \tag{4}$$

where **S** is a selector matrix which is 1 where data is present and 0 where it is not, **F** represents convolution with the non-stationary PEF, and **m** is the desired model.

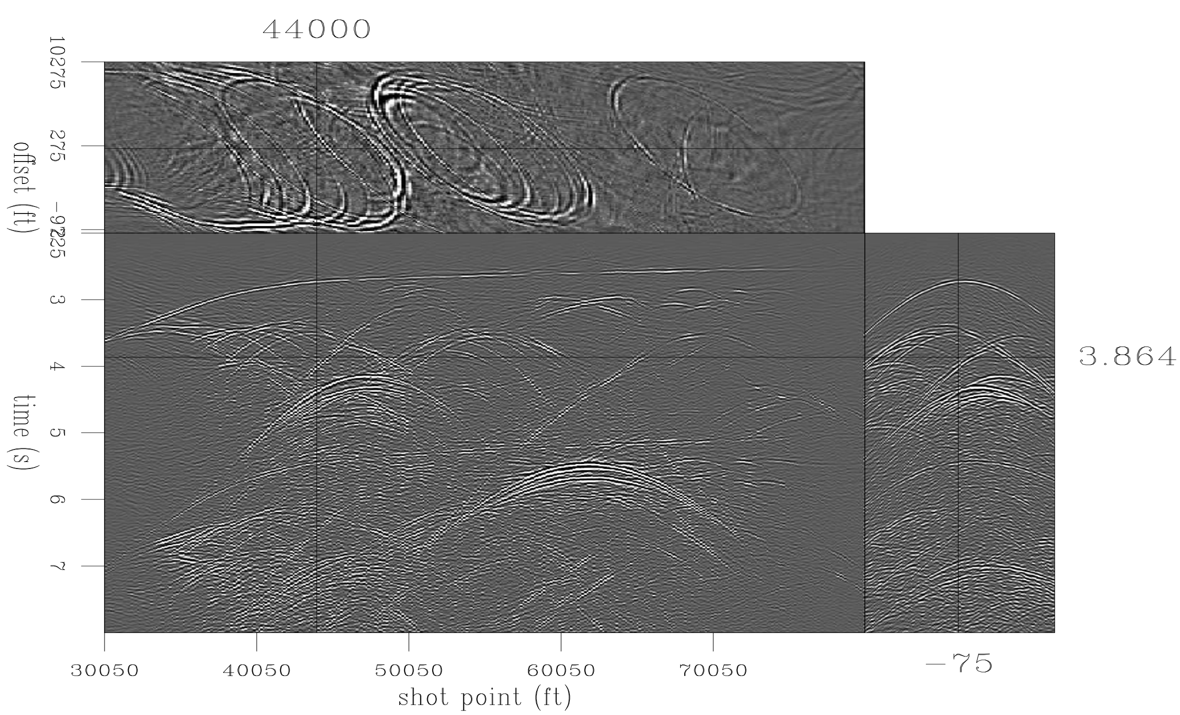

Figure 4: Pseudo-primaries located on a missing common-offset section from the input data set. bill2-pseudo2 [CR]

RESULTS

To increase the sampling by an integer factor, a PEF is typically estimated on the input data with some sort of change in sampling. In this example with a large gap this will not suffice. Instead, we estimate the PEF on the pseudo-primaries generated by equation 1 using equation 3 and then use that PEF to interpolate the recorded data with equation 4. The results of this experiment (using a non-stationary t-x filter) are shown in Figure 5.

The near offset gap is 4000 feet or 26 traces, as shown in Figure 1. A non-stationary PEF in the t-x domain is then independently estimated on each shot in Figure 4, and that PEF is then used to fill in the missing data in the input and the result is shown in Figure 5.

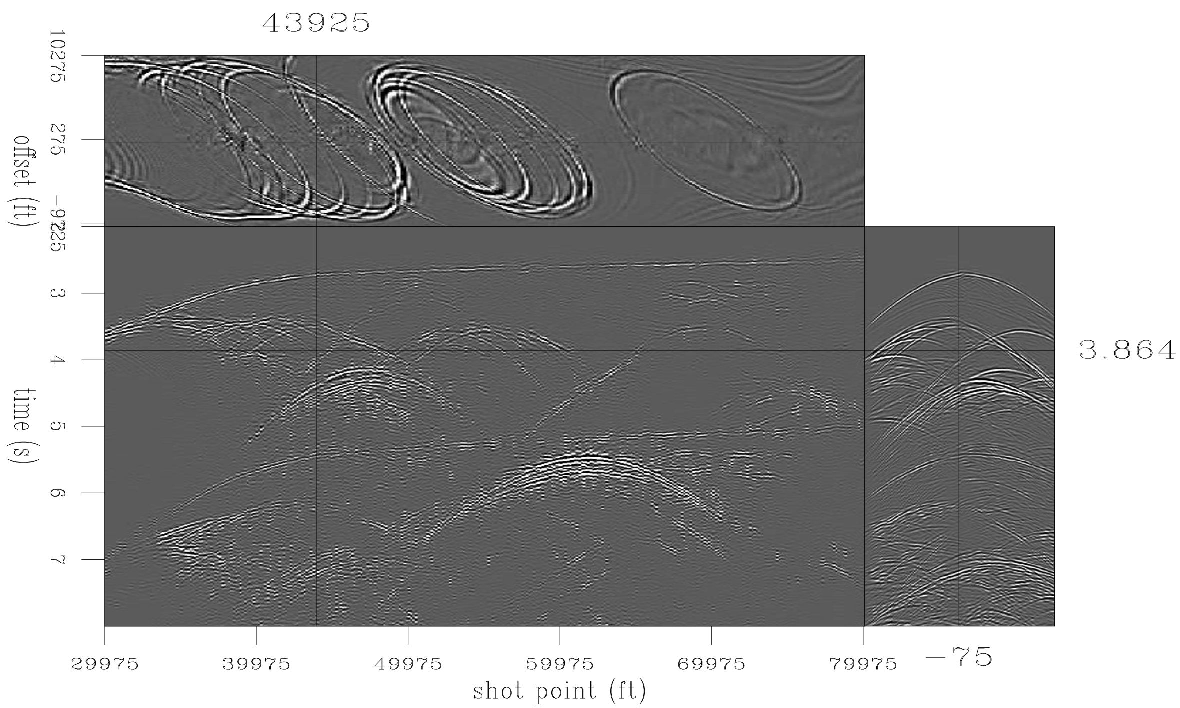

Figure 5: Near offsets interpolated with a non-stationary PEF trained on the pseudo-primaries in the t,offset domain. bill2-txinterp [CR]

The result in Figure 5 looks very good. The interpolated portion of the shot on the right side of Figure 5 looks very good, with diffractions and crossing events correctly interpolated. Most of the crosstalk present in the pseudo-primaries in Figure 4 is gone. However, there is a sight change in the wavelet, which can be explained by the squaring of the wavelet in the cross-correlation and the t-x domain PEF capturing spectral information.

Now looking at the front panel of Figure 5 that corresponds to a constant-offset section, another problem becomes more apparent. The section looks somewhat jagged from one shot to the next. This is largely because this problem was solved on a shot-by-shot basis so that there is no guarantee of lateral continuity between shots. This could be remedied by using a 3D PEF in t,offset, and shot so that correlations between shots would be taken into account.

Another less expensive method would be to use a non-stationary f-x-y domain PEF (Curry, 2007), described next.

Instead of interpolating in t-x, we can transform the data into the frequency domain and perform the interpolation individually for each frequency. This approach would largely use the same machinery as the t-x approach, except that all numbers are now complex and the filter would be 2D in shot and offset space. Since this filter does not operate along time or frequency no spectral information is captured, which should eliminate the wavelet issue in the t-x example and the added shot axis (due to the much lower memory requirements of the method) should reduce the jitteriness across the shot axis in the result.

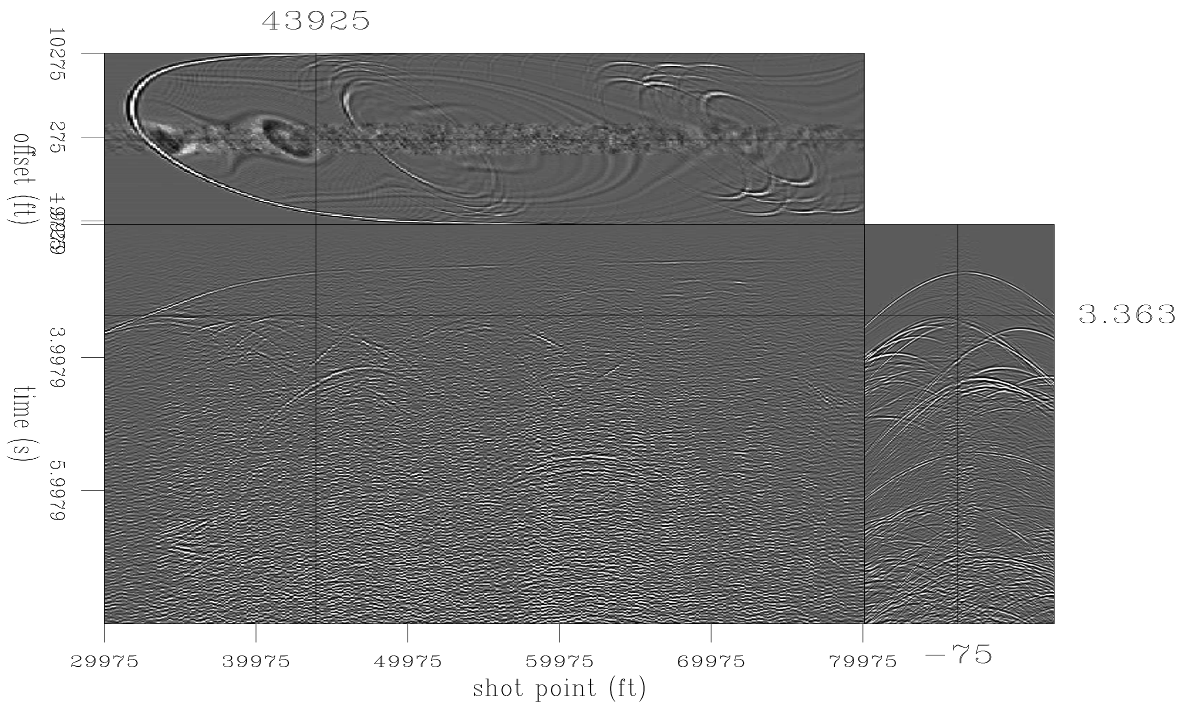

Figure 6: Near offsets interpolated with a non-stationary PEF trained on the pseudo-primaries in the f,offset,and shot domain. bill2-fxinterp [CR]

The f-x-y interpolation result, shown in Figure 6 does reduce both of these problems. The wavelet now remains consistent between the recorded data and the interpolated data. The interpolation result also seems less jagged along the shot axis as the PEF is also estimated along this axis. However, the problem of what appears to be crosstalk is much worse. This is due to assumed stationarity in time. Since the data were Fourier Transformed as a whole and the problem is solved in the frequency domain, all dips for a given spatial location appear to be present at all times. This problem could hopefully be addressed by breaking the input data into small time windows that are more stationary in time.

CONCLUSIONS AND FUTURE WORK

Incorporating pseudo-primary data into a non-stationary prediction-error filter based interpolation method gives promising results for large gaps in the near offset. This problem would be very difficult to solve without the additional information provided by the pseudo-primaries, and the prediction-error filter approach eliminates a lot of the crosstalk that a simple cut-and-paste approach would have.T-x filters give a much nicer result than f-x, even with the greater dimensionality of the f-x filters. This is largely due to the issue of non-stationarity in time, which may be addressed by using the f-x approach in small time windows.

This method has been attempted on real data, and would initially appear to have the most benefit in the cross-line direction by creating pseudo-source lines where the receiver cables are, but the signal-to-noise ratio of initial attempts is very poor and not useful to show here. This would be equivalent to reducing the number of samples along the shot axis in Figure 3, which would clearly present problems. Also, the issues of cable feathering, swerving sail lines, 3D geometry and coherent noise all present problems with using this approach with real data.

ACKNOWLEDGMENTS

The author would like to thank the SMAART JV for the Sigsbee2B dataset, and Guojian Shan for assistance with the pseudo-primaries.

REFERENCES

Berkhout, A. J., D. J. Verschuur, and R. Romijn, 2004, Reconstruction of seismic data using the focal transformation: Reconstruction of seismic data using the focal transformation:, Soc. of Expl. Geophys., 74th Ann. Internat. Mtg., 1993–1996.

Biondi, B. and I. Vlad, 2001, Amplitude preserving prestack imaging of irregularly sampled 3-D data: SEP–**110**, 1–18.

Claerbout, J., 1999, Geophysical estimation by example: Environmental soundings image enhancement: Stanford Exploration Project, `http://sepwww.stanford.edu/sep/prof/`.

Crawley, S., 2000, Seismic trace interpolation with nonstationary prediction-error filters: SEP–**104**.

Curry, W., 2006, Interpolation with pseudo-primaries: SEP–**124**.

Curry, W., 2007, Non-stationary interpolation in the f-x domain: SEP–**129**.

Dedem, E. V. and D. Verschuur, 2005, 3D surface-related multiple prediction: A sparse inversion approach: Geophysics, **70**, no. 03, V31–V43.

Duijndam, A. J. W. and M. A. Schonewille, 1999, Nonuniform fast Fourier transform: Geophysics, **64**, no. 2, 539–551.

Guitton, A., 2003, Multiple attenuation with multidimensional prediction-error filters: SEP–**113**, 57–74.

Liu, B. and M. D. Sacchi, 2004, Minimum weighted norm interpolation of seismic records: Geophysics, **69**, no. 6, 1560–1568.

Pica, A., G. poulain, et al., 2005, 3D surface-related multiple modeling: The Leading Edge, **24**, no. 3, 292–296.

Shan, G. and A. Guitton, 2004, Migration of surface-related multiples: Tests on Sigsbee2B data sets: Migration of surface-related multiples: Tests on Sigsbee2B data sets:, Soc. of Expl. Geophys., 74th Ann. Internat. Mtg., 1285–1288.

Spitz, S., 1991, Seismic trace interpolation in the F-X domain: Geophysics, **56**, no. 06, 785–794.

Trad, D. O., 2003, Interpolation and multiple attenuation with migration operators: Geophysics, **68**, no. 6, 2043–2054.

Xu, S., Y. Zhang, D. Pham, and G. Lambare, 2005, Antileakage Fourier transform for seismic data regularization: Geophysics, **70**, no. 04, V87–V95.