3D seismic data interpolation

Xishuo Wang, Geo-X systems Ltd., Calgary, Canada

Introduction

Conventional 3D seismic pre-stack data is usually gathered to CDP bins of size ($\Delta x, \Delta y$) in (x,y) space. Traces from shot point (x_s, y_s), receiver point (x_r, y_r), and with mid-point $x_m=(x_s+x_r)/2$ & $y_m=(y_s+y_r)/2$ falling within a bin will be stacked together to generate the stacked trace for that bin. The underlying assumption for this process is that the gathered traces to be stacked are virtually the same because of the small size of ($\Delta x, \Delta y$). When the size of bins increases or data has complex structure, the difference among these to-be-stacked traces increases. It is therefore desirable to generate stacked traces by way of interpolation in order to increase resolution.

Pre-stack data distribution is often less than perfect due to limitations in any particular survey design. Certain processing steps demand better data distribution than what is already collected, e.g., AVO analysis requires more uniform offset distribution, and pre-stack migration works better with densely and regularly distributed gathers. To achieve better data distribution with a denser survey grid can be costly. It is economical to do pre-stack data interpolation based on an “acceptable” input 3D data set. I'll explore interpolation of pre-stack 3D data and stacking by way of interpolation.

Interpolation to bin center as means of stacking

Pre-stack seismic data can be denoted as $A(x_s, y_s, x_r, y_r, t)$, or amplitude at time t, of a trace with source position (x_s, y_s) and receiver position (x_r, y_r). The time sampling is usually regular, $t=k\delta t$, $k=1,2,...n$, while the shooting geometry may have some irregularities. This notation contains all the sampling (time & space) geometry. An alternative notation is $A(x_m, y_m, t)$, where $x_m=(x_s+x_r)/2$, $y_m=(y_s+y_r)/2$, is the mid-point of the trace with source position (x_s, y_s) and receiver position (x_r, y_r). Note, however, (x_m, y_m) does not always coincide with a CDP bin center. This second notation, compared with the first, having lost all azimuthal and offset information, is much less complete, but this is all the geometrical information the conventional CDP stack demands. The third notation, $A(x_m, y_m, d_m, t)$, adds offset of a trace as an extra geometric parameter, with $d_m = \sqrt{[(x_s-x_r)^2+(y_s-y_r)^2]}$ as the source-receiver offset.

From these three notations, the interpolation algorithm can be formed in three to five dimensional space, if we handle one time slice at a time. Polynomials are chosen as the interpolation function for their simplicity. Moreover, conventional stack can be considered as a special case of such, as zero order polynomial in any of these spaces.

Interpolation in mid-point space to bin-center

In (x_m, y_m) space, the interpolation polynomial has the following form:

$$p(x_m, y_m, \boldsymbol{a}) = \sum_{i=0}^{N_x} \sum_{j=0}^{N_y} a_{ij} x_m^i y_m^j$$

in which, vector $\boldsymbol{a}$ represents the full set of polynomial coefficients (a_{ij}). With input data at (x_m, y_m), $m=1,2,\ldots N$, polynomial coefficients $\boldsymbol{a} = (a_{ij})$ can be solved by minimizing:

$$E^2 = \sum_{m=1}^{N_d} [A(x_m, y_m) - p(x_m, y_m, \boldsymbol{a})]^2$$

To ensure stability, order of polynomial, N_x & N_y, should be chosen such that the length of $\boldsymbol{a}$ is less than N_d. Substitute bin center coordinates (x_0, y_0) into polynomial $p(x_m, y_m, \boldsymbol{a})$, we get the interpolated amplitude at bin center as,

$$A(x_0, y_0) = p(x_0, y_0, \boldsymbol{a})$$

Interpolation in mid-point space to bin-center with AVO modeling

Expand the polynomial in the section above into (x_m, y_m, d_m) space, to include terms in the conventional AVO form, we have the interpolation polynomial modified to:

$$p(x_m, y_m, d_m, \boldsymbol{a}, \boldsymbol{b}) = \sum_{i=0}^{N_x} \sum_{j=0}^{N_y} (a_{ij} + b_{ij} d_m^2) x_m^i y_m^j \equiv p_0(x_m, y_m, \boldsymbol{a}) + d_m^2 p_2(x_m, y_m, \boldsymbol{b})$$

with intercept polynomial $p_0(x_m, y_m, \mathbf{a})$, and gradient polynomial $p_2(x_m, y_m, \mathbf{b})$:

$$p_0(x_m, y_m, \boldsymbol{a}) = \sum_{i=0}^{N_x} \sum_{j=0}^{N_y} a_{ij} x_m^i y_m^j$$

$$p_2(x_m, y_m, \boldsymbol{b}) = \sum_{i=0}^{N_x} \sum_{j=0}^{N_y} b_{ij} x_m^i y_m^j$$

And at the bin-center, zero-offset stack and AVO gradient is interpolated as,

$$p(x_0, y_0, d_m = 0, \boldsymbol{a}, \boldsymbol{b}) = p_0(x_0, y_0, \boldsymbol{a}) = \sum_{i=0}^{N_x} \sum_{j=0}^{N_y} a_{ij} x_0^i y_0^j, \quad \text{and}$$

$$p_2(x_0, y_0, \boldsymbol{b}) = \sum_{i=0}^{N_x} \sum_{j=0}^{N_y} b_{ij} x_0^i y_0^j .$$

With the same order N_x & N_y, both polynomial coefficients vectors, $\boldsymbol{a}$ and $\boldsymbol{b}$ have the same number of elements as in the last section. The number of unknowns double compared to the case which does not model AVO effects.

Interpolation in 4D space of source & receiver locations

In 4D space (x_s,y_s, x_r,y_r), the interpolation polynomial has the following form:

$$p(x_s, y_s, x_r, y_r, \boldsymbol{c}) = \sum_{i=0}^{N_{sx}} \sum_{j=0}^{N_{sy}} \sum_{k=0}^{N_{rx}} \sum_{l=0}^{N_{ry}} c_{ijkl} x_s^i y_s^j x_r^k y_r^l$$

Since the space (x_s,y_s, x_r,y_r) contains all geometrical information of input/output positions, it reflects source & receiver surface locations, source/receiver offset and azimuth. With high enough order of the polynomial it can model azimuthal structural response and AVO effects beyond gradient. The size of polynomial coefficient vector $\boldsymbol{c}$ grows fast with increased order. For all N's equal 1, 2 & 3, for example, the length of $\boldsymbol{c}$ is 16, 81 & 256.

Estimate equivalent AVO gradient from interpolated & conventional stack

At a fixed time sample, conventional stack, G, is a weighted average of amplitudes of all input traces:

(1) $$G = \sum_i w_i A(d_i) / \sum_i w_i$$

where, $A(d_i)$ is the amplitude of a trace at offset d_i and w_i its weight. The AVO model with simple gradient term assumes the amplitude as function of offset in the form:

(2) $$A(d) = a_0 + a_2 d^2$$

in which, a_2 is the AVO gradient, and a_0 is the intercept or zero offset stack that can be assumed to be the same as the interpolated stack at zero offset.

Substitute (2) into (1), we have

(3) $$a_2 = (a_0 - G)/E(d^2)$$

in which $E(d^2)$ is the weighted average of offset squared:

$$E(d^2) = \sum w_i d_i^2 / \sum W_i$$

In a simple case of no weighting and that traces in the gather are uniformly distributed from zero to maximum offset d_{max}, we have,

$$(4) \quad E(d^2) = \int_0^{d_{\max}} x^2 dx / d_{\max} = d_{\max}^2 / 3$$

Substitute (4) into (3), we have,

$$(5) \quad a_2 = 3(a_0 - G) / d_{\max}^2 = 3\,(\textit{Interpolated\_stack - conventional\_stack}) / d_{\max}^2$$

Numerical model test

A single 15-degree dipping reflector with known velocity over it is tested with some additive random white noise. The choice of such a simple model is to reveal the ability of interpolation in resolving such. Bear in mind that such a model can be viewed as a building block of seismic traces in general, the test results are revealing for a broad range of real data cases. The model also has an AVO gradient of $2*10^{-4}/m^2$ with maximum offset at 2000m.

I tested only the more complex algorithm, interpolation in (x_s, y_s, x_r, y_r) space. With polynomial orders increasing, the interpolation results improve. Even the lowest order of polynomial is better than zero order (conventional stack or Gaussian stack). The model input data has ideal distribution: S/R azimuth, and mid-point scatter from bin center are random. Offset distribution is random while following a trend of linearly increasing fold with offsets, which is characteristic of a uniform receiver grid.

Conclusions

Estimating stacked sections and AVO effects by interpolation is a natural extension of the existing conventional method. The latter is a special case of polynomial interpolation, i.e., zero order polynomial. Model test showed that the interpolation offers higher resolution and the ability to resolve the structure. Interpolation of pre-stack data is also possible with modestly structured data.

Acknowledgement

I am grateful to Geo-X Systems Ltd. for allowing me to publish this work. I thank my colleague Oliver Kuhn & Brent Sato for reading the manuscript and making suggestions.

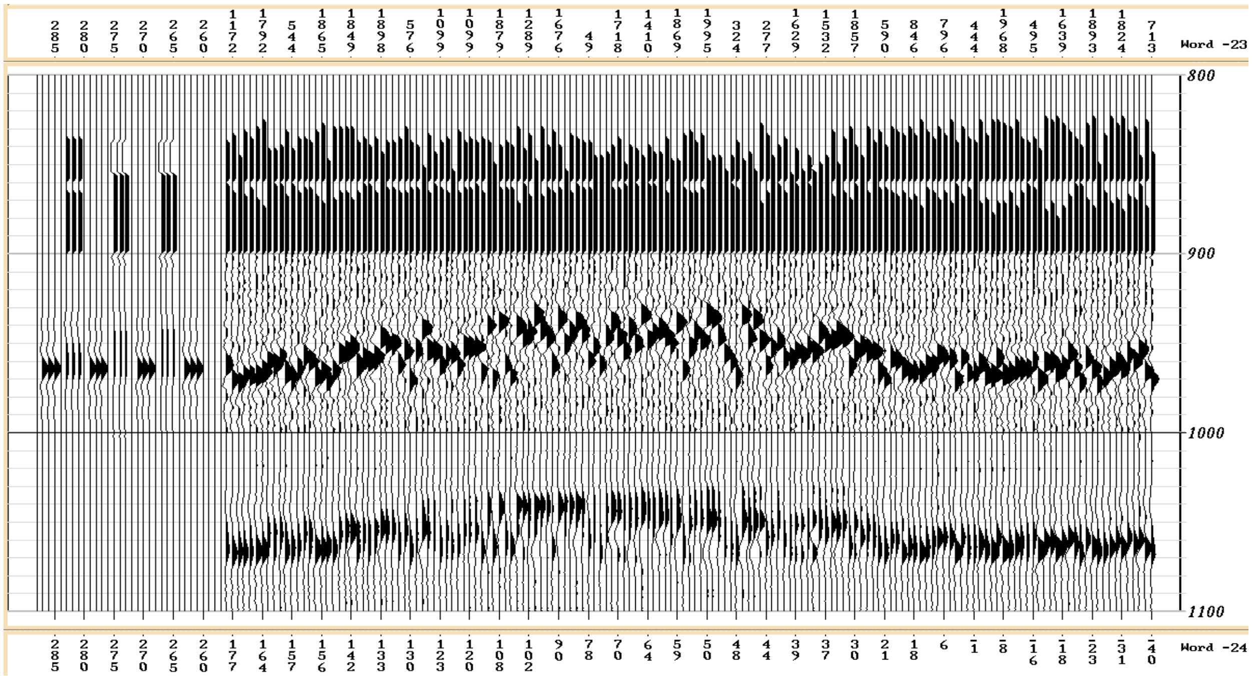

Figure 1. $N_{sx}=N_{sy}=N_{rx}=N_{ry}=1$, number of polynomial coef.=16. Calculated "true" traces are between 900-1000ms. Interpolated traces are between 1000-1100ms. Heights of 1st & 2nd bar (800-900ms) are cross-correlation & absolute relative misfit between the "true" and interpolated. On the left (top & bottom annotation 260-285 & beyond), the 4 groups of 3 identical traces are "true" zero-offset traces at bin-centre. Between these from right to left are 3 identical traces each for simple stack (265-267), Gaussian stack (273-275) and interpolated (281-283).

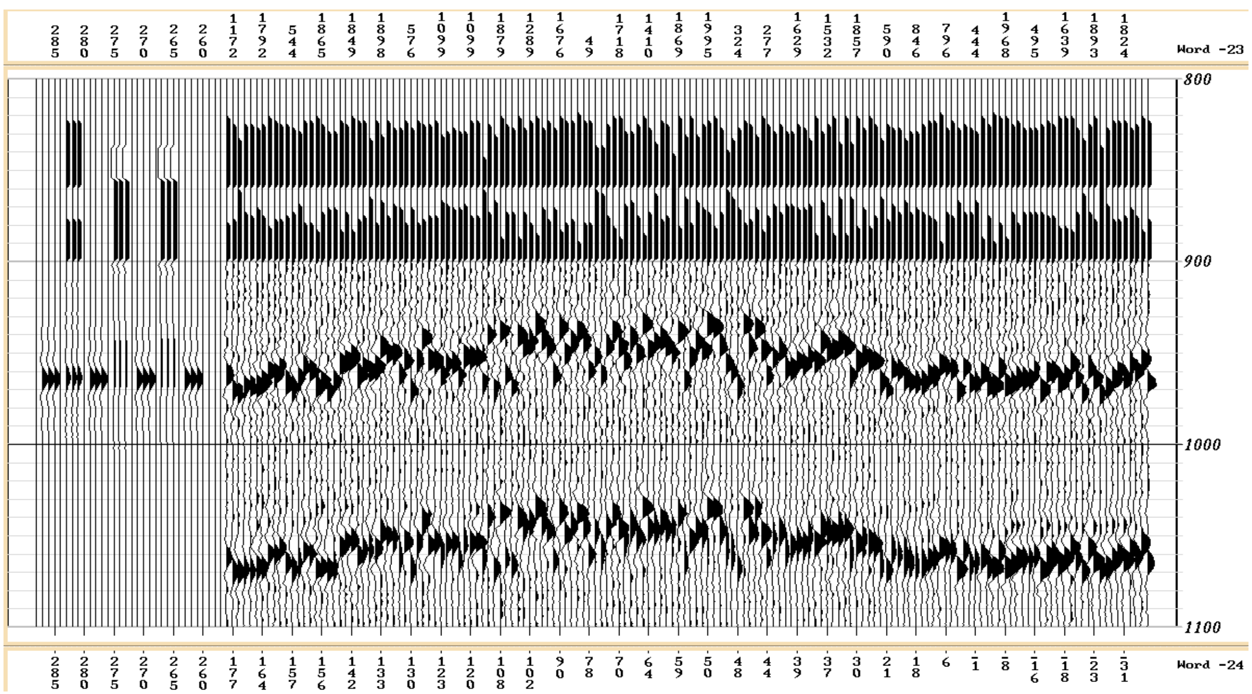

Figure 2. $N_{sx}= N_{sy}=N_{rx}=N_{ry}=3$, number of polynomial coef.=256 . Compare with Fig.1, higher order polynomial generates better gather and bin-center zero-offset traces: higher cross-correlation and lower misfit. The gather traces (only 156 of a total of 256 displayed) are arranged in increasing source/receiver azimuth (bottom annotation) and random source/receiver offset (top annotation), and are NMO corrected with true velocity.