3D seismic data interpolation via spatially localized f-xy prediction

Xishuo Wang, Ye Zheng and Mike Perz*
Geo-X Processing, Divestco Inc. Calgary, Canada

Summary

We have developed a new method for interpolating 3D prestack land seismic data. We employ a spatially localized frequency domain approach which operates on a data subset comprising traces with similar inline and crossline offsets and which performs spatial prediction over the midpoint coordinates. The approach provides flexibility in output data specification while utilizing the input data in a locally optimized way. For every single prestack target trace to be interpolated, all input traces proximal (in offset-midpoint coordinates) to the target trace are first subjected to certain selection criteria in order to derive an optimal subset of input data. Next these optimal data are transformed to the frequency domain and at each discrete frequency a 2D spatial prediction filter is estimated using existing traces within the subset. This filter is then used to predict data at the requested interpolated trace position. The filter size may vary spatially across the survey, expanding in areas associated with large gaps and shrinking in relatively well-sampled zones, Algorithm efficacy is assessed via extensive real and synthetic data testing.

Introduction

3-D land seismic data are often sparsely sampled in an attempt to reduce acquisition cost. In addition, surface obstacles can create irregular gaps within the survey. It is well understood that these shortcomings in spatial sampling can compromise final image quality, most notably at the stage of prestack migration, but also at other points in the processing. In recent years much interest has been cast on various techniques for data interpolation and regularization in an effort to compensate for imperfections in spatial sampling.

Spitz (1991) introduced a method of upsampling 2D data by an integer factor in the f-x domain, and this was extended to 3D (i.e. f-xy) by Huard et al. (1996). In our present work, we seek to *regularize*, rather than explicitly *upsample,* 3D prestack data. This helps to reduce artifacts in prestack migration and, in our most recent experience, helps to improve data sampling for AVAZ/VVAZ analysis. In contrast to certain "global" approaches to the interpolation problem, our algorithm interpolates one seismic trace at a time, thereby allowing great flexibility in specifying shot and receiver positions within the output interpolated volume. Our last work in this area was a time domain interpolation approach which involved fitting each time sample by a 4D polynomial (Wang, 2004). This approach is fast and avoids data gridding, but it cannot interpolate highly structured data. In this work, we develop a new interpolation kernel in the frequency domain which is capable of interpolating spatially aliased data.

We choose to describe the full 5-D prestack data volume in frequency-offset-midpoint coordinates. While in principle, spatial prediction could be done in any number of dimensions--at most in 4D or at least in 1D-- the higher the dimension count the larger the filter size and the more costly the algorithm. In our present approach we perform 2D prediction in CMP x and y while fixing offset x and offset y (i.e., inline and crossline offset) in order to interpolate a given target prestack trace. Of course in the case of irregularly sampled land data, input data usually do not exist exactly where we want them to be, and therefore we must employ some acceptance criteria for deeming which input traces ought to contribute to the current interpolation operation as discussed below.

2D frequency domain prediction

Let $A(f,x_k,y_k,x_d,y_d)$ denote a trace at mid-point (x_k,y_k) with offset vector $\boldsymbol{v}=(x_d,y_d)$ and frequency f. For locally linear seismic signals comprising planar dipping events, the trace can be approximated by 2D prediction in the CMP (x_k,y_k) plane:

$$A(f,x_k,y_k,x_d,y_d) = \sum_{i=-n_x}^{n_x}\sum_{j=-n_y}^{n_y} F_{ij}A(f,x_k-id_x,y_k-jd_y,x_d,y_d)\,, \qquad (1)$$

where F_{ij} is the 2D prediction filter (and therefore F_{00} is not included in the summation), d_x and d_y are CMP intervals in x and y direction and n_x and n_y are one-sided maximum filter lags in the x and y directions, respectively.

Eqn.(1) also forms the basis of the Spitz (1991) approach to interpolation, but as mentioned earlier, there are important differences between our respective algorithms. His method is global, and exploits the regular sampling of the input data volume in order to make it denser by an integer factor, resulting in an efficient but less flexible algorithm. By contrast, our approach requires good *local* data distribution and aims to regularize rather than upsample; as such it is flexible but less efficient.

In order to use the above scheme to interpolate a trace at $A(f,x_k,y_k,x_d,y_d)$ we need known data with the same offset (x_d,y_d) and CMP locations around (x_k,y_k) on the CMP grid of $(x_k \pm N_x, y_k \pm N_y)$, where N_x, *and* N_y define the input data block size ($N_x \geq n_x$ and $N_y \geq n_y$ to guarantee an overdetermined least squares system in solving the filter coefficients), and the maximum number of equations is $N_{eq\_max}=(2N_x+1)(2N_y+1)$ (i.e., the case where input data exists at all CMP locations within the grid). Setting up equation (1) with data CMP index k ranging through all the input data grid of N_{eq} points, we get N_{eq} equations to solve for the $N_{flt}=(2n_x+1)(2n_y+1)-1$ filter coefficients. With the filter solved, we then use equation (1) to interpolate the data at one frequency. Missing input data on the grid will degrade the filter solution, and this gets worse when input data becomes one-sided in any direction.

Input trace selection for prediction and QC

We will seldom get input data exactly as desired. The imperfection in input data sampling in the neighborhood of the interpolated trace is measured by two QC quantities.

First, the CMP misposition of one input trace relative to the *locally aligned* CMP grid is given by

$$\Delta^2_{k,ij} = [x_{input} - (x_k - id_x)]^2 + [y_{input} - (y_k - jd_y)]^2, \quad (2)$$

and secondly, the discrepancy between the offset vectors of input and interpolated traces, the *offset vector focusing factor,* is given by

$$Q_{fcs} = 1 - \frac{|\mathbf{v} \bullet \mathbf{v}_k|}{\max(|\mathbf{v}|^2, |\mathbf{v}_k|^2)}, \quad (3)$$

in which, $\boldsymbol{v}$ and $\boldsymbol{v}_k$ are the offset vectors of the interpolated and the input trace, respectively.

Note that the absolute value of the dot product in the numerator of eqn. (3) "honours" the reciprocity principle, since exchanging the source and receiver location of an input trace simply reverses the sign of the offset vector, which in turn will not change Q_{fcs}. Note also that the denominator of the ratio in (3) is always greater than or equal to the numerator, and therefore the dimensionless Q_{fcs} is always non-negative and between 0 (corresponding to $\boldsymbol{v}=\pm\boldsymbol{v}_k$) and 1 (corresponding to $\boldsymbol{v} \perp \boldsymbol{v}_k$). Just as in the case of Δ^2_{cm}, the smaller the value for Q_{fcs} the better, and the interpolation is "perfect" when both Δ^2_{cm} and Q_{fcs} are zero .

Input traces are selected on the $(2N_x+1)$ by $(2N_y+1)$ grid centered around the interpolated trace CMP position, and at each input grid point, the trace with the smallest Δ^2_{cm} and Q_{fcs} is chosen. The average of Δ^2_{cm} and that of Q_{fcs} over the $(2N_x+1)(2N_y+1)$ input grid points are noted in the trace header of this interpolated trace. These QC trace headers may be reviewed post-interpolation in order to "red flag" any interpolated output of questionable reliability.

Spatially variable filter size to accommodate irregular input sampling

Due to irregularities in the spatial sampling (i.e., holes, missing data etc) some filter lags, particulaly the small ones, may not be resolvable because they lack input data support and/or they may not contribute to the prediction of the target output trace. One way to avoid this pitfall is to simply increase the filter size (larger n_x and n_y) in a global sense, but that is costly and may degrade the lateral resolution in well sampled areas. Our approach is to allow the filter size to vary spatially across the survey, expanding in areas of poor sampling and shrinking in the well-sampled areas. The implementation entails specifying a large "default" filter, but truncating this filter in a local sense wherever the input sampling characteristics permit. The truncation is done in such a way that higher priority is given to small lags and to those lags which give rise to an even distribution among the four lag quadrants (i.e., +/-ve lags in xy).

Synthetic data tests

We generated constant velocity synthetic data corresponding to a real 3D survey geometry. These data contain a steeply dipping event which is spatially aliased in the midpoint-x direction. The interpolation algorithm was used to increase the number of shot and receiver lines (i.e., reduce the shot and receiver line spacing), while keeping shot and receiver intervals fixed. Figs.1 and 2 show an interpolated partial shot gather after constant velocity NMO correction (note the signficant residual moveout associated with the steeply dipping event). Interpolation by the aforementioned time domain polynomial method (Fig.1) fails somewhat and the frequency domain (Fig.2) method is better. Examination of the QC headers (not shown here) shows that the interpolation is quite robust with respect to imperfect offset focusing as some of the well aligned linear events in Fig.2 revealed offset vector focusing factors as high as 0.9. We repeated this synthetic test using both single and multiple (five) temporal windows and found that the use of multiple windows gave a marginally better result.

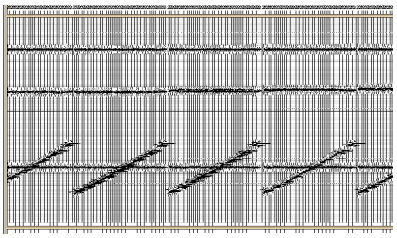

Figure 1: 4D polynomial interpolated partial shot gather

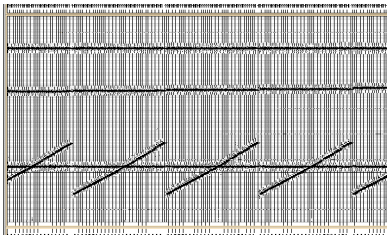

Figure 2: Localized f-xy interpolated partial shot gather

Real data tests

The first real data example concerns merging one sub-volume into another larger 3D data set. The combined 3D has a 30 m x 30 m subsurface grid, as dictated by the sampling characteristics of the larger 3D data set (the "processing grid"), but the natural subsurface grid of the sub-volume is 25 m x 25 m (the "native grid"). Fig 3 shows the mismatch between the actual source-receiver midpoints and the processing grid. Note the pattern of the fold essentially repeats itself every five grid points along an inline (inlines are vertical lines on this plot). When the sub-volume was prestack time migrated, systematic artifacts appeared (Fig 4) which exhibit a similar periodicity of five bins (i.e. 150 m) along an inline. Because 150 is the minimum common multiple of 25 and 30, we hypothesized that the problem was caused by the discrepancy between processing grid and native grid.

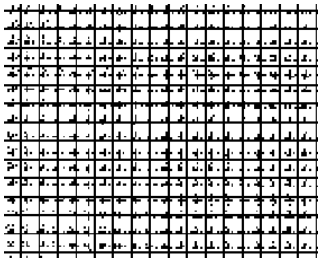

Figure 3: Midpoint scatter map with 30m bin grid

In order to test this hypothesis, we interpolated the real data of the sub-volume onto a coarser grid of 30 m in order to match the processing grid and then re-ran the prestack time migration. We found that the artifacts were significantly reduced (Fig. 5), strongly suggesting that the problem is caused by the conflict of grid sizes. To further confirm our suspicion, we generated a synthetic data set with some flat layers based on the original shot and receiver locations (native grid) and binned the data to the 30 m processing grid. Then, prestack time migration was applied to this synthetic data set (not shown here). Again, we saw strong, periodic artifacts, similar to what was observed on Fig. 4 of the real data. Next, we took these synthetic data set interpolated them onto the coarser 30 m grid, just as we did for the real data . As in the real data case, the artifacts were greatly reduced (not shown here).

From these tests, we can definitively conclude that the repeating artifacts on Fig. 4 is caused by the mismatch between native and processing grids and that prestack interpolation can help solve this problem.

The second test data set is collected over a potash mine whose many known mine entry locations provide a good test of the post-interpolation lateral and vertical resolution. Fig. 6 shows a time slice of the poststack-migrated section of the original data at the mine tunnel level with the known positions of the mine tunnel overlaid in black. Note that these data are well sampled and have produced an excellent image of the mine entries.

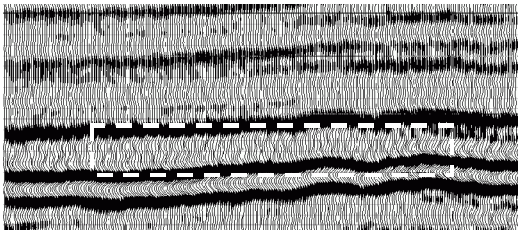

Figure 4: Inline of original sub-volume after PSTM (25m bin)

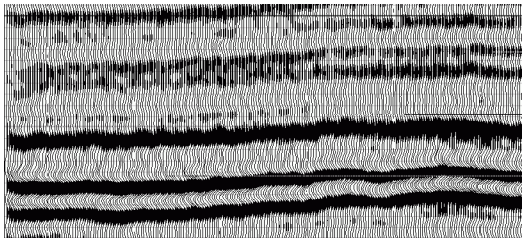

Figure 5: Inline of sub-volume after prestack interpolation to 30m bin size followed by PSTM

Fig.7 shows the same time slice generated with a decimated data set in which all even numbered shot gathers were

killed prior to stacking and poststack migration. The loss of resolution is obvious. Fig.8 shows the result of interpolating the decimated data back to the original geometry. Comparing figs. 6 and 8, we feel that the interpolation has faithfully restored the resolution of the original well-sampled data. Time domain polynomial method did not perform as good (though not shown here).

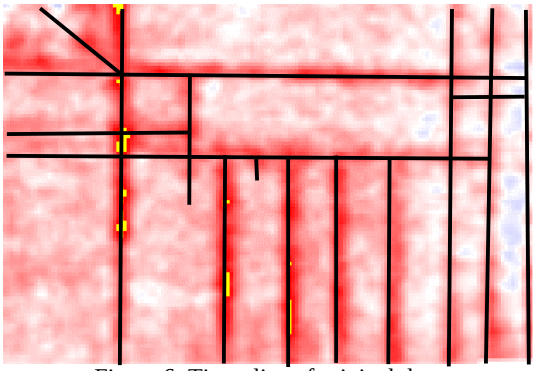

Figure 6: Time slice of original data

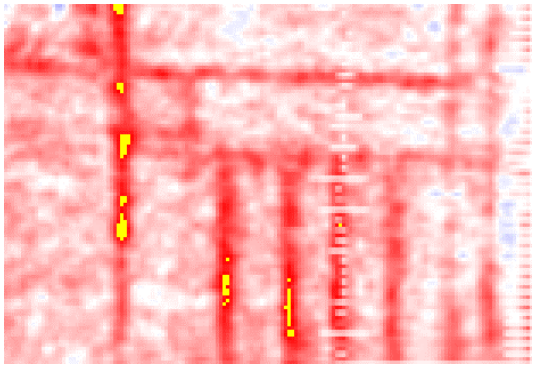

Figure 7: Time slice of 50% decimated data

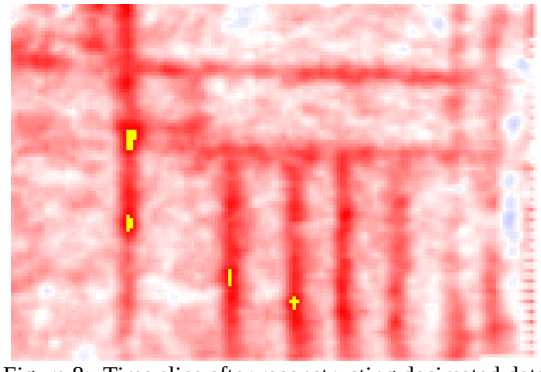

Figure 8: Time slice after reconstructing decimated data via localized f-xy interpolation

Conclusions

Localized interpolation offers ultimately flexible output volume specification and also facilitates the optimization of input data selection. Synthetic and real data tests suggest that the prediction interpolation tolerates a significant degree of irregularity in the input sampling, and attributes which quantify the imperfection of the data sampling serve as a useful QC tool. Spatially aliased data can be effectively interpolated (assuming the events show quasi-planar behavior in the midpoint coordinates). Truncation of filter size saves compute time while still allowing some degree of interpolation across relatively large gaps.

Real data tests show that the prediction method provides higher resolution over the time domain polynomial method, and that prestack interpolation can serve as a powerful tool to help reduce migration artifacts caused by irregular acquisition geometry.

Acknowledgement

We thank our anonymous clients for allowing us to use their data for this study. We thank Mark Ng for preparing the test data. Darren Betker, Earl Heather, Glenn Hauer (now at Aram), Leo Macht, Lyndsey Nicholas, Dan Norton, Wilf Reynish, Laurie Ross and Dick Yuen , carried out real data testing and application.

EDITED REFERENCES

Note: This reference list is a copy-edited version of the reference list submitted by the author. Reference lists for the 2007 SEG Technical Program Expanded Abstracts have been copy edited so that references provided with the online metadata for each paper will achieve a high degree of linking to cited sources that appear on the Web.

REFERENCES

Huard, I., S.Medina, and S. Spitz, 1996, F-XY wavefield de-aliasing for acquisition configurations leading to coarse sampling: 58th Annual Conference and Exhibition, EAGE, Extended Abstracts, B039.

Spitz, S., 1991, Seismic trace interpolation in the F-X domain, Geophysics, 56, 785–794.

Wang, X., 2004, 3D prestack seismic trace interpolation in time domain with input optimization: CSEG National Convention.