Efficient four-dimensional supergrouping algorithm for enhancement of high-channel count seismic data

Maxim Dmitriev, Andrey Bakulin, Pavel Golikov, Geophysics Technology, EXPEC ARC, Saudi Aramco

Summary

Modern seismic acquisition is trending toward recording high-channel count data with smaller field arrays or single sensors. Reducing the size of field arrays leads to a significant degradation of data quality that needs to be dealt with during processing. We present a computationally efficient approach for four-dimensional supergrouping of shots and receivers that significantly improves the pre-stack signal-to-noise ratio. This algorithm allows enhancement of 1.5 TB of data in less than 10 hours using a single compute node, and so enabling efficient iterative application required for modern high-channel count seismic data.

Introduction

Any seismic processing technique that requires pre-stack information is a challenge to apply to data acquired in the desert environment of Saudi Arabia. A complex near surface with strong acoustic contrasts and heterogeneities creates an incredibly complex wavefield. In the past, large arrays of shots and receivers were employed in the field to suppress surface waves and backscattered noise so that reflections can be recorded with reasonable signal-to-noise ratio (SNR). Modern seismic acquisition is steadily moving to finer spatial sampling with smaller field arrays or point sources and receivers while typical distances between shot and/or receiver lines remain relatively large compared to inline sampling. As a result, we obtain hundreds of terabytes of data (Figure 1) with low SNR. Processing huge amounts of data with poor SNR becomes challenging both geophysically as well as computationally. The traditional land seismic data processing workflow cannot compensate for the loss of signal making it very difficult to obtain reliable estimation of pre-stack parameters such as velocities, deconvolution operators, statics, and surface-consistent scalars. Proper estimation of these parameters is a prerequisite for any further processing. Supergrouping was introduced by Neklyudov et al. (2015) to effectively address these issues and Bakulin et al. (2016) showed various field examples demonstrating uplift in processing. Initial implementation was not optimal to support many iterative runs of flexible supergrouping on huge datasets. In this study we describe a new computationally efficient algorithm of four-dimensional supergorouping that handles high-channel count and single-sensor data comprising hundreds of terabytes of data. It gives noise suppression power similar to large field arrays, but allows task-specific flexibility and enables additional corrections before stacking.

Method

We consider conventional 3D land broadband seismic acquisition currently used in Saudi Arabia and elsewhere. In a typical orthogonal geometry, the receivers for each shot are closely spaced along parallel receiver lines, whereas shots are closely spaced along parallel shot lines that are perpendicular to receivers (Vermeer, 2005).

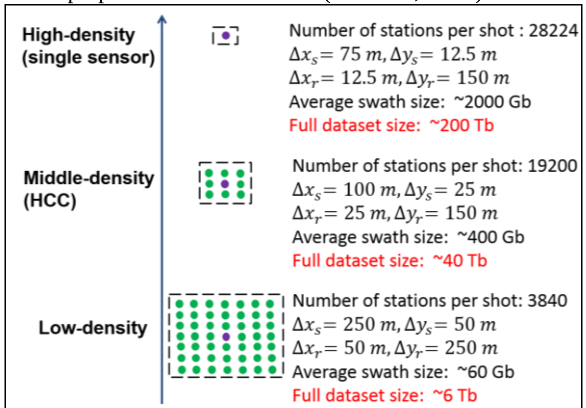

Figure 1: Examples of recent acquisition parameters illustrating the trend of increasing trace density. Currently, smaller field arrays of shots and receivers or single sensors are used to record broadband signal for imaging and inversion purposes.

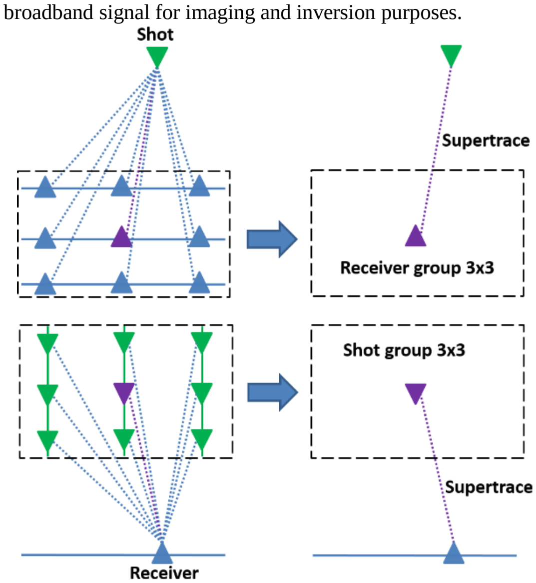

Figure 2: Scheme explaining supergrouping for 3D data.

Supergrouping can be implemented in the source domain, receiver domain or a combination. To construct a receiver group in the shot domain we fix a reference trace and find all traces $x_k(t), k = 1, \ldots, N$ located within a user defined stacking window (Figure 2). Stacking itself can be achieved with several stacking methods. Similarly to conventional group-forming, we can use straight summation:

$$y_k(t) = \frac{1}{N}\sum_{k=1}^{N} x_k(t).$$

Likewise, source supergrouping can be similarly performed in the common-receiver domain. While simple summation is robust and fast, it has three main limitations. First, it does not account for moveout differences. Second, it assumes that traces are well balanced inside the stacking window. Third, it suffers from intra-array statics. While these limitations are uncorrectable for field source and receiver arrays, they all can be compensated for during supergrouping. To address the first limitation, we apply supergrouping after normal moveout corrections using preliminary velocity information, which is always available. It allows us to preserve signals at higher frequencies and use larger summation apertures (Bakulin et al., 2016). The second limitation is addressed by using a diversity stack (Martinez et al., 1993) that weights each trace by its smoothed envelope before summation, whereas the final sum is then normalized by the sum of the original weights:

$$y_k(t) = \left(\frac{1}{N}\sum_{k=1}^{N}\frac{x_k(t)}{w_k(t)}\right)\left(\frac{1}{N}\sum_{k=1}^{N} w_k(t)\right).$$

Diversity stack helps to eliminate high-amplitude noise bursts and prevent them from smearing to adjacent traces. More advanced weighted summation taking into account intra-array statics is proposed by Neklyudov et al. (2015). In this paper we will focus on efficient implementation of the basic simple stack that is the backbone of all approaches.

Efficient implementation of four-dimensional supergrouping in practice requires proper organization of the input data. We propose to use shot records, which are organized by swaths (Figure 3). For such input one needs to include neighboring swaths to construct a shot group to avoid edge effects. The number of swaths depends on shot group size in the crossline direction. Figure 4 shows a workflow for supergrouping of shots and receivers. It includes two sorting steps and two supergrouping steps, in common-receiver and common-shot domains respectively. The first pass of supergrouping is done in the common-receiver domain to minimize the amount of data for the second sorting and supergrouping phase. One of the most time consuming operations when dealing with high-channel count and single-sensor data is searching for neighboring traces from the large ensemble of traces (common shot or common receiver). Brute-force searching that sequentially checks coordinates of each trace becomes extremely slow for larger ensembles. Instead, we propose to use a R-tree data structure (Manolopoulos et al., 2010) that is designed for efficient processing of the spatial queries. The main idea of the R-tree is to group nearby objects and represent such groups using minimum bounding rectangles (MBR). Each internal node of the R-tree contains a set of MBRs and corresponding pointers to the child nodes (Figure 5). Each node in the R-tree bounds their children. For a given query window, we recursively check only the nodes and their children whose MBRs overlap the query window. It allows the algorithm to skip most of the nodes in the R-tree and optimize searching time. Figure 6 shows a comparison of the searching time within common shot gathers using an R-tree index provided by Boost libraries (http://www.boost.org/).

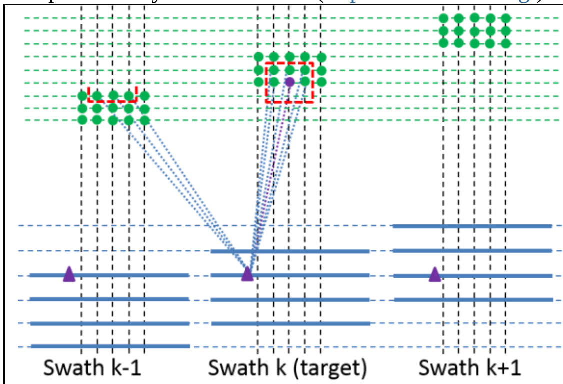

Figure 3: Data required for four-dimensional shot and receiver supergrouping of the central swath of the data. To construct a shot supergroup without edge effects we have to use neighboring swaths. Here, for visualization purposes, the horizontal extents of the receiver lines are shown separately for each swath although in reality they should coincide.

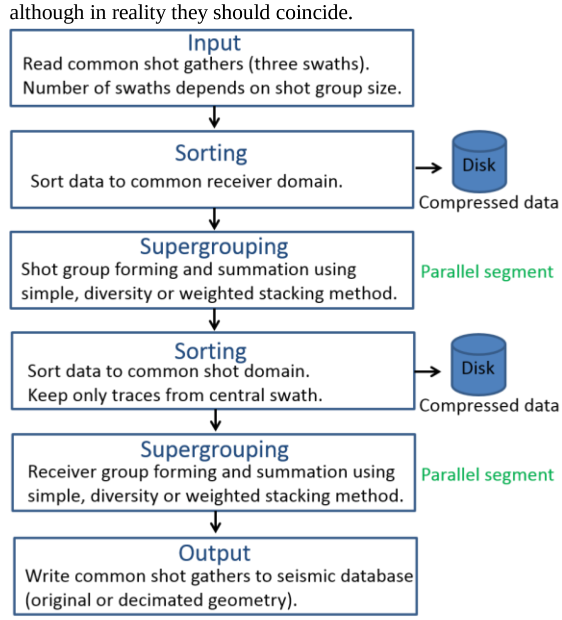

Figure 4: A computationally efficient workflow for four-dimensional supergrouping (shot and receiver domains).

One can see that using R-tree data structure greatly reduces search time especially for high-channel count and single-

sensor data, providing a computation speedup of 24X and 41X respectively. To sort huge data volumes we use an efficient sorting technique proposed by Luo et al. (2003).
This computationally efficient implementation of four-dimensional supergrouping allows to process one swath of single-sensor data within 10 hours using only one computational node. Note that the input is three swaths of data (main swath plus two neighbors) comprising a data volume of around 1.5 TB.

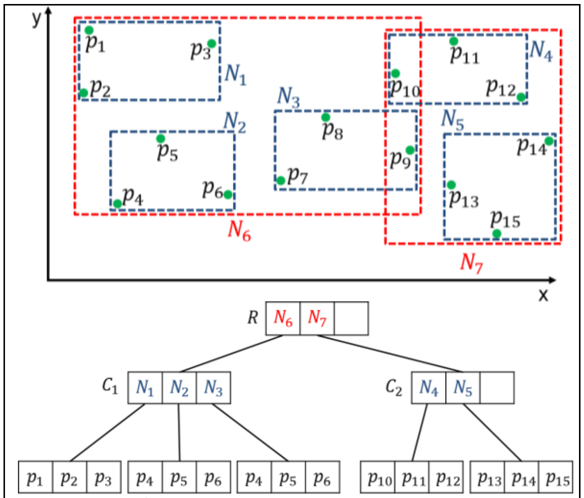

Figure 5: Set of points indexed by R-tree (top) and corresponding R-tree data structure (bottom).

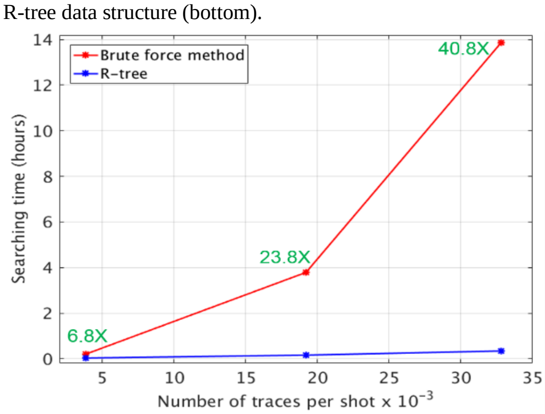

Figure 6: Search time comparison for 50,000 shot gathers using a 5x3 receiver supergroup. Observe significant speedup of the method based on the R-tree compared to a brute-force approach especially for larger ensemble sizes typical of high-channel count and single-sensor data.

Enhancement of low frequencies for FWI

Let us show application of supergrouping for FWI. A lack of low frequencies in seismic data is a central issue in FWI (Virieux et al., 2009). Due to strong nonlinearity, the natural way to apply FWI on bandlimited data is to use a multi-scale hierarchical approach proceeding sequentially from low to higher frequencies. The key element of this multi-scale strategy is availability of low frequencies such as 1.5 to 3 Hz that are at the edge of the spectrum of conventional broadband data. Obtaining such information from broadband land data in an arid desert environment represents a significant challenge because of poor SNR. We demonstrate the effect of supergrouping on 3D broadband land seismic dataset from Saudi Arabia acquired using a nonlinear sweep from 2 to 90 Hz.

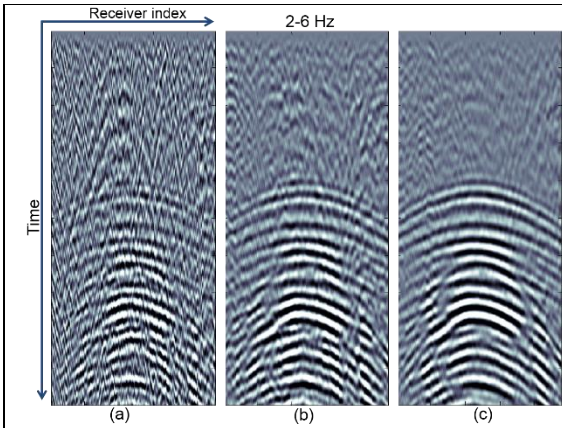

Figure 7: Broadband field data for the far cable bandpass filtered to 2-6 Hz showing (a) raw data (single shot gather), (b) standard symmetric supergroup 5x5 (200 m for shots in crossline direction and 200 m for receivers in inline direction), and (c) symmetric supergroup 5x5 followed by common-offset summation in the inline direction using five shots.

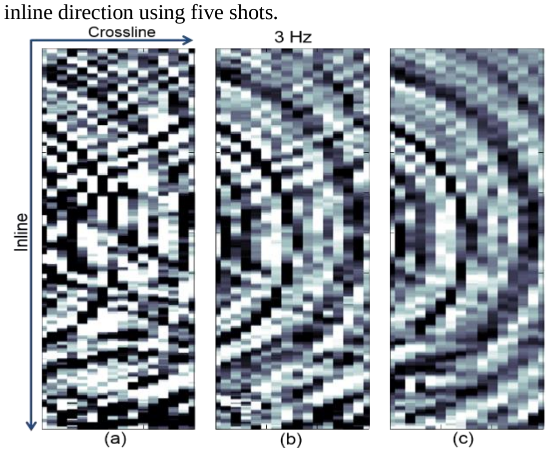

Figure 8: Frequency-domain panels at 3 Hz including (a) raw data (single shot gather), (b) standard symmetric supergroup 5x5 (200m for shots in crossline direction and 200 m for receivers in inline direction), and (c) symmetric supergroup 5x5 followed by common offset-summation in inline direction using five shots.

Acquisition parameters for this dataset are shown in Figure 1 (low-density data). We perform basic preprocessing before supergrouping comprising suppression of noisy traces with high amplitudes and linear noise removal for groundroll. Figure 7 shows the low-frequency component (2-6 Hz) of the data using a two-dimensional symmetric 5x5 shot and receiver supergroup (200 m for shots in crossline direction and 200 m for receivers in inline direction) and more general three-dimensional supergroup obtained as 5x5 shot and receiver group followed by common-offset summation in the inline direction using five shots. One can see that supergrouping significantly improves SNR in both time and frequency domains (Figure 8). The two-stage three-dimensional supergroup produces the best results. Stacking the common-offset traces from several neighboring shot lines allows additionally to suppress incoherent events in the sparsely sampled inline direction and highlight reflections.

Automatic first-break picking

It may sound surprising, but for modern data with reduced field arrays even the strongest arrivals representing refracted waves may lack proper SNR needed for successful first-break picking and near-surface modeling. For large 3D datasets, automatic first-break picking is a must and some minimum SNR is required for these algorithms to work. Figures 9a and 9b show that such quality is no longer achieved with high-channel count data using small field arrays, in this case - nine geophones in a sensor array and two vibrators in a source group. Other typical acquisition parameters are shown in Figure 1 for high channel count acquisition. Clearly, picking fails in many places because of lack of sufficient signal (Figure 9a and 9b). Standard ways of data preconditioning help, but still lead to inconsistent time picks (Figures 9c and 9d). As a consequence, tomographic inversion and resulting statics will have higher degree of error and uncertainty and this is just the first challenge for this modern data. When we apply a combination of 3x5 shot supergrouping on top of conventional preprocessing we obtain results shown in Figures 9e and 9f. There is a clear improvement for both near- and far-offset regions. While we may lose some details, they represent high-frequency statics that can be addressed by residual statics algorithms. We obtain much more reliable travel times for refraction tomography, which constrains the long-wavelength statics critical for proper delineation of subtle structures.

Conclusion

A simple and computationally efficient algorithm for four-dimensional supergrouping was presented to enhance challenging pre-stack data. The four dimensions are typically represented by inline and crossline directions of both sources and receivers, although other domains are also possible. Despite its simplicity, supergrouping becomes computationally demanding for high-channel count data where data volumes reach hundreds of terabytes or more. To achieve computational efficiency we use an optimized R-tree method for searching neighboring traces. We have shown that the R-tree method provides significant computational speedup that increases with the ensemble size. Without such optimization, search times alone could become overwhelming for high-channel count data. A two-stage implementation of supergrouping (each using two coordinates e.g. source X and source Y) allows to process large data ensembles independently. Supergrouping allows to obtain significantly improved signal-to-noise ratio for pre-stack data that is fundamental to the entire land processing sequence from first-break picking to time processing to velocity model building using tomography or FWI. We have presented field examples that demonstrate these improvements for FWI and first-break picking. In addition, supergrouping can be used as a data decimation tool to reduce the total computational cost of FWI. Each pre-stack trace can be efficiently utilized to suppress noise and enhance the resulting signal and no single trace is left behind. Supergroup geometry is decided in processing and as such can be adaptively changed for the specific scale of interest.

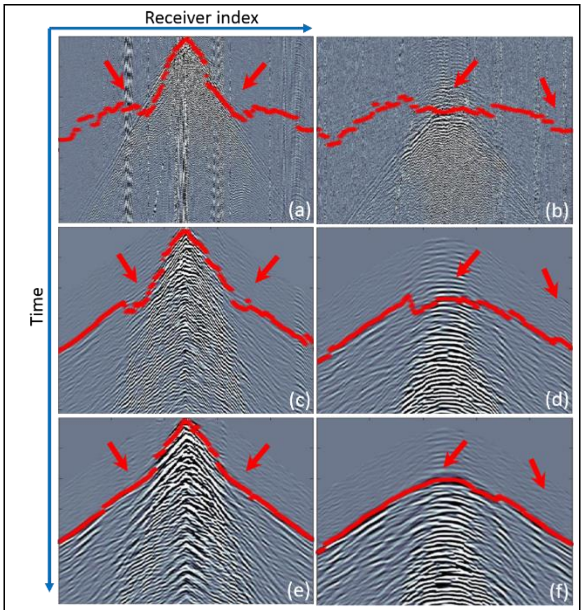

Figure 9: An example of data enhancement for automatic first break picking showing (a) and (b) raw displays of field data, (c) and (d) the same data after standard preprocessing for automatic first-break picking and (e) and (f) after 3x5 shot supergrouping (5-element group in crossline and 3-element group in sparsely sampled inline direction implement using common-offset summation). Observe more robust picks on both near and far cables using supergrouped data.

Acknowledgments

We would like to thank Saudi Aramco for support and permission to publish this work. We thank Saud Saadi, Saudi Aramco, for his assistance with the second data example. We are grateful to Simon Cordery and Pierre Leger, Saudi Aramco, for their valuable comments and discussions.

EDITED REFERENCES

Note: This reference list is a copyedited version of the reference list submitted by the author. Reference lists for the 2017 SEG Technical Program Expanded Abstracts have been copyedited 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

Bakulin, A., P. Golikov, M. Dmitriev, V. Dolgov, P. Leger, and D. Neklyudov, 2016, Application of supergrouping to land seismic data in a desert environment: 86th Annual International Meeting, SEG, Expanded Abstracts, 4649–4563, https://doi.org/10.1190/segam2016-13838765.1.

Luo, Y., M. Huwadi, K. P. Gunaratnam, and M. N. Alfaraj, 2003, A fast sorting algorithm for large volume datasets: First Break, **2**, 29–31.

Martinez, R. D., J. R. Masters, and D. Fryar, 1993, Prestack noise attenuation using the diversity Tau-p transform: 63rd Annual International Meeting, SEG, Expanded Abstracts, 1187–1190, https://doi.org/10.1190/1.1822330.

Manolopoulos, Y., A. Napoulos, A. N. Papadopoulos, and Y. Theodoridis, 2010, R-Trees: Theory and Applications. Springer Science & Business Media.

Neklyudov, D., A. Bakulin, M. Dmitriev, and P. Golikov, 2015, Enhancement 3D broadband land seismic data with smart super groups for processing and FWI: 77th Annual International Conference and Exhibition, EAGE, Extended Abstracts, P7 06, https://doi.org/10.3997/2214-4609.201413435.

Vermeer, G. J. O., 2005, 3-D seismic survey design: SEG, 342 pp.

Virieux, J., and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, **74**, no. 6, WCC1–WCC26, https://doi.org/10.1190/1.3238367.