ENHANCING MASSIVE LAND 3D SEISMIC DATA USING NONLINEAR BEAMFORMING: PERFORMANCE, QUALITY AND PRACTICAL TRADE-OFFS

I. Silvestrov[1], A. Bakulin[1], D. Nekluydov[2], K. Gadylshin[2], M. Protasov[2]

[1] Saudi Aramco; [2] Institute of Petroleum Geology and Geophysics

Summary

Modern land seismic data are typically acquired using high spatial trace density but single sensors or small source and receiver arrays. These datasets are challenging to process due to their massive size and rather low signal-to-noise ratio caused by scattered near surface noise. Prestack data enhancement becomes a critical step in processing flow. Nonlinear beamforming was proven very powerful for 3D land data. It requires computationally costly estimations of local coherency on dense spatial/temporal grids in 3D prestack data cubes and poses inevitable trade-off between performance of the algorithm and quality of the obtained results. In this work, we study different optimization schemes and discuss practical details required for applications of the algorithm to modern 3D land datasets with hundreds of terabytes of data.

Introduction

The land seismic industry is moving toward high-density surveys using smaller source/receiver arrays or single sensors. Large field arrays with 5-10 m intra-arrays spacing used in the past were designed to attenuate strong noise caused by ground-roll and multiple scattering in the near-surface. Theoretically, denser data sampling and decreased size of the arrays should allow better sampling of the noise wavefield facilitating its attenuation during processing stage. In practice, high-density surveys with uniform and dense 5-10 m sampling in all directions remain prohibitively expensive with current sensor technology. Current practice is to acquire orthogonal 3D surveys with smaller inline and much larger crossline spacing. With small arrays or single sensors, such acquisition results in massive datasets with rather low signal-to-noise (SNR) ratio that are challenging to process. Conventional processing techniques such as surface-consistent scaling and deconvolution, statics estimation and velocity analysis, all require good prestack SNR and deliver suboptimal results otherwise. This particularly affects quality of prestack inversion that requires reliable and accurate prestack amplitudes in the gathers. Therefore, data enhancement step becomes critical component in processing of modern high-channel count and single-sensor seismic surveys.

Coherency-based data enhancement techniques are widely used in the post-stack domain, and can be applied to improve the signal level in prestack data in pre-migrated and post-migrated domains. The main techniques from this category are based on local slant stacking of the data and selection of the most coherent components. More advanced approaches, coming mainly from common-reflection surface (CRS) and multifocusing theory, utilize second-order approximations of the wavefront to better describe kinematic of the events and to stack locally along them to increase SNR. The moveout can be described using either by a global operator or a local one. The latter type of the methods appear most flexible for enhancing challenging 3D land data where static issues often invalidate any global moveout behavior. At the same time, these methods rely on intensive numerical search for optimal coherency over the entire 5D data domain making them computationally expensive. By representing a traveltime surface locally as a general second-order curve (Buzlukov and Landa, 2013), a nonlinear beamforming algorithm was introduced for enhancing massive challenging land datasets (Bakulin et al., 2019). In this work, we discuss implementation details of this algorithm and provide insights into its quality and performance, and the trade-offs for 3D land processing operating on massive datasets with petabytes of data.

Method

The main idea of the nonlinear beamforming method is to describe traveltime moveout locally as a second-order surface, estimate its parameters and perform local summation along this local moveout to improve the signal-to-noise ratio. Considering the data space with a coordinate vector $\vec{x} = (x_s, y_s, x_r, y_r)$ defined by source and receiver *x* and *y* coordinates, the traveltime, *t*, can be locally represented using a Taylor series expansion as:

$$t(\vec{x}_0 + \Delta\vec{x}) = t(\vec{x}_0) + \vec{A}^T \Delta\vec{x} + \Delta\vec{x}^T B \Delta\vec{x}, \quad (1)$$

where $\vec{A}$ is a first-derivative gradient vector and B is a matrix of second derivatives. In total, 14 unknown coefficients of vector A and matrix B define the local traveltime surface at a current sample. Estimation of all these kinematic parameters is too costly from the computational point of view and simplifications of formula (1) are usually required. In the current work, we fix two arbitrary directions in the data space and consider sections of the traveltime surface along two other directions only. This reduces the number of the unknown parameters to five. We estimate these local kinematic parameters by scanning many different trajectories and finding one, with the best coherency defined by the maximum value of a semblance function. Optimization can be implemented either as a simultaneous global 5D search for all five parameters, or sequential estimation of one parameter after another as in a coordinate-descent method. The sequential strategy is more computationally attractive; however, it may get trapped into a local maxima and produce erroneous results when the signal-to-noise ratio in the data is low. Another hybrid solution is to utilize a sequential “2+2+1” strategy (Hoecht et al., 2009), searching first for a pair of parameters (first and second traveltime derivatives) in one plane, then in another plane, and then for mixed derivative coupling both directions. Figure 1 shows real-data example illustrating comparison of the global 5D search and the sequential "2+2+1" strategy. The global 5D search was done using two approaches. In the first one, a brute force method was applied where we scan all possible values of five kinematic parameters to maximize semblance function. In the second one, we use adaptive simulated annealing method to make global optimization more efficient. The "2+2+1" strategy was implemented using the brute force method only. As one can see, the "2+2+1" strategy provides a slightly noisier semblance panel showing that this method is less robust with respect to noise. At the same time, overall semblance behaviour is similar to 5D case, and the enhanced data themselves are comparable.

To evaluate computational performance of the approaches, we compare number of semblance calculations required for all the three methods. Search intervals encompass range of possible dips and curvatures of seismic events (measured as maximum moveout in ms over a defined aperture length). Figure 2 depicts relative computational effort involved in a single parameter estimation step at one time sample for a synthetic numerical test. As expected, number of calculations for 5D brute-force approach significantly exceeds that of two other methods. In addition, it rapidly increases with increasing size of the searching intervals. The number of calculations required for 5D simulated annealing is also big, but it remains more or less constant when intervals increase. The "2+2+1" strategy shows best performance with number of required semblance calculations being one to two orders of magnitude less than for 5D simulated annealing approach. Taking into account massive size of the prestack seismic datasets and comparable quality of the results, we conclude that "2+2+1" strategy is a reasonable compromise between quality and performance, at least for the data with a moderate noise level as considered in this example.

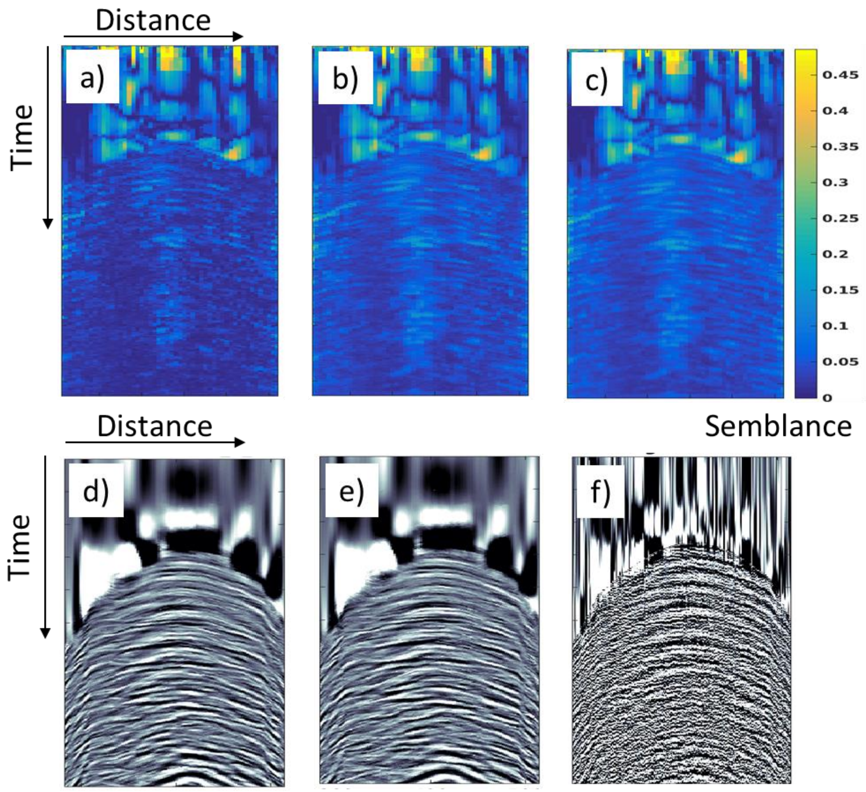

Figure 1 *Quality comparison between different semblance optimization strategies. Estimated semblance values for brute-force hybrid "2+2+1" approach (a) are slightly noisier than for simulated annealing 5D optimization (b) and brute-force hybrid 5D approach (c). Data enhancement results are comparable between brute-force hybrid "2+2+1" (d), simulated annealing 5D method (e), and brute-force global 5D approach (not shown here). Original data is depicted in (f) for reference.*

Optimizing spatial and temporal estimation grid

Once we settled optimization strategy for each spatial/temporal location, next step is to optimize the global grid used for parameter estimation in the entire dataset. Using every trace and every time sample is not computationally feasible for massive datasets. More optimal method is a so-called operator-oriented approach (Hoecht et al., 2009), where "operator" defines a traveltime surface along which the local moveout correction is done. In this approach, kinematic parameters are estimated and stored at samples of sparse so-called parameter traces that are located at decimated uniform grid in data space. After estimation step, these parameters define estimated travel-time trajectories spread around the neighbourhood and used to perform local summation for each of the actual data traces. The coarser the spatial grid of parameter traces, the faster algorithm performs. At the same time, too coarse grid may reduce accuracy of estimated moveouts and reduce quality of enhanced data traces located far from the operator traces. It is convenient to introduce a ratio (K_{est}) between the step size of the estimation grid and a summation aperture that defines size of the area used for stacking (beamforming). The value of K_{est} between 0.3 and 1 was found experimentally to provide reasonable results. Smaller value of K_{est} usually leads to better quality of the summation results, but the computational effort increases. For a fixed ratio of K_{est} (say 0.7 in the current case), there are ways to further improve computational efficiency by estimating parameters on a coarser regular or random grid, and then interpolating parameters to the original grid defined by K_{est}. While advanced interpolation and inpainting techniques are subject of other studies (Gadylshin et al., 2019), here we focus on simplest linear interpolation solution. We define the ratio between the grid steps of this new coarser grid and the originally chosen estimation grid as K_{int}^{x}, assuming that the same steps are used in both spatial dimensions. Figure 3 shows a comparison between the originally enhanced data ($K_{int}^{x} = 1$) and the data, where the estimation of parameters was done every second spatial point following by parameter interpolation ($K_{int}^{x} = 2$). The achieved speedup in the second case is four times (due to 2D estimation grid), while the enhancement results are comparable (Figure 4a). It is interesting to note that performing summation at this coarser grid without additional parameter interpolation (equivalent to the usage of $K_{est} = 1.4$), significantly deteriorates the results (Figure 4b).

To increase performance of the algorithm, similar approach can be applied along the temporal axis as well. Instead of estimating parameters at every time sample, we prescribe a coarser time grid with a ratio of K_{est}^{t} with respect to the sampling. Since the semblance during a coherency search is calculated in a certain time window, the half-window size appears as good candidate for a grid step in time direction as confirmed by numerical computations. NRMS values (Figure 4c) obtained with such a step ($K_{est}^{t} = 11$ in this case) are comparable to the previous spatial interpolation results with $K_{int}^{x} = 2$. In the current example, additional performance speedup is a factor of four. Increasing K_{est}^{t} beyond 11 (half-window size) produces data with larger NRMS values indicating unacceptable deviation from a reference dataset without use of interpolation (Figure 4d).

Conclusions

Efficient implementation of nonlinear beamforming based on estimating local coherency on dense spatial/temporal grids in 3D prestack data cubes poses inevitable trade-off between performance of the algorithm and quality of the obtained results. We demonstrate that hybrid "2+2+1" optimization scheme provides reasonable compromise between speed and quality and lead to reliable results on both synthetic and real data. It remains computationally demanding for massive 3D land datasets. To achieve additional speedup, we implement estimation of kinematical parameters on coarser grids in space and in time followed by interpolation back to original grid. With not too aggressive grid decimation, straightforward linear interpolation delivers very similar enhanced data as measured by sensitive NRMS repeatability metric. For aggressive decimation, interpolation can still be applicable but more advanced parameter interpolation schemes are required.

References

Hoecht, G., Ricarte, P., Bergler, S. and Landa, E. [2009] Operator-oriented CRS interpolation. *Geophysical Prospecting*, **57**, 957-981.

Buzlukov, V. and Landa, E. [2013] Imaging improvement by prestack signal enhancement. *Geophysical Prospecting*, **61**, 1150-1158.

Bakulin, A. Silvestrov, I., Dmitriev, M., Neklyudov, D., Protasov, N., Gadylshin, K., and Dolgov, V.. [2019] 3D Prestack land data enhancement based on nonlinear beamforming in cross-spread domain. 81st EAGE Conference and Exhibition, Extended Abstracts, Tu_R06_07

Gadylshin K., Silvetrov I., Bakulin A. [2019] Inpainting of local wavefront attributes using artificial intelligence, SEG Technical Program Expanded Abstracts, 2212-2216

Maximum searching intervals	5D Simulated Annealing	5D Brutal force	2+2+1 strategy
Dip = 20 ms, Curvature = 4 ms	5 006 (3D semblance)	50 000 (3D semblance)	160 (2D semblance) + 4 (3D semblance) ~ 20 (3D semblance)
Dip = 60 ms, Curvature = 4 ms	6 115 (3D semblance)	230 000 (3D semblance)	480 (2D semblance) + 4 (3D semblance) ~ 52 (3D semblance)
Dip = 100 ms, Curvature = 10 ms	3 861 (3D semblance)	10*10^6 (3D semblance)	2000 (2D semblance) + 10 (3D semblance) ~ 210 (3D semblance)
Dip = 200 ms, Curvature = 20 ms	6 489 (3D semblance)	320*10^6 (3D semblance)	8000 (2D semblance) + 20 (3D semblance) ~ 820 (3D semblance)

***Figure 2** Relative computational effort estimated for different optimization strategies. First column specifies values of maximum dip and curvature that define search intervals. Remaining columns shows number of required semblance calculations to obtain a solution. Observe significantly smaller number of calculations in "2+2+1" strategy.*

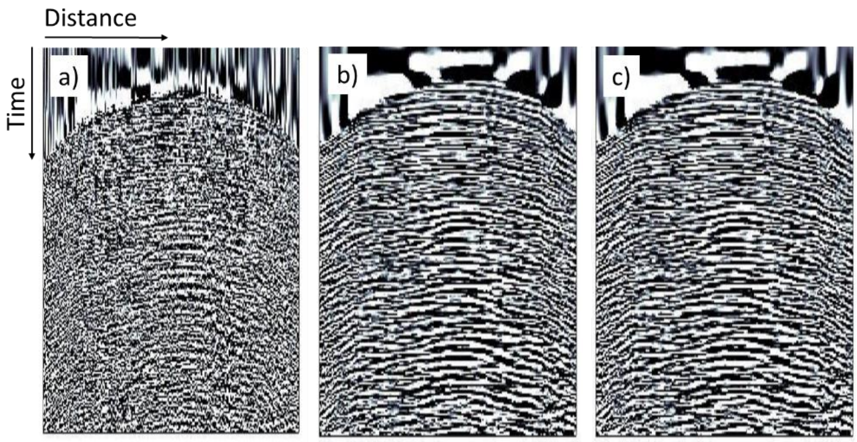

***Figure 3** Original data (a), enhanced data without parameter interpolation ($K_{int}^{x} = 1$) (b) and with interpolation (c) where the estimation of parameters was done at every second spatial point following by parameter interpolation ($K_{int}^{x} = 2$). Observe similar quality of (b) and (c).*

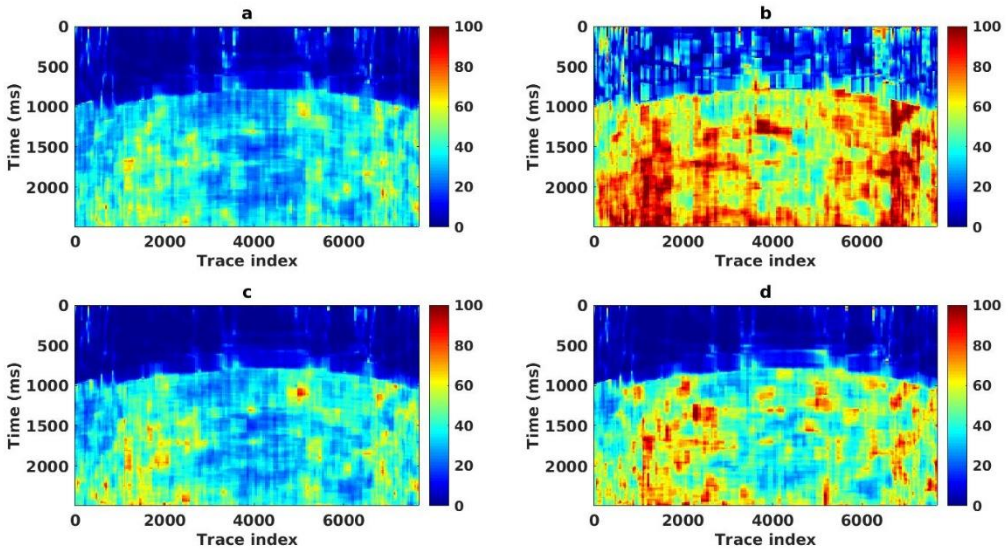

***Figure 4** NRMS values quantify repeatability or similarity between reference dataset enhanced using original dense grid and other versions of enhanced data obtained using different interpolation approaches: (a) $K_{est} = 0.7, K_{int}^{x} = 2, K_{int}^{t} = 1$; (b) $K_{est} = 1.5, K_{int}^{x} = 1, K_{int}^{t} = 1$; (c) $K_{est} = 0.7, K_{int}^{x} = 1, K_{est}^{t} = 11$; (d) $K_{est} = 0.7, K_{int}^{x} = 1, K_{est}^{t} = 32$. Average NMRS values are 27% for (a), 51% for (b), 30% for (c) and 36% for (d) with smaller values indicating closer resemblance to reference dataset. We consider NRMS $\leq$ 30% as an acceptable range here.*