

Challenges for land data interpolation

Daniel Trad^{*}, Jeff Deere and Scott Cheadle, Veritas DGC Inc.

2005 CSEG National Convention



Summary

Advances in processing techniques have led to more stringent requirements on information content in the input seismic data. Although not a substitute for well sampled field data, interpolation has provided some useful data preconditioning that in particular allows migration techniques to work better.

Seismic data interpolation has been around for long time, but only recently have we been able to use complex multidimensional and global algorithms that have the capability to infill large gaps in 3D land surveys. This innovation brings a great potential for improvement but, for this technology to become useful, first many questions need answers: which are the best domains to interpolate? What is the optimal size of operators given a particular level of structural complexity? Should we pursue a ideal geometry for migration or should we stay close to the input geometry in order to minimize distortions? How does sampling in multiple dimensions affect our traditional aliasing constraints? How can we infill large gaps without using a model for our data? Are irregularities in sampling beneficial?

In this paper, we show some examples with Minimum Norm Weighted Interpolation, sparse DFTs and other related techniques, and we discuss some of these issues when trying to extend real data to well sampled migration-friendly data.

Introduction

Interpolation of land data presents many challenges, some of them quite different from those of interpolating marine data sets. In general, all 3D geometries have very poor sampling along one or more dimensions. As a consequence migration suffers from artifacts when applied to poorly sampled data, because migration algorithms work based on the principle of constructive and destructive interference.

There are many different approaches to attack this problem. The only perfect solution is to acquire well sampled data; all the other approaches attack the symptoms of the problem rather than the problem itself, and there is no guarantee that they will really solve it. However, given that in the real world often we cannot go back to the field and fix the actual problem, we have to address this issue using all the processing tools at our disposal.

It is important to realize that most seismic algorithms implicitly apply some sort of interpolation because they assume correct sampling. If a sample is missing the missing value is supposed to be either zero or similar to its neighbouring values. If this assumption is not good enough we have to apply a separate interpolation algorithm where explicit assumptions can be made, for example, assuming that frequencies beyond Nyquist are zero (sinc interpolation). This is what interpolation algorithms are useful for, preconditioning the data with a reasonable constraint.

There are many interpolation methods, some more successful than others and some more complex than others. As a general classification, methods are local or global (Figure 1). Local methods (for example triangulation, nearest neighbour, etc.) tend to be quite robust, fast, adaptable and easy to implement. Unfortunately, they cannot handle very large gaps, because either they need nonexistent local information (there is nothing around the trace to interpolate) or they do not have enough information to overcome aliasing. On the contrary, global methods are often slow, not very adaptable, and difficult to implement. However, they can fill in missing data by using information obtained from more distant data. This allows them to overcome aliasing and fill in very large gaps. Most practical methods will fall between these two extremes.

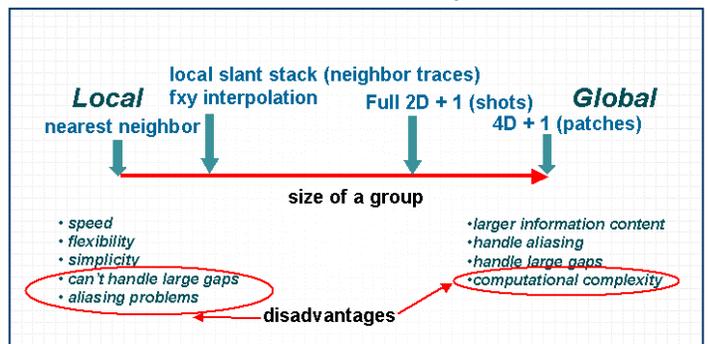


Figure 1 - Local vs. global interpolation methods.

A related distinction is the number of dimensions the algorithm can handle simultaneously. When applying one-dimensional methods the user has to decide in which order to interpolate the different dimensions. Often one of the dimensions is so poorly sampled that it becomes almost impossible to fill in the gaps. Multidimensional methods can, in principle, obtain information from a well sampled dimension to infill the poorly sampled dimension automatically.

Most well-known and accepted interpolation algorithms have been developed for marine surveys. Marine data usually have a very well sampled direction (inlines) and a very coarsely sampled direction (crosslines). Many algorithms have been quite successful in infilling the crossline direction, even in the presence of aliasing and complex structure. Most of these methods do not try to achieve full azimuthal coverage. In interpolation for land data we need to consider sampling along the offset and azimuth direction. This makes the problem more difficult than for marine data.

A few global multidimensional methods have been published recently. One of the most promising algorithms is Minimum Weighted Norm Interpolation (Liu and Sacchi, 2004). This algorithm can operate on the data in all dimensions simultaneously. For example, we have applied this method to four spatial dimensions (Liu et al. 2004). Although this new capability of filling in several dimensions offers a world of possibilities, it also creates numerous questions: Which dimensions to use? How large should the input be? How regular do these dimensions have to be? How to cope with efficiency issues in particular for irregular sampling? Indeed, for fully irregular data, expensive Discrete Fourier computations (as opposed to Fast Fourier Transforms) are required and four dimensional methods seem to be out of reach for now.

The issue of which dimensions to interpolate is critical. If we want to precondition the data properly for Kirchhoff migration, then a full and uniform coverage of offsets and azimuths would be desirable. Unfortunately, this would increase the size of the survey by an order of magnitude and make migration too expensive. In addition, it is well known that good results can be achieved with much less than perfectly sampled data, but it is not clear how sparse the data can be. By simulating full coverage surveys it is easy to see that they lead to geometries very different from what we typically acquire. Therefore, going from one to the other is a serious challenge. For example, Figure 2 shows original and interpolated receivers for a geometry where a constant (far) offset plane was interpolated to achieve full coverage in cmps (no azimuth coverage was attempted). The points represent receiver locations, the colour the shot to which the receiver belongs. The receiver locations created with interpolation tend to fill the space between the receiver lines that were acquired in the field.

A reasonable approach is to stay as close as possible to the original geometry. This implies infilling shot and receiver lines, while keeping the original traces unchanged. This method tends to improve on acquisition artifacts of large amplitude and wavenumber. Regularity can introduce short wavelength artifacts while irregularity seems to eliminate them. This leads to another question: How regular do we want the output data to be? It seems certain that some amount of irregular sampling is beneficial.

Methods

Some of the methods we currently use are:

Minimum Norm Weighted Interpolation MWNI (Liu and Sacchi, 2004).

The available data, \mathbf{d} , are the result of a picking matrix (sampling), \mathbf{T} , acting on an unknown fully sampled data set, \mathbf{m} . The unknown data set is subject to the constraint of that it has the same multidimensional spectrum as the original data. This constraint can be calculated with multidimensional Fast Fourier Transforms \mathbf{F}_{nd} . To solve for the unknown data set (the interpolated data set), a cost function is defined and minimized using standard optimization techniques. The cost function J is defined as follows:

$$J = \|\mathbf{d} - \mathbf{Tm}\|^2 + \lambda \|\mathbf{m}\|_w,$$

with a norm calculated as:

$$\|\mathbf{m}\| = \mathbf{m} \mathbf{F}_{nd}^{-1} |p_k|^{-2} \mathbf{F}_{nd} \mathbf{m}.$$

\mathbf{F}_{nd} is the multidimensional Fourier transform and ND is the dimension of the data with ND=2, 3 or 4 for 2D data, common azimuth, or 3D data, respectively. p_k is the spectrum of the unknown data, obtained by bootstrapping or iterations. This method is very fast in spite of being very global, because it uses FFTs. The main drawback is that the true sampling of the data has to be an integer multiple of the required final sampling along each dimension. This is equivalent to binning the data along the chosen dimensions.

Sparse DFT (Duijndam et al., 1999),

This method is also based on optimization and consists of:

- 1) Calculating the multidimensional spectrum of the data subject to a constrain of sparseness. This is, as above, achieved by defining a cost function and minimizing it:

$$J = \|\mathbf{d} - \mathbf{Fm}\|^2 + \lambda \|\mathbf{m}\|_w$$

A difference from the previous case is that the operator \mathbf{F} is itself a multidimensional Fourier transform. In practice only two dimensions are used, because this operator can become very large.

- 2) Predicting the data from the model: $\mathbf{d} = \mathbf{Fm}$.

We implemented this method using discrete Fourier transforms with the purpose of taking into account the sampling interval. In principle, this method is intended to deal with very irregular data where the assumptions in MWNI are violated. Alternatively, this method can be used as a first regularization of the data to a grid, without attempting to infill them densely. Once the data have been moved to a regular grid, MWNI can be used to upsample it to a denser grid.

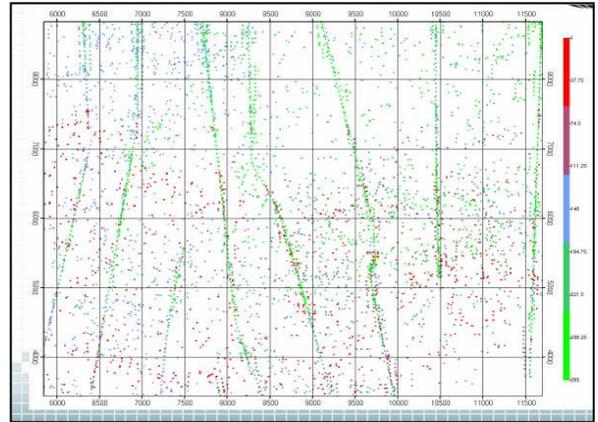


Figure 2 - Receiver positions for full CMP coverage in one offset range.

Antileakage Fourier Reconstruction (Xu et. al., 2004): this method has some similarities with the sparse DFT, but the model is obtained by iteratively eliminating the non-orthogonality that arises from irregularly sampled Fourier transforms. It has been used successfully in complex marine environments and currently we are adapting it to land interpolation (Xu et.al. 2005).

Fxy interpolation : This method assumes that the data are locally linear, which is a valid assumption as long as the size of the operators is small. As a consequence of using this model, aliasing can be successfully overcome, although big gaps are difficult to infill. Several passes of this method can be used to progressively infill data. In our experiments we use several passes along inlines and crosslines in the cmp domain.

Local methods: We are also experimenting with simpler local methods like triangulation or local tau-p transforms since they are very reliable for the interpolation of small gaps.

Examples

Here we show some examples with MWNI, leaving more complete examples for the actual presentation. Figure 3 shows a MegaBin survey (provided by Encana). In this case an extra line of receivers is added to reduce the sparse crossline sampling.

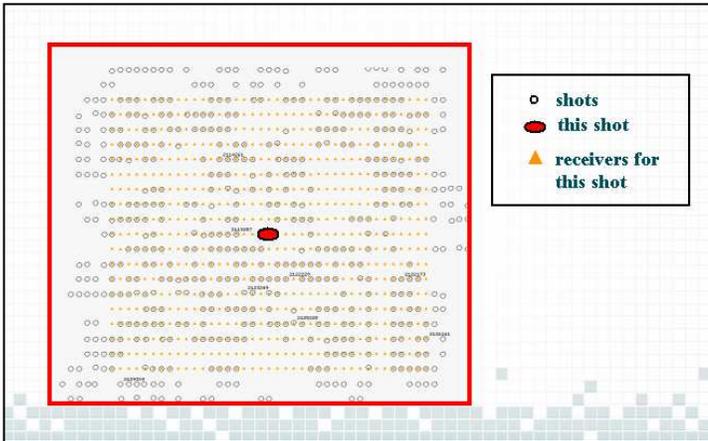


Figure 3 - Shot patch for Megabin survey before interpolation.

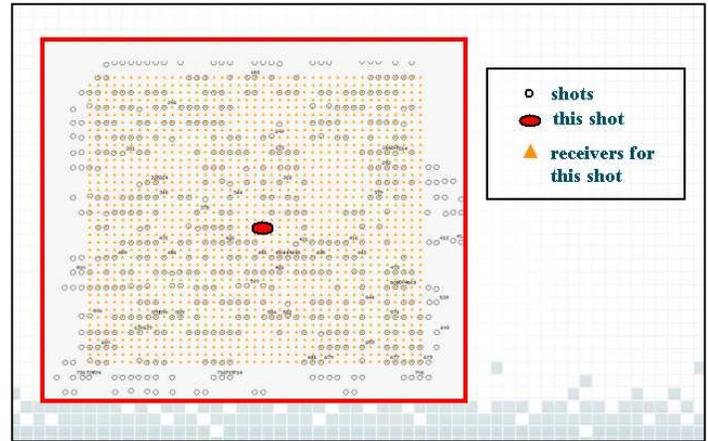


Figure 4 - Shot patch after interpolation.

In this example, only receivers have been interpolated but a two pass approach is sometimes used where we first interpolate receivers, then we sort to receiver gathers and interpolate shots. Interpolating both shots and receivers has the advantage of filling the near offsets but it can lead to data with a higher level of smoothing, since the second pass uses interpolated data to create more interpolated data. There are many unanswered questions in this regard.

Figures 5 and 6 and show a view of a shot before and after interpolation. Notice that big gaps in one direction may be smaller in the other (orthogonal direction). This is what makes multidimensional interpolation very powerful in filling large gaps.

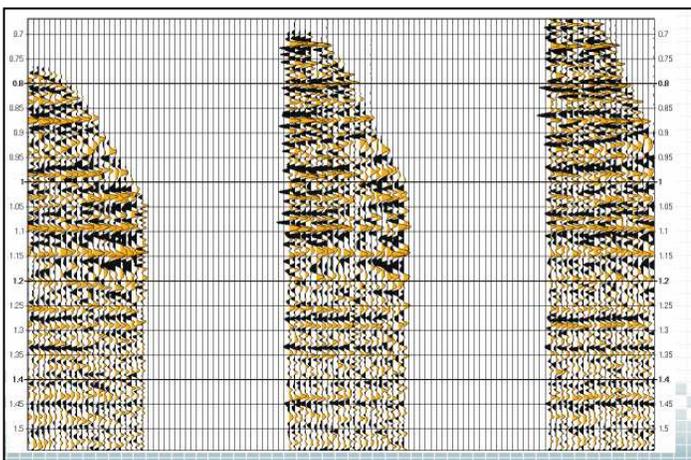


Figure 5 - Shot before interpolation.

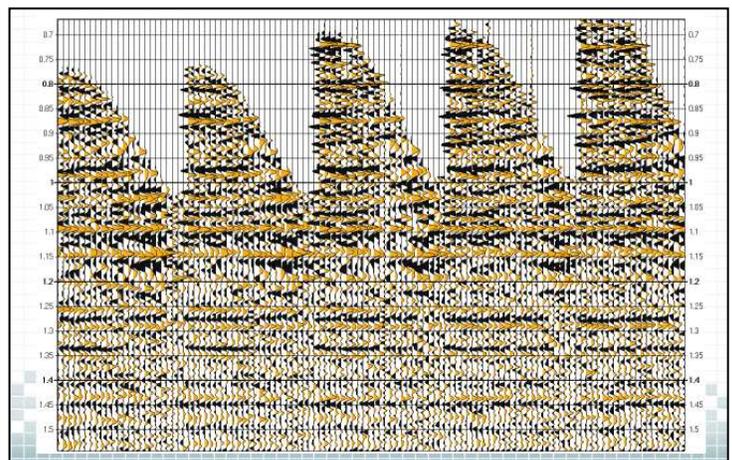


Figure 6 - Shot after interpolation.

Figures 7 and 8 show a time slice of the stack from a MegaBin survey where two receiver inlines and one receiver crossline have been inserted between existing ones. The initial been size was reduced from 35x70m to 17.5x17.5 m.

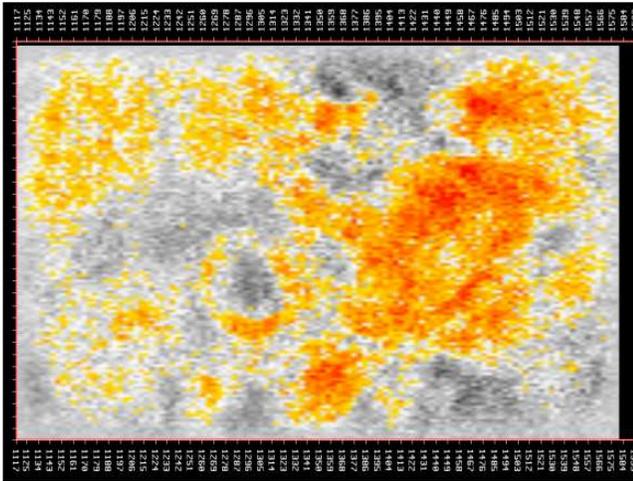


Figure 7 - Time slice from stack of original data.

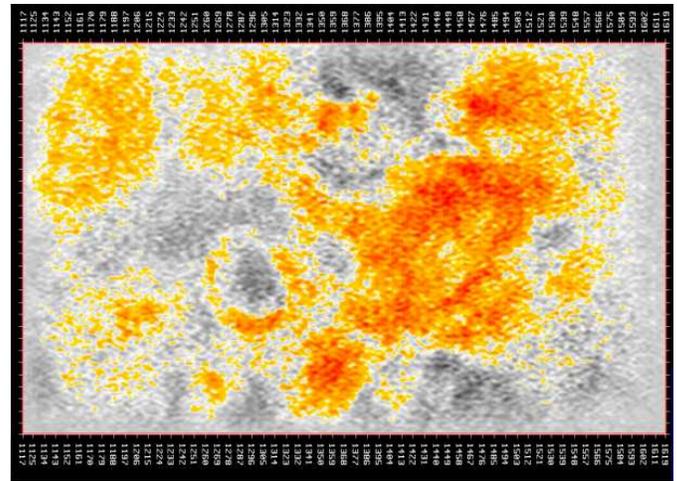


Figure 8 - Time slice from stack of interpolated data.

Conclusions

3D land geometries are always subsampled along one or more dimensions. As seismic processing becomes more demanding in terms of analyzing prestack data in detail, interpolation has become a very useful tool to condition the data for migration. Given that seismic data live in five dimensions and there are different choices for them, interpolation is far from being a solved problem. The perfect input data set for Kirchhoff migration may contain full cmp coverage for all offset ranges, and uniform azimuth distribution, but this is far cry from the data obtained with typical acquisition geometries. Therefore, there is a conflict between the goal of a perfect data for migration and preserving as much as possible from the original data.

New interpolation algorithms allow us to look at several dimensions at once. This has a large number of benefits but also unsolved problems, particularly if the method is global. On the other hand, relaxing the goal of globality, the method may fall into the problem of not having enough input to predict the output, in particular for large gaps. An optimal solution is somewhere in between global and local methods.

In this paper we have discussed several approaches MWNI, sparse DFT and other methods, but more than the algorithm used to calculate the interpolated traces, it seems that the more important issue to decide is the domain where to apply it, and the best choice in terms of size of operators. Finally, but still important, is the question of exactly how much data we need to interpolate, where to put them, and how much we can trust them.

These decisions are related to considerations of aliasing, complexity of the structure in different domains, efficiency, capability to manage large gaps and very irregular sampling. We are just starting to scratch the surface of a difficult problem. The more we learn about it, the more complex it seems.

Acknowledgments

We would like to acknowledge Veritas for permission to publish this paper, Encana for the data sets shown in this abstract, and all clients who have encouraged for this research

References

- Duijndam, A. J. W., Schonewille, M. A., and Hindriks, C. O. H., 1999, Reconstruction of band-limited signals, irregularly sampled along one spatial direction: *Geophysics*, **64**, no. 2, 524-538.
- Liu, B. and Sacchi, M. 2004, Minimum Weighted Norm Interpolation of Seismic Records, *Geophysics* **69**, 6,
- Liu, B., Sacchi, M.D., and Trad, D., 2004, Simultaneous interpolation of 4 spatial dimensions, 74th Ann. Internat. Mtg: Soc. of Expl. Geophys., Expanded Abstracts (SP 3.6).
- Xu, S., Zhang, Y., Pham, D. and Lambaré, G., 2004, On the orthogonality of anti-leakage Fourier transform based seismic trace interpolation. 74th Ann. Internat. Mtg: Soc. of Expl. Geophys., Expanded Abstracts (SP 3.7).
- Xu, S., Zhang, Y. and Lambaré, G., 2005, Recovering densely and regularly sampled 5D seismic data for current land acquisition. 67th Ann. Internat. Mtg., EAGE, Expanded Abstracts