Seismic data interpolation beyond aliasing using regularized nonstationary autoregression

Yang Liu¹ and Sergey Fomel²

ABSTRACT

Seismic data are often inadequately or irregularly sampled along spatial axes. Irregular sampling can produce artifacts in seismic imaging results. We have developed a new approach to interpolate aliased seismic data based on adaptive prediction-error filtering (PEF) and regularized nonstationary autoregression. Instead of cutting data into overlapping windows (patching), a popular method for handling nonstationarity, we obtain smoothly nonstationary PEF coefficients by solving a global regularized least-squares problem. We employ shaping regularization to control the smoothness of adaptive PEFs. Finding the interpolated traces can be treated as another linear least-squares problem, which solves for data values rather than filter coefficients. Compared with existing methods, the advantages of the proposed method include an intuitive selection of regularization parameters and fast iteration convergence. The technique was tested on benchmark synthetic and field data to prove it can successfully reconstruct data with decimated or missing traces.

INTRODUCTION

The regular and fine sampling along the time axis is common, whereas good spatial sampling is often more expensive or prohibitive and therefore is the main bottleneck for seismic resolution. Too large a spatial sampling interval may lead to aliasing problems that adversely affect the resolution of subsurface images. An alternative to expensive dense spatial sampling is interpolation of seismic traces. One important approach to trace interpolation is prediction interpolating methods (Spitz, 1991), which use low-frequency nonaliased data to extract antialiasing prediction-error filters (PEFs) and then interpolates high frequencies beyond aliasing. Claerbout (1992) extends Spitz’s method by using PEFs in the t-x domain. Porsani (1999) proposes a half-step PEF scheme that makes the interpolation process more efficient. Huard et al. (1996) and Wang (2002) extend f-x trace interpolation to higher spatial dimensions. Gulunay (2003) introduces an algorithm similar to f-x prediction filtering, which has an elegant representation in the f-k domain. Curry (2006) uses multidimensional nonstationary PEFs to interpolate diffracted multiples. Naghizadeh and Sacchi (2009) propose an adaptive f-x interpolation using exponentially weighted recursive least squares. More recently, Naghizadeh and Sacchi (2010a) propose a prediction approach similar to Gulunay’s method but using the curvelet transform instead of the Fourier transform. Abma and Kabir (2005) compare the performance of several different interpolation methods.

Correcting irregular spatial sampling is another application for seismic data interpolation algorithms. A variety of interpolation methods have been published in the recent years. One approach is to estimate the PEF on multiple rescaled copies of the irregular data (Curry, 2003), where the data are rescaled with a number of progressively larger bin sizes. Curry (2004) further improves the rescaling method by introducing multiple scales of the data where the location of the grid cells are varied in addition to the size of the cells. Curry and Shao (2008) use pseudoprimary data by crosscorrelating multiples and primaries to estimate nonstationary PEF and then interpolated missing near offsets. Naghizadeh and Sacchi (2010b) propose autoregressive spectral estimates to reconstruct aliased data and data with gaps.

Seismic data are nonstationary. The standard PEF is designed under the assumption of stationary data and becomes less effective when this assumption is violated (Claerbout, 1992). Cutting data into overlapping windows (patching) is a common method to handle nonstationarity (Claerbout, 2010), although it occasionally fails in the presence of variable dips. Crawley et al. (1999) propose smoothly varying nonstationary PEFs with micropatches and radial smoothing, which typically produces better results than the rectangular
patching approach. Fomel (2002) develops a nonstationary plane-wave destruction (PWD) filter as an alternative to t-x PEF (Claerbout, 1992) and applies the PWD operator to trace interpolation. The PWD method depends on the assumption of a small number of smoothly variable seismic dips. Curry (2003) uses Laplacian and radial rougheners to ensure a nonstationary PEF that varies smoothly in space, which specifies an appropriate regularization operator.

In this paper, we use the two-step strategy, similar to that of Claerbout (1992) and Crawley et al. (1999), but calculate the adaptive PEF by using regularized nonstationary autoregression (Fomel, 2009) to handle both nonstationarity and aliasing. The key idea is the use of shaping regularization (Fomel, 2007) to constrain the spatial smoothness of filter coefficients. We provide an approach to nonstationary data interpolation, which has an intuitive selection of parameters and fast iteration convergence. We test the new method by using several benchmark synthetic examples. Results of applying the proposed method to a field data example demonstrate that it can be effective in trace interpolation problems, even in the presence of multiple strongly variable slopes.

**THEORY**

A common constraint for interpolating missing seismic traces is to ensure that the interpolated data, after specified filtering, have minimum energy (Claerbout, 1992). Filtering is equivalent to spectral multiplication. Therefore, specified filtering is a way of prescribing a spectrum for the interpolated data. A sensible choice is a spectrum of the recorded data, which can be captured by finding the data’s PEF (Spitz, 1991; Crawley 2000). The PEF, also known as the autoregression filter, plays the role of the “inverse-covariance matrix” in statistical estimation theory. A signal is regressed on itself in the estimation of PEF. The PEF can be implemented in either t-x (time-space) or f-x (frequency-space) domain. Time-space PEFs are less likely to create spurious events in the presence of noise than f-x PEFs (Abma 1995; Crawley, 2000). When data interpolation is cast as an inverse problem, a PEF can be used to find missing data. This involves a two-step approach. In the first step, a PEF is estimated by minimizing the output of convolution of known data with an unknown PEF. In the second step, the missing data are found by minimizing the convolution of the recently calculated PEF with the unknown model, which is constrained where the data are known (Curry, 2004).

**Step 1: Adaptive PEF estimation**

**Regular trace interpolation**

An important property of PEFs is scale invariance, which allows estimation of PEF coefficients $B_n$ (including the leading “−1” and prediction coefficients $B_n$) for incomplete aliased data $S(t,x)$ that include known traces $S_{known}(t,x)$ and unknown or zero traces $S_{zero}(t,x)$. For trace decimation, zero traces interlace known traces. To avoid zeros that influence filter estimation, we interlace the filter coefficients with zeros. For example, consider a 2D PEF with seven prediction coefficients:

$$
\begin{bmatrix}
B_3 & B_4 & B_5 & B_6 & B_7 \\
\cdot & \cdot & \cdot & -1 & B_1 & B_2
\end{bmatrix}
$$

(1)

Here, the horizontal axis is time, the vertical axis is space, and “·” denotes zero. Rescaling both time and spatial axes assumes that the dips represented by the original filter in equation 1 are the same as those represented by the scaled filter (Claerbout, 1992):

$$
\begin{bmatrix}
B_3 & B_4 & B_5 & B_6 & B_7 \\
\cdot & \cdot & \cdot & -1 & B_1 & B_2
\end{bmatrix}
$$

(2)

For nonstationary situations, we can also assume locally stationary spectra of the data because trace decimation makes the space between known traces small enough, thus making adaptive PEFs locally scale-invariant. For estimating adaptive PEF coefficients, nonstationary autoregression allows coefficients $B_n$ to change with both $t$ and $x$. The new adaptive filter can look something like

$$
\begin{bmatrix}
B_3(t,x) & B_4(t,x) & B_5(t,x) & B_6(t,x) & B_7(t,x) \\
\cdot & \cdot & \cdot & -1 & B_1(t,x) & B_2(t,x)
\end{bmatrix}
$$

(3)

In other words, prediction coefficients $B_n(t,x)$ are obtained by solving the least-squares problem,

$$
\hat{B}_n(t,x) = \arg \min_{B_n} \left\| S(t,x) - \sum_{n=1}^{N} B_n(t,x)S_n(t,x) \right\|^2_2 + \epsilon^2 \sum_{n=1}^{N} \left\| D[B_n(t,x)] \right\|^2_2,
$$

(4)

where $S_n(t,x) = S(t - m_i \Delta t, x - m_j \Delta x)$, which represents the causal translation of $S(t,x)$, with time-shift index $i$ and spatial-shift index $j$ scaled by decimation interval $m$. Note that predefined constant $m$ uses the interleaving value as an interval; i.e., the shift interval equals 2 in equation 3. Subscript $n$ is the general shift index for both time and space, and the total number of $i$ and $j$ is $N$. $D$ is the regularization operator, and $\epsilon$ is a scalar regularization parameter. All coefficients $B_n(t,x)$ are estimated simultaneously in a time/space variant manner. This approach was described by Fomel (2009) as regularized nonstationary autoregression (RNA). If $D$ is a linear operator, least-squares estimation reduces to linear inversion

$$
b = A^{-1}d,
$$

(5)

where

$$
b = \begin{bmatrix} B_1(t,x) & B_2(t,x) & \cdots & B_N(t,x) \end{bmatrix}^T,
$$

(6)

$$
d = \begin{bmatrix} S_1(t,x) & S_2(t,x) & \cdots & S_N(t,x) \end{bmatrix}^T,
$$

(7)

and the elements of matrix $A$ are

$$
A_{nk}(t,x) = S_n(t,x)S_k(t,x) + \epsilon^2 \delta_{nk}D^T D.
$$

(8)

Shaping regularization (Fomel, 2007) incorporates a shaping (smoothing) operator $G$ instead of $D$ and provides better numerical properties than Tikhonov’s regularization (Tikhonov, 1963) in
Regularized nonstationary autoregression

Step 2: Data interpolation with adaptive PEF

In the second step, a similar problem is solved, except that the filter is known, and the missing traces are unknown. In the decimated-trace interpolation problem, we squeeze (by throwing away alternate zeroed rows and columns) the filter in equation 3 to its original size and then formulate the least-squares problem,

$$\hat{S}(t,x) = \arg \min_S \left\{ \left\| S(t,x) - \sum_{n=1}^{N} \hat{B}_n(t,x)S_n(t,x) \right\|^2_2 \right\},$$

subject to

$$\hat{S}(t,x_k) = S_{\text{known}}(t,x_k),$$

where \(\hat{S}(t,x)\) represents the interpolated output, and \(i\) and \(j\) use the original shift as the interval; i.e., the shift interval equals 1.

We carry out the minimization in equations 4, 13, and 14 by the conjugate-gradient method (Hestenes and Stiefel, 1952). The constraint condition (equation 15) is used as the initial model and constrains the output by using the known traces for each iteration in the conjugate-gradient scheme. The computational cost is proportional to \(N_{\text{iter}} \times N_f \times N_x \times N_y\), where \(N_{\text{iter}}\) is the number of iterations, \(N_f\) is the filter size, and \(N_x \times N_y\) is the data size. In our tests, \(N_f \) and \(N_{\text{iter}}\) were approximately equal to 100. Increasing the smoothing radius in shaping regularization decreases \(N_{\text{iter}}\) in the filter estimation step.

SYNTHETIC DATA TESTS

Aliasing decimated-trace interpolation test

We start with a strongly aliased synthetic example from Claerbout (2009). The sparse spatial sampling makes the gather severely aliased, especially at the far-offset positions (Figure 1a). For comparison, we used PWD (Fomel, 2002) to interpolate the traces (Figure 1b). Interpolation with PWD depends on dip estimation. In this example, the true dip is nonnegative everywhere and is easily distinguished from the aliased one. Therefore, the PWD method recovers the interpolated traces well. However, in the more general case, an additional interpretation may be required to determine which of the dip components is contaminated by aliasing. According to the theory described in the previous section, the PEF-based methods use the lower (less aliased) frequencies to estimate PEF coefficients, and then interpolate the decimated traces (high-frequency information) by minimizing the convolution of the scale-invariant PEF with the unknown model, which is constrained where the data are known. We designed adaptive PEFs using 10 (time) × 2 (space) coefficients for each sample and a 50-sample (time) × 2-sample (space) smoothing radius and then applied them so as to interpolate the aliased trace. The nonstationary autoregression algorithm effectively removes all spatial aliasing artifacts (Figure 1c). The proposed method compares well with the PWD method. The CPU times, for single 2.66 GHz CPU used in this example, are 20 s for adaptive PEF estimation (step 1) and 2 s for data interpolation (step 2).

Abma decimated-trace interpolation tests

A benchmark example created by Raymond Abma (personal communication, 2003) shows a simple curved event (Figure 2a).
The challenge in this example is to account for both nonstationarity and aliasing. Figure 2b shows the interpolated result using Claerbout's stationary \( t \times x \) PEF, which was estimated and applied in one big window, with each PEF coefficient \( B_n \) constant at every data location. Note that the \( t \times x \) PEF method can recover the aliasing trace only in the dominant slope range. The trace-interpolating result using regularized nonstationary autoregression is shown in Figure 2c. The adaptive PEF has \( 20 \times 3 \) coefficients for each sample and a 20-sample \( t \times 3 \) smoothing radius. The proposed method eliminates all nonstationary aliasing and improves the continuity of the curved event.

Abma and Kabir (2005) present a comparison of several algorithms used for trace interpolation. We chose the most challenging benchmark Marmousi example from Abma and Kabir to illustrate the performance of RNA interpolation. Figure 3a shows a zero-offset section of the Marmousi model, in which curved events violate the assumptions common for most trace-interpolating methods. Figure 3b shows that our method produces reasonable results for both curved and weak events and does not introduce any
undesirable noise. The adaptive PEF parameters correspond to 7 (time) × 5 (space) coefficients for each sample and a 40-sample (time) × 30-sample (space) smoothing radius.

Missing-trace interpolation test

A missing trace test is shown in Figure 4a and comes from decimated-trace interpolation result (Figure 2c) after removing 70% of randomly selected traces. The curved event makes it difficult to recover the missing traces. The interpolated result is shown in Figure 4b, which uses a regularized adaptive PEF with 4 (time) × 2 (space) coefficients for each sample and a 50-sample (time) × 10-sample (space) smoothing radius. In the interpolated result, it is visually difficult to distinguish the missing trace locations, which is an evidence of successful interpolation. The filter size along space direction needs to be small to generate enough regression equations.

FIELD DATA EXAMPLES

We use a set of marine 2D shot gathers from a deepwater Gulf of Mexico survey (Crawley et al., 1999; Fomel 2002) to further test the proposed method. Figure 5 shows the data before and after subsampling in the offset direction. The shot gather has long-period multiples and complicated diffraction events caused by a salt body. Amplitudes of the events are not uniformly distributed. Subsampling by a factor of two (Figure 5b) causes visible aliasing of the steeply dipping events. We designed a nonstationary PEF, with 15 (time) × 5 (space) coefficients for each sample and a 50-sample (time) × 20-sample (space) smoothing radius to handle the variability of events. Figure 6 shows the interpolation result and the difference between interpolated traces and original traces plotted at the same clip value. The proposed method succeeds in the sense that it is hard to distinguish interpolated traces from the interpolation result alone. A close-up comparison between the original and interpolated traces (Figure 7) shows some small imperfections. Some energy of the steepest events is partly missing. Coefficients of the adaptive PEF are illustrated in Figure 8, which displays the first coefficient \( B_1 \) and the mean coefficient of \( B_n \), respectively. The filter coefficients vary in time and space according to the curved events. The interpolated results are relatively insensitive to the smoothing parameters.

For a missing trace interpolation test (Figure 9a), we removed 40% of randomly selected traces from the input data (Figure 5a). Furthermore, the first five traces were also removed to simulate traces missing at near offset. The adaptive PEF can only use a small number of coefficients in the spatial direction because of a small number of fitting equations (where the adaptive PEF lies entirely on known data). However, it also limits the ability of the proposed method to interpolate dipping events. We used a nonstationary PEF with 4 (time) × 3 (space) coefficients for each sample and a 50-sample (time) × 10-sample (space) smoothing radius to handle...
the missing trace recovery. The result is shown in Figure 9b. By comparing the results with the original input (Figure 5a), the missing traces are interpolated reasonably well, except for weaker amplitude of the steeply dipping events.

An extension of the method to 3D is straightforward and follows a two-step least-squares method with 3D adaptive PEF estimation. We use a set of shot gathers as the input data volume to further test our method (Figure 10a). We removed 50% of randomly selected traces and five near offset traces for all shots (Figure 10b). For comparison, we used PWD to recover the missing traces (Figure 11a). The PWD method produces a reasonable result after carefully estimating dip information, but the interpolated error is slightly larger in the diffraction locations. (Figure 11b). The additional direction provided more information for interpolation but also increased the number of zeros in the mask operator $K(t,x)$, which constrains enough fitting equations in equation 13. To use the available fitting equations for adaptive PEF estimation, we chose a smaller number of coefficients in the spatial direction. The proposed method is able to handle conflicting dips, although it does not appear to improve the dipping-event recovery compared to the 2D case. This characteristic partly limits the application of RNA in the 3D case. We used a 3D nonstationary PEF with $4 \times 2 \times 2$ coefficients for each sample and a 50-sample smoothing radius was used in each direction.

Figure 5. A 2D marine shot gather. (a) Original input and (b) input subsampled by a factor of 2.

Figure 6. (a) Shot gather after trace interpolation (adaptive PEF with $15 \times 5$) and (b) difference between original gather (Figure 5a) and interpolated result (Figure 6a).
Figure 7. (a) Close-up comparison of original data and (b) interpolated result by RNA.

Figure 8. Adaptive PEF coefficients. (a) First coefficient $B_1$ and (b) mean coefficient of $B_n$.

Figure 9. (a) Field data with 40% randomly missing traces and (b) reconstructed data using RNA.
CONCLUSIONS

We have introduced a new approach to adaptive prediction-error filtering for seismic data interpolation. Our approach uses regularized nonstationary autoregression to handle time-space variation of nonstationary seismic data. We apply this method to interpolating seismic traces beyond aliasing and to reconstructing data with missing and decimated traces. Experiments with benchmark synthetic examples and field data tests show that the proposed filters can depict nonstationary signal variation and provide a useful description of complex wavefields having multiple curved events. These properties are useful for applications such as seismic data interpolation and regularization. Other possible applications may include seismic noise attenuation.

ACKNOWLEDGMENTS

We are grateful to Raymond Abma for providing benchmark data sets. We thank Eric Verschuur, Bill Curry, and two anonymous reviewers for helpful suggestions, which improved the quality of the paper. This work is supported by National Natural Science Foundation of China (Grant No. 41004041) and 973 Programme of China (Grant No. 2009CB219301). This publication was authorized by the Director, Bureau of Economic Geology, The University of Texas at Austin.

REFERENCES