Sparsity-promoting recovery from simultaneous data: a compressive sensing approach

Haneet Wason, Felix J. Herrmann and Tim T.Y. Lin, Seismic Laboratory of Imaging and Modeling - The University of British Columbia*

Summary

Seismic data acquisition forms one of the main bottlenecks in seismic imaging and inversion. The high cost of acquisition work and collection of massive data volumes compel the adoption of simultaneous-source seismic data acquisition - an emerging technology that is developing rapidly, stimulating both geophysical research and commercial efforts. Aimed at improving the performance of marine- and land-acquisition crews, simultaneous acquisition calls for development of a new set of design principles and post-processing tools. Leveraging developments from the field of compressive sensing the focus here is on simultaneous-acquisition design and sequential-source data recovery. Apart from proper compressive sensing sampling schemes, the recovery from simultaneous simulations depends on a sparsifying transform that compresses seismic data, is fast, and reasonably incoherent with the compressive-sampling matrix. Using the curvelet transform, in which seismic data can be represented parsimoniously, the recovery of the sequential-source data volumes is achieved using the sparsity-promoting program — SPGL1, a solver based on projected spectral gradients. The main outcome of this approach is a new technology where acquisition related costs are no longer determined by the stringent Nyquist sampling criteria.

Introduction

Seismic exploration techniques involve the collection of massive data volumes—where sampled wavefields exhibit up to 5-dimensional structure—and their exploitation during processing. Consequently, some of the fundamental shortcomings in our workflow are related to the Nyquist sampling criteria and the 'curse of dimensionality' which results in an exponential increase in volume on addition of extra dimensions to our data collection.

In this paper, we propose an alternate sampling scheme adapted from the field of "compressive sensing" which is aimed at removing these impediments via dimensionality reduction techniques based on randomized subsampling. With this dimensionality reduction, we arrive at a sampling framework where the sampling rates no longer scale directly with the size and the desired resolution of our acquisition areas, but with the transform-domain compression; more compressible data requires less sampling.

Compressive sensing overview

Compressive sensing (abbreviated as CS throughout the paper) is a process of reconstructing a signal utilizing the prior knowledge that it is sparse or compressible in some transform domain. The core idea of CS is a novel sampling technique, which under certain conditions can lead to smaller sampling rate compared to the conventional Nyquist sampling rate. CS is based on three key elements: randomized subsampling, sparsifying transforms and sparsity-promotion recovery by convex optimization.

One of the main advantages of CS is that it combines transformation and encoding in a single linear step, resulting in a direct application of this technology in seismic acquisition where the acquisition costs are quantified by the transform-domain sparsity of seismic data instead of the grid size. This scheme aims to design acquisition surveys in a way that renders the randomized subsampling related artifacts—whether caused by periodic missing traces or by crosstalk between simultaneous sources—harmless by turning them into incoherent Gaussian noise that can be easily removed during processing.

By solving a sparsity-promoting problem (Candès et al., 2006; Donoho, 2006; Herrmann et al., 2007; Mallat, 2009), we reconstruct high-resolution data volumes from the randomized samples at the moderate cost of a minor oversampling factor compared to data volumes obtained after conventional compression (see e.g. Donoho et al., 1999, for wavelet-based compression). With sufficient sampling, this nonlinear recovery outputs a set of largest transform-domain coefficients that produces a reconstruction with a recovery error comparable with the error incurred during conventional compression. As in conventional compression this error is controllable, but in the case of CS this recovery error depends on the sampling ratio—i.e., the ratio between the number of samples taken and the number of samples of the high-resolution data. Because compressively sampled data volumes are much smaller than high-resolution data volumes, we reduce the dimensionality and hence the costs of acquisition, storage, and possibly of data-driven processing.

The sparse recovery problem involves solving an undertermined system of equations

$$\mathbf{b} = \mathbf{A}\,\mathbf{x}_0 \qquad (1)$$

where $\mathbf{b} \in R^n$ represents the compressively sampled data of n measurements, the compressive sensing (or measurement) matrix $\mathbf{A} \in R^{n \times N}$ represents the sampling operator that collects the acquired samples from a high-resolution data $\mathbf{f_0} \in R^N$, which has a sparse reperesentation $\mathbf{x_0} \in R^N$. Therefore, the sparsity-promoting recovery is achieved via solving the convex optimization problem (also known as the 'Basis Pursuit' (BP) problem)

$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{subject to} \quad \mathbf{b} = \mathbf{A}\,\mathbf{x_0}, \qquad (2)$$

with $\mathbf{x}$ representing the estimated coefficients and the l_1 norm $\|\mathbf{x}\|_1$ is the sum of absolute values of the coefficients in the vector $\mathbf{x}$. The optimization problem, therefore, finds a sparse or (under some conditions) the sparsest solution that explains data exactly.

Invoking techniques of CS, the recovery of a k-sparse signal—i.e., with $k \leq N$ non-zeros in $\mathbf{x}$—is possible as long as any subset S of k columns of the $n \times N$ matrix $\mathbf{A}$ behaves approximately like an orthogonal basis. In this case, a restricted isometry constant δ_k —which bounds the energy of the signal—is the smallest quantity that can be defined such that

$$(1 - \delta_k)\, \|\mathbf{x}_S\|_2 \;\leq\; \|\mathbf{A}_S\mathbf{x}_S\|_2 \;\leq\; (1 + \delta_k)\, \|\mathbf{x}_S\|_2 \qquad (3)$$

for all subsets S, with the cardinality of S — $|\,S\,| \;\leq k$. The mutual coherence between the columns of $\mathbf{A}$, $\mu(\mathbf{A})$, provides a bound on the restricted isometry constant δ_k as

$$\delta_k \leq (\,k - 1\,)\,\mu(\mathbf{A}) \qquad (4)$$

with

$$\mu(\mathbf{A}) \;=\; \max_{1 \leq i \neq j \leq N} \;|\,\mathbf{a}_i^H\,\mathbf{a}_j\,| \,/\, (\,\|\,\mathbf{a}_i\,\|_2 \cdot \|\,\mathbf{a}_j\,\|_2\,), \qquad (5)$$

where $\mathbf{a}_i$ is the i^{th} column of $\mathbf{A}$ and H denotes the Hermitian transpose. Hence, the mutual coherence of a matrix $\mathbf{A}$ is the largest absolute normalized inner product between different columns from $\mathbf{A}$, which is a way to characterize the dependence between the columns of $\mathbf{A}$ (Bruckstein et al., 2009). For a successful (CS) recovery, the mutual coherence between the columns of the measurement matrix $\mathbf{A}$ should be small, consequently, a smaller δ_k captures more signal-energy leading to a more stable inversion of $\mathbf{A}$ for signals $\mathbf{x}$ with maximally k non-zero entries.

Random matrices with Gaussian *i.i.d.* entries with variance n^{-1} have a small mutual coherence—they contain subsets of k columns that are incoherent, and hence a small δ_k. The sparse recovery problem (Equation (2)) recovers the k non-zero coefficients exactly as long as

$$k \;\leq\; C \cdot \mathrm{n} \,/\, (\log_2 (N/n)) \qquad (6)$$

with C as a constant. This proves an important result wherein the recovery of k non-zeros only requires an oversampling ratio of $n/k \approx C \cdot \log_2 N$, as opposed to taking all N measurements, for large N - typical for seismic data. Therefore, only a few number of measurements ($n << N$) are required to recover the nonzeros.

In a nutshell, according to CS, high accuracy recovery is possible when — i) $\mathbf{x_0}$ is sufficiently sparse—i.e., it has a sparsifying representation that exploits the structure of the signal by mapping the signal energy into a small number of significant transform-domian coefficients, ii) the measurement matrix renders the coherent subsampling artifacts incoherent, i.e., behaves like a Gaussian matrix.

Experimental setup

Compressive sensing provides powerful tools for acquiring signals that have a sparse representation in some transform domain via sampling strategies/rates that are small compared to the conventional Nyquist sampling rate. However, exploiting these tools in exploration seismology is not readily feasible, based on the previously outlined mathematical formulation of CS, due to the complex nature of seismic data in high dimensions. Hence, our focus is specifically on the design of source subsampling schemes that favor recovery and on the selection of the appropriate sparsifying transform.

Seismic data permits sparse representation with multiscale and multidirection transforms that capture the "wavefront-set" of the subsurface reflectors. By construction, curvelets are well adapted to data with wavefront-like features, hence, are well suited for representing seismic data parsimoniuosly as the elements of this tranform behave approximately as high-frequency asymptotic eigenfunctions of wave equations (see e.g. Smith, 1998; Candès and Demanet, 2005; Candès et al., 2006a; Herrmann et al., 2008). Therefore, we use curvelet transform as the sparsifying transform in our study.

During seismic data acquisition, data volumes are collected that represent discretizations of analog finite-energy wavefields in two or more dimensions including time. We recover the discretized wavefield $\mathbf{f}$ by inverting the compressive-sampling matrix

$$\mathbf{A} \overset{\mathbf{def}}{=} \mathbf{R}\,\mathbf{M}\,\mathbf{S^H}, \qquad (7)$$

where $\mathbf{R}$ is a restriction matrix acting on the measurement matrix $\mathbf{M}$, and $\mathbf{S^H}$ is the sparsifying synthesis matrix, with the sparsity-promoting program

$$\tilde{\mathbf{f}} = \mathbf{S^H}\,\mathbf{x} \quad \text{with} \quad \tilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{subject to} \quad \mathbf{b} = \mathbf{A}\,\mathbf{x} \quad (8)$$

This formulation differs from standard compressive sensing because we allow for a wavefield representation that is redundant—i.e., $\mathbf{S} \in \mathrm{C}^{P \times N}$ with $\mathrm{P} \geq \mathrm{N}$. The algorithm for the recovery problem involves the following steps:

1. Inversion of an underdetermined system by promoting signal sparsity.
2. Exploit sparsity in the transform (curvelet in our case) domain as a prior.
3. Find the sparsest set of coefficients that explain/match the data using the one-norm minimization solver that is capable of handling large systems. We use SPGL1 (Berg and Friedlander, 2008).

4. Recover the sequential-source data via synthesis, i.e., apply the inverse curvelet transform to the coefficients.

We present two simultaneous-source acquisition approaches by studying the performance of our algorithm on a seismic line from the Gulf of Suez (Figure 1). The two approaches differ in the definition of the measurement matrix.

Approach I. Random amplitude encoding

This approach can be considered as the simultaneous 'Land' acquisition where the sequential impusive sources are replaced with the impulsive simultaneous 'phase-encoded' sources. Mathematically, measurements with the phase-encoded sources are obtained by replacing the identity matrix (the source-sampling matrix in sequential source acquisition) with the measurement matrix defined by:

$$\mathbf{M} \stackrel{\text{def}}{=} [\, \mathbf{I} \otimes \mathbf{diag}(\eta)\, \boldsymbol{F_s}^{*} \mathbf{diag}(\boldsymbol{e^{i\theta}})\, \boldsymbol{F_s} \otimes \mathbf{I} \,]\,, \qquad (9)$$

where the action of the identity matrix is replaced with the action of 'phase-encoded' matrix, which is a combination of applying the Fourier transform (F_s) along the source coordinate, uniformly drawn random random phase rotations $\theta \in [0,\pi]$, an inverse Fourier transform (F_s^*) and multiplication with a random sign vector diag(sign(η)) with $\eta \in \mathrm{N}(0,1)$. This formulation follows from the work by Romberg (2009) and Herrmann et al. (2009) wherein the above mentioned combination corresponds to the action of a Gaussian matrix. Application of this measurement matrix turns the sequential sources into simultaneous sources, i.e., one simultaneous 'supershot' wherein all the sources fire simultaneously, and the restriction matrix selects a subset (n_s') of these supershots generated by different randomly-weighted simultaneous sources. The (restricted) measurement matrix now has an aspect ratio of n_s' / n_s and the recovery problem boils down to solving an underdetermined system (n_s' << n_s) of linear equations. Figure 2 illustrates the transition in the data collected from a conventional sequential-source experiment to simultaneous source experiments. In reality, this sort of sampling is perhaps physically unrealizable—i.e., we typically do not have large numbers of vibroseis trucks available—it gives us the most favorable recovery conditions from the CS perspective. Therefore, our 'Land' acquisition will serve as a benchmark with which we can compare alternative and physically more realistic acquisition scenarios.

Approach II. Random time dithering

This approach can be considered as the simultaneous 'Marine' acquisition where the sequential acquisition with a single airgun is replaced with continuous acquisition with multiple airguns firing at random times and at random locations. Here, the sampling operator is defined as:

$$\mathbf{RM} \stackrel{\text{def}}{=} [\, \mathbf{I} \otimes \mathbf{T} \,]\,, \qquad (10)$$

where the linear operator T turns the sequential source recordings into continuous recordings with n_s* impulsive sources firing at random positions, selected uniformly-random from [1.... n_s] discrete source indices and from discrete random time indices, selected uniformly from (0....(n_s* - 1) x n_t) time indices. The operator T acts on both the time and the source coordinate. Hence, in this scenario, a seismic line is mapped into a single long 'supershot' that consists of a superposition of n_s* impulsive shots. Figure 3 represents a subset of this long record. Notice that this type of 'Marine' acquisition is physically realizable as long as the number of simultaneous sources involved is limited.

Discussion and conclusions

We simulate 'Land' data for (subsampling ratio) δ = 0.5, i.e., 64 simultaneous source experiments with all sources firing and study the recovery based on 3-D curvelets. We conduct a similar experiment for the 'Marine case' where we randomly select 128 shots from the total survey time $T = \delta \times (n_s - 1) \times T_0$, yielding the same aspect ratio for the sampling matrix. Using the 3-D curvelet transform, which attains higher sparsity because it explores continuity of the wavefield along all three coordinate axes, the recovery results for 'Land' and 'Marine' acquisition are summarized in Figures 4 and 5. One of the observations made is that it is clear that accurate recovery is possible by solving an L1-optimization problem using SPGL1 (Berg and Friedlander, 2008) while limiting the number of iterations to 200. Second, 'Land' acquisition clearly favors recovery by curvelet-domain sparsity promotion compared to 'Marine' acquisition. This is true despite the fact that the subsampling ratio, i.e., the aspect ratio of the sampling matrices, are the same. This becomes clear by realizing that the difference lies in the mutual coherence of the two sampling matrices where the columns of the sampling matrix for 'Land' acquisition are more incoherent and hence more independent. This favors (better) recovery in the 'Land' acquisition scenario. The SNR for 'Land' acquisition is computed to be 11.5dB, and 11.1dB for 'Marine' acquisition, which justify the above made observations. Notice that 'Marine' acquisition works relatively so well because it randomizes along two coordinates—time and source, as opposed to randomizing along the source coordinate only as in 'Land' acquisition.

In summary, following ideas from CS, seismic wavefields can be reconstructed from randomized subsamplings. Acquisition and processing costs are no longer determined by the resolution and size of the acquisition survey, rather, they scale with transform - domain sparsity of the wavefield, and a new paradigm for randomized processing and inversion. Recovery from simultaneous simulations depends on transform-domain sparsity wherein sparser signals—i.e., signals with a small number of significant transform-domain coefficients permit better recovery. This new sampling paradigm can be successfully exploited in various problems in exploration seismology to effectively repulse the curse of dimensionality.

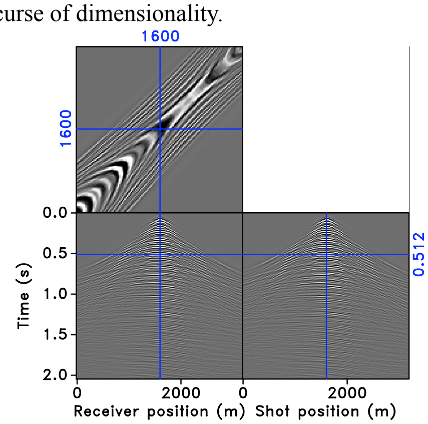

Figure 1. Fully sampled sequential data from conventional sequential acquisition with 128 shots, 128 receivers and 512 time samples.

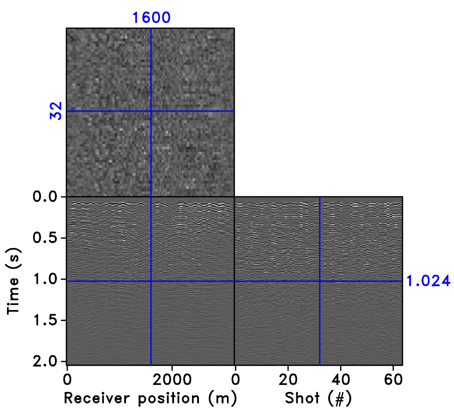

Figure 2. Compressively sampled ‘Land’ data ($\delta = 0.5$).

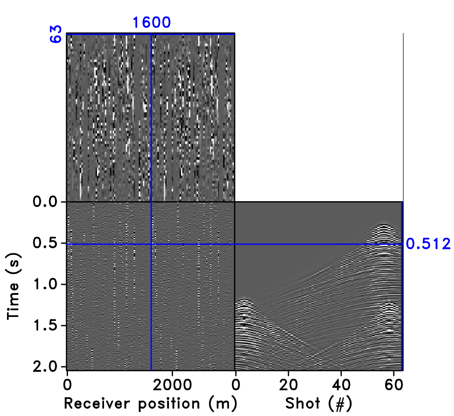

Figure 3. Compressively sampled ‘Marine’ data ($\delta = 0.5$).

Acknowledgements

We would like to thank the authors of CurveLab (curvelet.org), a toolbox implementing the Fast Discrete Curvelet Transform, WaveAtom a toolbox implementing the Wave Atom transform (http://www.waveatom.org/), Madagascar (rsf.sf.net), a package for reproducible computational experiments, SPGL1(cs.ubc.ca/labs/scl/spgl1), SPOT (http://www.cs.ubc.ca/labs/scl/spot/), a suite of linear operators and problems for testing algorithms for sparse signal reconstruction, and pSPOT, SLIM’s parallel extension of SPOT. The industrial sponsors of the Seismic Laboratory for Imaging and Modelling (SLIM) BG Group, BP, Chevron, ConocoPhillips, Petrobras, Total SA, and WesternGeco are also gratefully acknowledged.

[Note: SNR = $-20 \log(\| \mathbf{f} - \tilde{\mathbf{f}} \|_2 / \| \mathbf{f} \|_2)$]

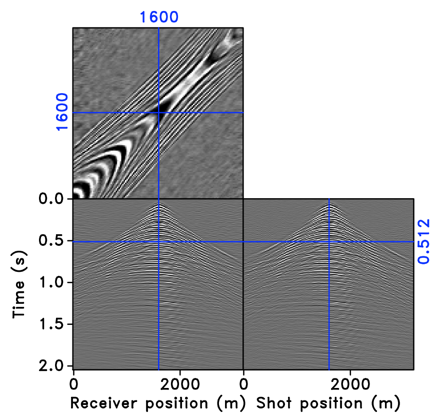

Figure 4. Recovery from ‘Land’ data (SNR = 11.5dB).

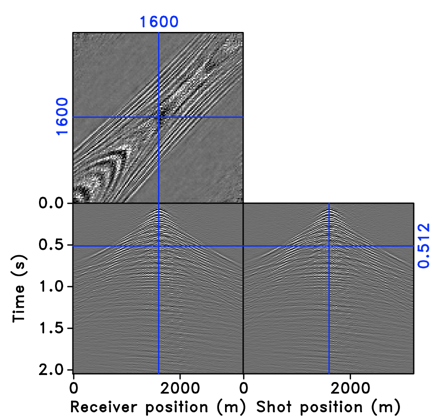

Figure 5. Recovery from ‘Marine’ data (SNR = 11.1dB).

References

Berg, E. v., and M. P. Friedlander, 2008, Probing the Pareto frontier for basis pursuit solutions: Technical Report 2, Department of Computer Science, University of British Columbia.

Bruckstein, A. M., D. L. Donoho, and M. Elad, 2009, From Sparse Solutions of Systems of Equations to Sparse Modeling of Signals and Images: SIAM Review, **51**, 34-81.

Candès, E., J. Romberg, and T. Tao, 2006, Stable signal recovery from incomplete and inaccurate measurements: Comm. Pure Appl. Math., **59**, 1207–1223.

Candès, E. J., and L. Demanet, 2005, The curvelet representation of wave propagators is optimally sparse: Comm. Pure Appl. Math, **58**, 1472–1528.

Candès, E. J., L. Demanet, D. L. Donoho, and L. Ying, 2006a, Fast discrete curvelet transforms: Multiscale Modeling and Simulation, **5**, 861–899.

Donoho, D. L., 2006, Compressed sensing: IEEE Trans. Inform. Theory, **52**, 1289–1306.

Donoho, P., R. Ergas, and R. Polzer, 1999, Development of seismic data compression methods for reliable, low-noise performance: SEG International Exposition and 69th Annual Meeting, 1903–1906.

Herrmann, F. J., P. P. Moghaddam, and C. C. Stolk, 2008, Sparsity- and continuity- promoting seismic imaging with curvelet frames: Journal of Applied and Computational Harmonic Analysis, **24**, 150–173. (doi:10.1016/j.acha. 2007.06.007).

Herrmann, F. J., U. Boeniger, and D. J. Verschuur, 2007, Non-linear primary-multiple separation with directional curvelet frames: Geophysical Journal International, **170**, 781–799.

Herrmann, F. J., Y. A. Erlangga, and T. Lin, 2009, Compressive simultaneous full-waveform simulation: Geophysics, **74**, A35.

Mallat, S. G., 2009, A Wavelet Tour of Signal Processing: the Sparse Way: Academic Press.

Romberg, J., 2009, Compressive sensing by random convolution: SIAM Journal on Imaging Sciences, **2**, 1098–1128.

Smith, H. F., 1998, A Hardy space for Fourier integral operators.: J. Geom. Anal., **8**, 629–653.