3D Prestack Kirchoff Beam Migration for Depth Imaging: Theory and Data Examples

Y. Sun, F. Qin, S. Checkles, J. Leveille
Amerada Hess Corporation, 500 Dallas Street, Houston, Texas 77002

Summary

We present a beam implementation for 3D prestack Kirchhoff depth migration of seismic data. Unlike conventional Kirchhoff migration in which the input seismic traces in time are migrated one at a time into the 3D image volume for the Earth's subsurface, the beam migration processes a group of input traces at a time. The migration of each group of traces, or supergather, consists of two major steps: slant stacking the traces into a $\tau - \mathbf{P}$ domain beam volume, and mapping the beams into the image volume. Since the beam volume is much smaller than the image volume, the beam migration cost is roughly proportional to the number input supergathers instead of the number input traces. The computation speedup of beam migration over conventional Kirchhoff migration is roughly proportional to N_g, the average number of traces per supergather, resulting a theoretical speedup up to two orders of magnitudes ($10 \sim 100$). The beam migration was successfully implemented and has been in production use for several years in Amerada Hess Corporation. A factor of 5-25 speedup has been achieved in our in-house depth migrations. The implementation made 3D prestack full volume depth imaging feasible in parallel distributed environment.

Introduction

Kirchhoff migration using raytracing or traveltimes is currently the tool of choice for 3D prestack depth imaging. Kirchhoff migration is extremely suitable for parallel implementation, partitioning both input and output volumes, or targeting subvolumes. Due to the shear number of summations, Kirchhoff migration is expensive. The flexibility of target-oriented processing helps when quick processing turn-around time is essential. However, in practice, interpreters almost always prefer larger volumes during a prospect definition, compromising between the length of runtime, size of the output, or sometimes, the quality of image.

We use the following well-known and rather general form (see e.g., Bleistein [1], H. Sun and Schuster [2]) for the Kirchhoff migration image $M(\mathbf{x})$ at subsurface point $\mathbf{x}$,

$$M(\mathbf{x}) = \sum_{\mathbf{x}_s} \sum_{\mathbf{x}_r} \int d\omega A(\mathbf{x}_s, \mathbf{x}_r, \mathbf{x}, \omega) \tilde{P}_{sr}(\omega) \exp[i\omega(\tau_s(\mathbf{x}_s, \mathbf{x}) + \tau_r(\mathbf{x}, \mathbf{x}_r)], \tag{1}$$

where $\tilde{P}_{sr}(\omega)$ is the input seismogram in the circular frequency domain ω, for the source at $\mathbf{x}_s$ and receiver at $\mathbf{x}_r$, $\tau_s(\mathbf{x}_s, \mathbf{x})$ is the traveltime from the source to the subsurface point $\mathbf{x}$, $\tau_r(\mathbf{x}, \mathbf{x}_r)$ is the traveltime from the subsurface point to the receiver, and, $A(\mathbf{x}_s, \mathbf{x}_r, \mathbf{x}, \omega)$ contains geometric spreading information and factors necessary to generate correct migration output amplitudes. The subscripts in τ_s and τ_r is used to distinguish these travel time functions from the total travel time function $\tau = \tau_s + \tau_r$ and from the τ variable in the $\tau - \mathbf{p}$ domain. Issues such as amplitudes (with damping), distance aperture, angular muting, antialiasing, are important for a successful migration implementation. Post migration processing such as AVO and velocity analysis usually requires that the Kirchhoff summations be separated into intervals of offsets between sources and receivers. We will pay special attention to the computation count of the migrations. To make our discussions more specific, let N_x, N_y, and N_z be the 3D image volume grid sizes. The output image volume size per offset is $N_{out} = N_x \times N_y \times Nz$. The number of offset ranges is N_{off}. The total output volume over all offsets is $N_{off} \times N_{out}$. The number of input traces N_{in} is determined by the number of shots (proportional to $N_x \times N_y$) and number of receivers per shot. Let the number of samples per trace be N_t. We will use the following values for our "typical" 3-D marine survey: $N_x = N_y = N_z = 1000$, $N_t = 2000$, $N_{in} = 50,000,000$, and $N_{off} = 50$.

The computation cost of the Kirchhoff migration (Eq. 1) is proportional to the number of input traces and the number of output image points,

$$N_{kirchhoff} = C_k \cdot N_{in} N_x N_y N_z. \tag{2}$$

where C_k is the number of (arithmetic or floating point) operations per input trace per image point. As can be seen, the cost of Kirchhoff migration increases rapidly with the survey size and become prohibitive for large surveys. Our beam migration concepts and implementation in this paper overcomes this difficulty.

Beam Migration Methodology

Introduction

In a postal distribution center, the postal service bundles the mail according to the neighborhoods of the recipients before delivery. This is more efficient than making one trip from the distribution center to the same neighborhood for each piece of mail. In a Kirchhoff implementation, if the contribution to the output image by one input trace is

computed independent of those of other input traces, that implementation is the migration equivalent of "one-trip-per-piece-of-mail". In such an implementation, the computation time of migrating two traces is roughly double that of migrating one trace.

The input traces of a 3D reflection seismic survey are densely spaced in both the source and receiver coordinates, implying large dimensional spatial adjacency. The postal analogy suggests that we should take advantage of the spatial adjacency of traces, and migrate a bundle of many traces for the cost of migrating one or two traces. Specifically, we would like to regroup the input traces into a collection of supergathers, with each supergather containing traces that are adjacent to each other. The patches of a supergather are two *small* areal neighborhoods containing the sources and receivers of the traces in the gather. Let $\mathbf{x}_{sg}$ and $\mathbf{x}_{rg}$ be some reference points within source and receiver patches respectively. We will refer to $\mathbf{x}_{sg}$ and $\mathbf{x}_{rg}$ as the centers of the patches, and require that the patches be limited in size by some radius R. Equation 1 can be rewritten as summations over the supergathers,

$$M(\mathbf{x}) = \sum_{g=1}^{N_{gthrs}} M_g(\mathbf{x}), \tag{3}$$

where the summation is over the supergather index, N_{gthrs} is the number of supergathers, and M_g is the migration contribution from the g^{th} supergather,

$$M_g(\mathbf{x}) = \sum_{\substack{n=1 \\ |\Delta\mathbf{X}_s|<R, |\Delta\mathbf{X}_r|<R}}^{N_g} \int d\omega A(\mathbf{x}_s, \mathbf{x}_r, \mathbf{x}, \omega) \tilde{P}_{sr}(\omega) \cdot \exp[i\omega(\tau_s(\mathbf{x}_s, \mathbf{x}) + \tau_r(\mathbf{x}, \mathbf{x}_r)], \tag{4}$$

where N_g is the number of traces in the g^{th} gather, and the source and receiver coordinates $\mathbf{x}_s$ and $\mathbf{x}_r$ vary with the trace summation index n. In Eq. 4, $\Delta\mathbf{x}_s = \mathbf{x}_s - \mathbf{x}_{sg}$ and $\Delta\mathbf{x}_r = \mathbf{x}_r - \mathbf{x}_{rg}$ are the coordinate deviations of the traces relative to the reference points. The number of traces vary from supergather to supergather, for our analysis, we will use $N_g = 50$, $N_{gthrs} \sim N_{in}/N_g = 1000000$.

First Order Approximation

For a given supergather, since the source and receiver patch sizes are small, we can approximate the travel times within a patch using the time map information associated with a patch center, by interpolation or Taylor expansions. Taylor expansions of the time map functions $\tau_s(\mathbf{x}_s, \mathbf{x})$ and $\tau_r(\mathbf{x}, \mathbf{x}_r)$ about the source and receiver centers lead to (in lowest order):

$$M_g(\mathbf{x}) = A_c(\mathbf{x}_{sg}, \mathbf{x}_{rg}, \mathbf{x}) B_g^{slant2}(\tau, \mathbf{p}_s, \mathbf{p}_r), \tag{5}$$

$$B_g^{slant2}(\tau, \mathbf{p}_s, \mathbf{p}_r) = \sum_{\substack{n=1 \\ |\Delta\mathbf{X}_s|<R, |\Delta\mathbf{X}_r|<R}}^{N_g} P_{sr}^c(\tau + \mathbf{p}_s \cdot \Delta\mathbf{x}_s + \mathbf{p}_r \cdot \Delta\mathbf{x}_r), \tag{6}$$

$$\tau = \tau_s(\mathbf{x}_{sg}, \mathbf{x}) + \tau_r(\mathbf{x}, \mathbf{x}_{rg}), \mathbf{p}_s = \nabla_{\mathbf{x}_s} \tau_s(\mathbf{x}_s, \mathbf{x})|_{\mathbf{x}_s = \mathbf{x}_{sg}}, \text{ and } \mathbf{p}_r = \nabla_{\mathbf{x}_r} \tau_r(\mathbf{x}_r, \mathbf{x})|_{\mathbf{x}_r = \mathbf{x}_{rg}}. \tag{7}$$

For simplicity, we have replaced $A(\mathbf{x}_{sg}, \mathbf{x}_{rg}, \mathbf{x}, \omega)$ with its frequency independent component $A_c(\mathbf{x}_{sg}, \mathbf{x}_{rg}, \mathbf{x})$, and absorbed the frequency dependence of A in the seismogram by replacing $P_{sr}(\omega)$ with $\tilde{P}_{sr}^c(\omega)$, or with $P_{sr}^c(t)$ in the time domain. It is important to point out that the expression for $B_g^{slant2}(\tau, \mathbf{p}_s, \mathbf{p}_r)$ is a slant stack of the seismogram with two slant parameters $\mathbf{p}_s$ and $\mathbf{p}_r$. If the source and receiver coordinates are not independent, the two parameter slant stack can be reduced to a single parameter slant stack:

$$\tau + \mathbf{p}_s \cdot \Delta\mathbf{x}_s + \mathbf{p}_r \cdot \Delta\mathbf{x}_r \Rightarrow \tau + \mathbf{P} \cdot \mathbf{X} \tag{8}$$

where, for example, under the sort restrictions, $\Delta\mathbf{x}_s = 0$ (CSG), $\Delta\mathbf{x}_r = 0$ (CRG), $\Delta\mathbf{x}_s - \Delta\mathbf{x}_r = 0$ (COG), or $\Delta\mathbf{x}_s + \Delta\mathbf{x}_r = 0$ (CMG) with $\mathbf{P}$ and $\mathbf{X}$ defined by

$$\mathbf{P} = \mathbf{p}_r, \qquad \mathbf{X} = \Delta\mathbf{x}_r \qquad \text{(CSG)} \tag{9}$$

$$\mathbf{P} = \mathbf{p}_s, \qquad \mathbf{X} = \Delta\mathbf{x}_s \qquad \text{(CRG)} \tag{10}$$

$$\mathbf{P} = \mathbf{p}_s + \mathbf{p}_r, \quad \mathbf{X} = (\Delta\mathbf{x}_r + \Delta\mathbf{x}_s)/2 \quad \text{(COG)} \tag{11}$$

$$\mathbf{P} = -\mathbf{p}_s + \mathbf{p}_r, \quad \mathbf{X} = (\Delta\mathbf{x}_r - \Delta\mathbf{x}_s)/2 \quad \text{(CMG)}. \tag{12}$$

Here we have assumed that patch centers are chosen to be the source and receiver locations of one of the traces in the gather. The replacing the two-parameter slant stack expression in Eq. 6 with the one-parameter slant stack expression (Eq. 8), we have:

$$B_g^{slant1}(\tau, \mathbf{P}) = \sum_{\substack{n=1 \\ |\Delta\mathbf{X}_s|<R, |\Delta\mathbf{X}_r|<R}}^{N_g} P_{sr}^c(\tau + \mathbf{P} \cdot \mathbf{X}). \tag{13}$$

where C[SROM]G stands for CSG, CRG, COG, or CMG sort restrictions. Like $\mathbf{x}_s$ and $\mathbf{x}_r$, the variable $\mathbf{X}$ varies with the trace summation index n. $B_g^{slant1}(\tau, \mathbf{P})$ is to be used in place of $B_g^{slant2}(\tau, \mathbf{P})$ in Eq. 5 to get,

$$M_g(\mathbf{x}) = A_c(\mathbf{x}_{sg}, \mathbf{x}_{rg}, \mathbf{x}) B_g^{slant1}(\tau, \mathbf{P}). \tag{14}$$

Equations 3, 7, 9–12, and 13–14 form the basis for the one-parameter slant stack migration of (CSG, CRG, COG, CMG) supergathers using the first order time map approximation.

Computation Counts

The $\tau - \mathbf{P}$ volume of the slant stack (Eq. 13) can be discretized by $N_{px} \times N_{py} \times N_\tau$ grid points, where $N_{px} \times N_{py}$ is the size of the sampling grid for $\mathbf{P}$, and N_τ is the number of samples in slant stack time. We use the following typical values: $N_{px} = N_{py} = 50$, $N_\tau = N_t = 2000$.

The beam migration using the slant stack consists of the two major steps: the beam forming (Eq.13), and mapping from $\tau - \mathbf{P}$ space to the output volume $\mathbf{x}$ (using Eqns. 7, 9–12, and 14). The cost of performing a beam migration therefore has two dominant terms,

$$N_{beam} = N_{gthrs}(C_b N_g N_{px} N_{py} N_\tau + C_m N_x N_y N_z), \tag{15}$$

where C_b and C_m are machine dependent constants. This is to be compared with that of Kirchhoff migration (Eq. 2): $N_{kirchhoff} = C_k N_{gthrs} N_g N_x N_y N_z$, with $N_{in} = N_{gthrs} N_g$. For sufficiently large output volume, the beam forming cost (first term in Eq. 15) can be considered a small, leading to, $N_{beam} \approx C_m N_{gthrs} N_x N_y N_z$. Thus the speedup of the beam migration over the Kirchhoff migration can be estimated by

$$Speedup \equiv N_{kirchhoff}/N_{beam} \approx (C_k/C_m) N_g. \tag{16}$$

and is roughly proportional N_g, the number of traces per supergather, with C_k/C_m being of order unity. With typical seismic surveys, N_g ranges between 20 to 200. This theoretical speedup is hard to achieve.

Generalized Slant Stack

Sort restrictions (such as CSG, CRG, CMG, or COG) on the supergathers are required to reduce the two-parameter ($\mathbf{p}_s$ and $\mathbf{p}_r$) slant stack (Eq. 6) to a one-parameter ($\mathbf{P}$) slant stack (Eq. 13). Sort restrictions degrade the speedup (due to smaller N_g). One way to overcome the restriction is to simply impose no sort in supergathering. A gather can be corrected to a specific sort type by relocating the traces slightly within the source and receiver patches in a model or moveout consistent manner before forming the beams. This correction process introduce errors. For shallower depth, keeping only the leading terms also leads to large errors.

Notice that in the computation count for the one-parameter beam migration, the exact attribute of $\mathbf{P}$ (such as $\mathbf{P} = \mathbf{p}_s + \mathbf{p}_r$ (COG)) does not affect the outcome (Eq. 15) of the qualitative analysis. We can choose the $\tau - \mathbf{P}$ representation such that the mapping $\mathbf{x}(\mathbf{P}, \tau)$ and the inverse mappings $\mathbf{P}(\mathbf{x})$ and $\tau(\mathbf{x})$ are simple and unique. For example, if we choose $\mathbf{P}$ to be $\mathbf{p}_s$, then $\mathbf{x}(\mathbf{P}, \tau)$ for a constant $\mathbf{P}$ correspond to ray paths emanating from the source patch center, and $\tau(\mathbf{x})$ and $\mathbf{P}(\mathbf{x})$ are supplied by the maps of the travel times and their gradients with respect to the source. The $\mathbf{P} = \mathbf{p}_s + \mathbf{p}_r$ is probably desirable because of its symmetry with respect to the source and receiver. Directional path normals to the total travel time contours are also good candidates. Normals to constant velocity migration ellipses or straight lines from the source-receiver midpoints are found to be robust choices. Given that we have picked the beam directional parameterization $\mathbf{x}(\mathbf{P}, \tau)$, we can now do beam forming, or migrate into the $\tau - \mathbf{P}$ volume using Eq. 4 as follows,

$$M_g(\mathbf{x}) = B_g(\tau(\mathbf{x}), \mathbf{P}(\mathbf{x})), \tag{17}$$

$$B_g(\tau, \mathbf{P}) = \sum_{\substack{n=1 \\ |\Delta \mathbf{X}_s| < R, |\Delta \mathbf{X}_r| < R}}^{N_g} A_c(\mathbf{x}_s, \mathbf{x}_r, \mathbf{x}(\mathbf{P}, \tau)) P_{sr}^c(\tau_s(\mathbf{x}_s, \mathbf{x}(\mathbf{P}, \tau)) + \tau_r(\mathbf{x}(\mathbf{P}, \tau), \mathbf{x}_r)), \tag{18}$$

where all occurences of $\mathbf{x}$ is replaced with $\mathbf{x}(\mathbf{P}, \tau)$ including those implicit in $\Delta \mathbf{x}_s$ and $\Delta \mathbf{x}_r$. The expression for computation counts, Eq. 15, remains valid for the beam forming and migration using Eqns. 17 and 18, as exactly the same arguments that lead to Eq. 15 still apply. The proportionality constant C_b for the beam-forming varies somewhat depending the choice of the beam parameters $\mathbf{P}$ and τ. Also notice that the sort restrictions on the supergathers are no longer required if Eq. 18 above or its Taylor expansion versions are used for the beam forming.

Implementation Notes and Strategy

Our implementation is based on a master-slave model with a distributed homogeneous set of worker nodes. We also tried to minimize the communication between worker nodes during the migration by precomputing time maps for a regular grid of source points on the model surface. Each node stores the entire set of time maps (compressed) on its local disk, so that no nodes will need time maps from other nodes during migration. Due to disk space and communication limitations, slant stacking of the supergathers are not precomputed. Each node is responsible for migrating a fixed offset range. Each worker sorts the input traces it obtained from the master into supergathers based on the source and receiver coordinates of the traces. These sorted gathers are stored on local disks of the worker nodes, and removed from the disks as they are migrated. The beam migration implementation was initiated in 1994. We first tested our beam migration implementation in production around the end of 1995 on North Sea data set. The data set was very small (25 million traces, covering $11.25 \times 10.0 km^2$ or 5.4 Gulf of Mexico blocks), so

we were able to output a full volume skipping every other line and every other crossline line, using a 4-node IBM SP2 (each node was an IBM 590, with a 66Mhz clock, 512Mb RAM). Tight parameters were used for additional runs targeting the shallow depths. We were encouraged to find that the beam migration ran roughly 25 times faster than the Kirchhoff migration. The migration of the first 9 out of 31 offset ranges of input took two weeks. The beam migration has been in production processing since it was first tested successfully.

Practical Considerations and Data Examples

Two examples are presented illustrating the results of implementing beam migration. Both are marine streamer acquisition with typical modern parameters of approximately: 6000 meter cable length, 40 meter subsurface line spacing, 12.5 meter subsurface crossline spacing and 60 fold. The first example illustrates a subsalt migration from the Gulf of Mexico. Figure 1(left) shows the final image obtained by full volume Beam migration with the top and base of salt identified. The processing flow in summary was: (1) Target oriented Kirchhoff migration to build the velocity model to the top of salt (2) Poststack depth migration to image and pick top of salt (3) Full volume Beam prestack depth migration to image and pick base of salt (4) Full volume Beam prestack depth migration with final velocity model. A well has been drilled confirming both the thickness of the salt in the model and the presence of hydrocarbons associated with the bright subsalt reflection. Figure 1(right) illustrates the second example, the results of Beam migration in an area with complex faulting. Here the velocity model was derived from target oriented Kirchhoff migration followed by a full volume Beam migration. Note the clear and crisp imaging of the faults.

By assuming that the run time of a standard Kirchhoff migration increases linearly with the size of the output volume a relative speedup of 7 to 10 times has been consistently observed between standard Kirchhoff and efficient the Beam implementations.

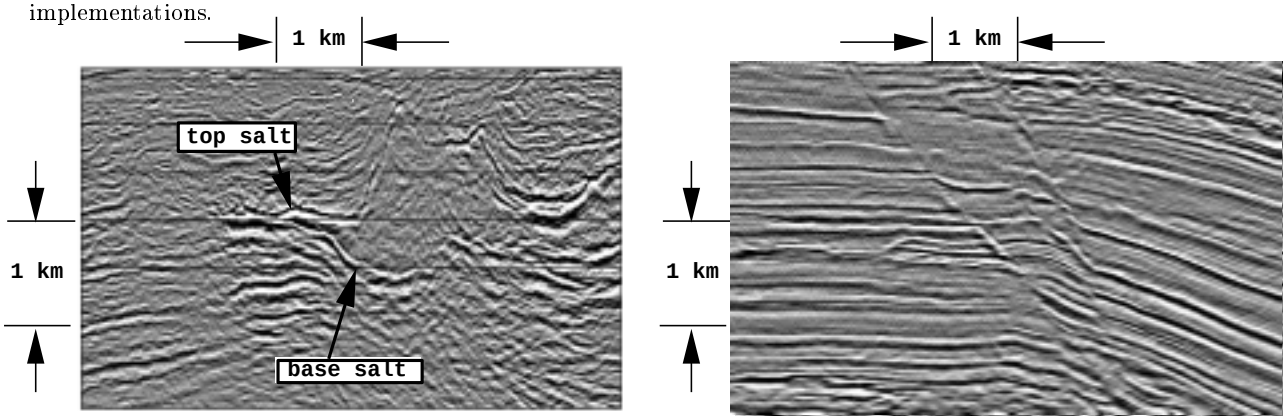

Fig. 1: Left: 3D prestack beam migration with salt. Right: 3D prestack beam migration with faults.

Discussion and Conclusions

We have presented a beam approach to Kirchhoff migration. In this beam approach, the input traces are migrated one supergather at a time. The source and receiver positions of traces in each supergather are scattered in small patches so that the migration of the gather can be performed by stacking the contribution of the gather into a beam volume before projecting beams into the much larger output imaging volume. This approach has the theoretical speedup of up to two orders of manitude over Kirchhoff migration. A speedup closer to 10 has been consistently achieved in production processing, subject to practical survey geometry limitations and computational resource compromises.

Acknowledgement

This work would not have been possible without the support of many people in the Geoscience Technology and MIS Geoscience groups. We particularly thank Vic Forsyth, Jeff Davis, Mike Oyler, Gary Donathan and Gary Whittle. Gerald Neale, formerly of Amerada Hess deserves special thanks for his guidance and sharing his many insights at the early stages of this project. Bee Bednar was project supervisor when the work was accomplished and we thank him for his support and patience. Finally John Weigant contributed much to our understanding through his massive amount of processing and code breaking. We thank Scott Morton, J. C. Wan, John Potter, and Simon Edwards for discussions. We also thank Amerada Hess Corporation for permission to publish this paper.

References

[1] N. Bleistein. *Mathematical Methods for WavePhenomena.* **Academic Press**, NY, NY, 1984.

[2] H. Sun and G. T. Schuster. *Wavepath Migration versus Kirchhoff migration: theory and poststack examples*, **Utah Tomography and Modelling/Migration Development Project**, 1999 Annual Report.