geoconvention Virtual Event September 21-23 2020
CSPG • CSEG • CWLS • GAC • MAC • IAH

Sensor Placement Optimization for Seismic Data Acquisition and Shot Reconstruction

Yi Guo, Rongzhi Lin and Mauricio D. Sacchi

Signal Analysis and Imaging Group (SAIG), Department of Physics, University of Alberta

Summary

This article explores a recently proposed method for optimizing sensor placement that we adapted for optimal shot gather reconstruction. The method uses Proper Orthogonal Decomposition (POD) to extract bases from a training dataset. Dimensionally reduction methods such as singular value decomposition, and QR pivoting are used to determine optimal receiver geometry. The technique of least-squares is used to reconstruct the shot from the POD bases.

Introduction

In applications where either deploying sensors are costly, or constraints limit a dense coverage, the design of optimal sensor placement can reduce the overall acquisition cost. However, determining optimal sensor locations is an NP-hard problem. Instead of trying a brute-force search among all the combinatorial possibilities formed by the positions of the sensors, we test a new sensor placement design method recently proposed by Manohar et al. (2018).

Machine learning techniques usually take advantage of the dominant features which are exhibited in a dataset with similar patterns for classification or other purposes. These features can often be identified using dimensionality reduction techniques, for instance, via Proper Orthogonal Decomposition (POD). In this paper, optimized receiver locations for seismic data acquisition and reconstruction are designed based on a tailored set of features extracted from either synthetic or real seismic data. These bases are extracted via the POD method. The optimal receiver locations are computed using QR pivoting and singular value decomposition (SVD) methods. Shot gathers can then be reconstructed with minimal distortion via a least-squares inversion that uses training data.

Theory

In general, natural signals are highly compressible. That is to say, when the signal is transformed into another appropriate domain, an economical representation in terms of a few basis functions is attainable, and missing data could be received from a small number of measurements. Rather than acquiring the full measurements in the first stage, the theory of compressed sensing states that it may be possible to collect a compressed version of the data directly and then estimate the coefficients that model the data via sparse inversion.

Compressed sensing strategies are ideal for the recovery of unknown compressible signals from random measurements on a universal basis (Baraniuk, 2007). However, if the information is available about the signal, it is possible to design bases that are particularly tailored for a given signal (Manohar et al., 2018) and, therefore, find a representation that can be used for signal recovery.

Proper orthogonal decomposition (POD) is a data-driven dimensionality reduction technique, which expresses $\mathbf{x} \in \mathbb{R}^n$ as linear combinations of several orthonormal eigenmodes Ψ. On a tailored basis $\Psi_r \in \mathbb{R}^{n \times r}$, the data $\mathbf{x}$ may have a low-rank representation

$$\mathbf{x} = \Psi_r \mathbf{a}, \tag{1}$$

where $\mathbf{a} \in \mathbb{R}^r$. The eigenmodes Ψ_r and POD coefficients $\mathbf{a}$ can be obtained via the SVD. This low-rank embedding requires data for training. Given a data matrix $\mathbf{X} = [\mathbf{x}_1\ \mathbf{x}_2 \ldots \mathbf{x}_m]$, the resulting eigenmodes are the orthonormal left singular vectors Ψ of $\mathbf{X}$ obtained via the SVD

$$\mathbf{X} = \Psi \Sigma \mathbf{V}^T \approx \Psi_r \Sigma_r \mathbf{V}_r^T. \tag{2}$$

The matrix Ψ_r and $\mathbf{V}_r$ contains the first r columns of Ψ and $\mathbf{V}$, and the diagonal matrix Σ_r contains the first $r \times r$ block of Σ. For low-rank datasets, the singular values have a fast decay. In this paper, the optimal hard threshold algorithm based on the singular value distribution and aspect ratio of the data matrix is used for choosing the target rank r (Gavish and Donoho, 2013).

The measurement or sampling matrix $\mathbf{C} \in \mathbb{R}^{p \times n}$ is constructed in the following way

$$\mathbf{C} = \left[\mathbf{e}_{\gamma 1}\ \mathbf{e}_{\gamma 2} \ldots \mathbf{e}_{\gamma p}\right]^T, \tag{3}$$

where $\mathbf{e}_j$ represents the basis vector with a unit entry at index j and zeros elsewhere. Then, the measurement $\mathbf{y}$ consists of p elements selected from $\mathbf{x}$

$$\mathbf{y} = \mathbf{C}\mathbf{x} = \left[\mathbf{x}_{\gamma 1}\ \mathbf{x}_{\gamma 2} \ldots \mathbf{x}_{\gamma p}\right]^T, \tag{4}$$

where the index set of $\gamma = [\gamma_1 \ldots \gamma_p]$ denotes sensor location. Combining equation 1 and equation 4 yields the optimal sensing problem

$$\mathbf{y} = \mathbf{C}\Psi_r \mathbf{a} = \Theta \mathbf{a}. \tag{5}$$

When $\mathbf{x}$ is unknown, the reconstruction can be obtained using (Manohar et al., 2018)

$$\hat{\mathbf{x}} = \Psi_r \hat{\mathbf{a}}, \text{ where } \hat{\mathbf{a}} = \begin{cases} \Theta^{-1}\mathbf{y} = (\mathbf{C}\Psi_r)^{-1}\mathbf{y}, & p = r, \\ \Theta^{\dagger}\mathbf{y} = (\mathbf{C}\Psi_r)^{\dagger}\mathbf{y}, & p > r \end{cases}. \tag{6}$$

The sensor placement design finds rows of Ψ_r, which corresponds to the sensor locations that optimally condition the inversion of the sensing matrix Θ. Following (Manohar et al., 2018), we adopted the QR factorization with column pivoting to decompose the matrix Ψ_r

$$\Psi_r^T \mathbf{C}^T = \mathbf{Q}\mathbf{R}, \tag{7}$$

where $\mathbf{Q}$ is a unitary matrix, $\mathbf{R}$ is an upper-triangular matrix, and $\mathbf{C}$ represents a column permutation matrix. The algorithm of QR column pivoting increments the volume of the submatrix by enforcing a diagonal dominance structure. Therefore, it optimally conditions the measurement or row-selected POD bases and yields the r point sensors that best sample the r basis modes Ψ_r.

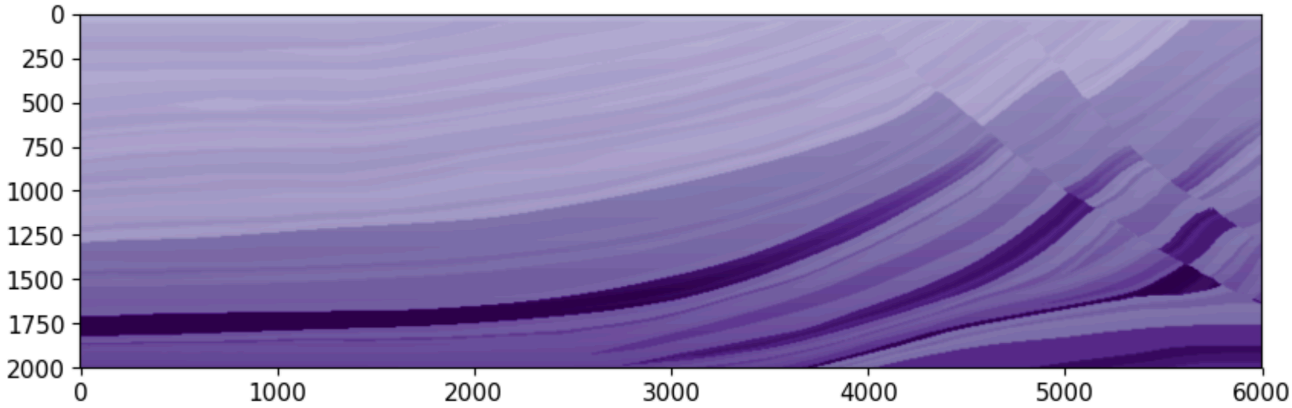

Figure 1: 2D velocity model.

Example

For the synthetic data, we generate 2D shot gathers by finite-difference modeling using the SeisAcoustic package from SeismicJulia (Stanton and Sacchi, 2016). The velocity model of the Marmousi model (Figure 1) is relatively flat on the left side but becomes complicated in the right side. Both the shots and receivers are placed at the surface of the model, and there are in total 160 receivers and 120 shots.

In this example, the first 60 shot gathers are incorporated into the training dataset, and the shot number 61 is the testing gather, which is not included in the training dataset. The data are low rank, and Figure 2 illustrates the decay of the singular values. In Figure 2, the blue dots represent the modes that are used in the reconstruction stage by applying the optimal hard threshold method for singular values.

In this case, the minimal number of sensors used is 24. Figure 3a) is the original shot gather for shot number 61. The difference between random sensors and optimal sensors are compared. The randomly selected data adopted the jittered sampling strategy proposed by Hennenfent and Herrmann (2008) to control the maximum gap size. Figure 3b) is the randomly sampled data with 24 receivers, and Figure 3c) is the corresponding reconstruction. On the other hand, Figure 3d) is the optimal sampled data with 24 optimal sensors, and Figure 3e) is the reconstructed result based on optimal sampling. It is noticeable that the optimal sampled reconstruction almost recovered all the essential features of the original dataset. In contrast, the randomly sampled reconstruction result is worse than the optimal result. Further, the optimal sensor locations are prone to be the place where the data is more complex.

A real marine seismic dataset from the Gulf of Mexico is also used to test the method. The data are composed of 180 receivers and 267 shots. The first 100 shots are the training dataset. The shot number 101, the testing shot gather, is shown in Figure 4a). As in the synthetic example, the 34 random sampled traces are shown in Figure 4b), and the reconstructed result in Figure 4c). While Figure 4d) is the optimal sensing case where we used 34 sensors. Figure 4e) is the recovered shot gather based on the optimal sampled data.

Conclusion

In this study, we adopted a technique proposed by Manohar et al. (2018) to optimal receiver design for seismic acquisition. We apply the method to both synthetic and real seismic datasets. Instead of using random measurements, sensor positions can be selected by this method. However, the drawback of the method is that it requires properly sampled shots to extract the natural bases of the

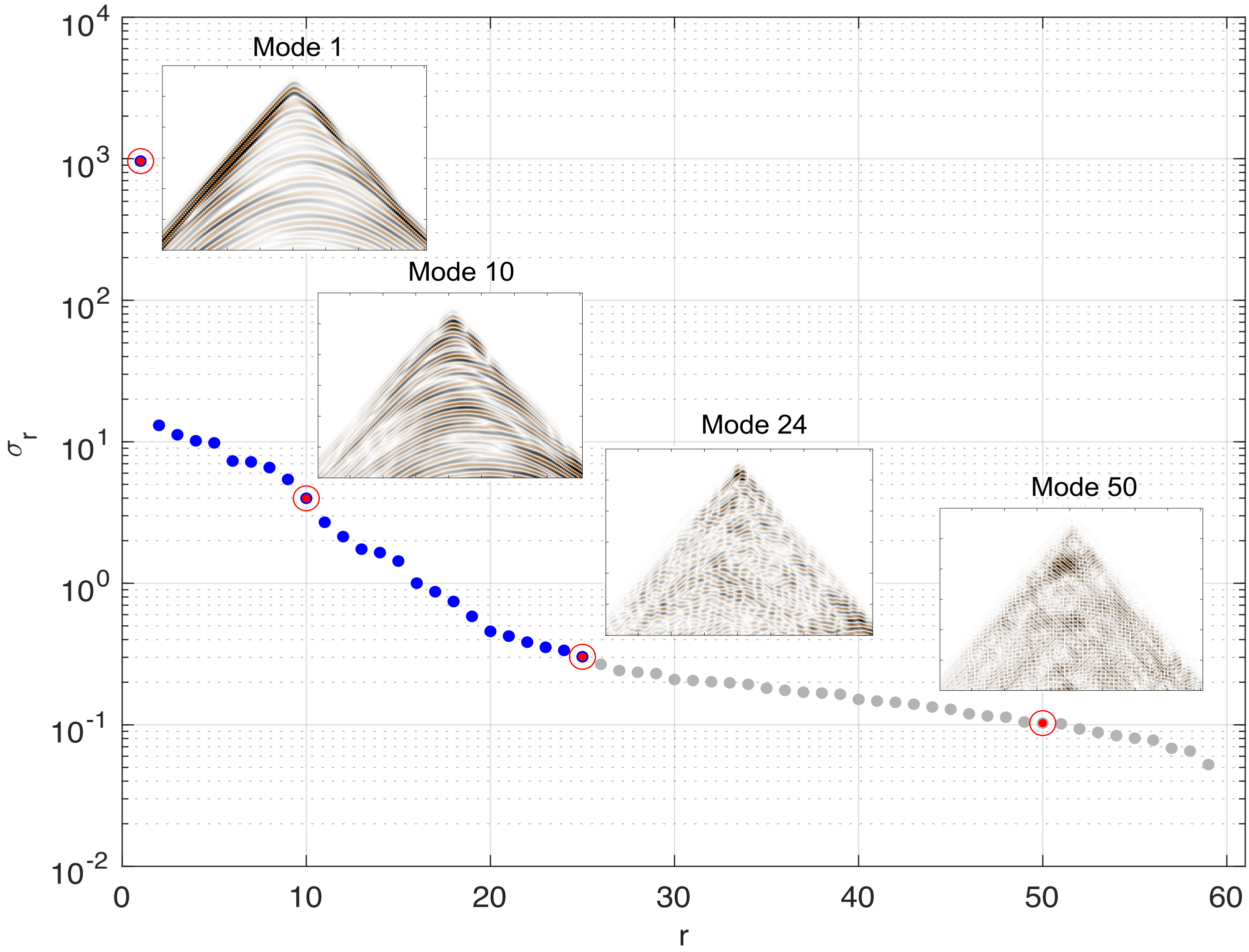

Figure 2: Singular values and four selected POD modes. The optimal rank truncation threshold occurs at r = 24.

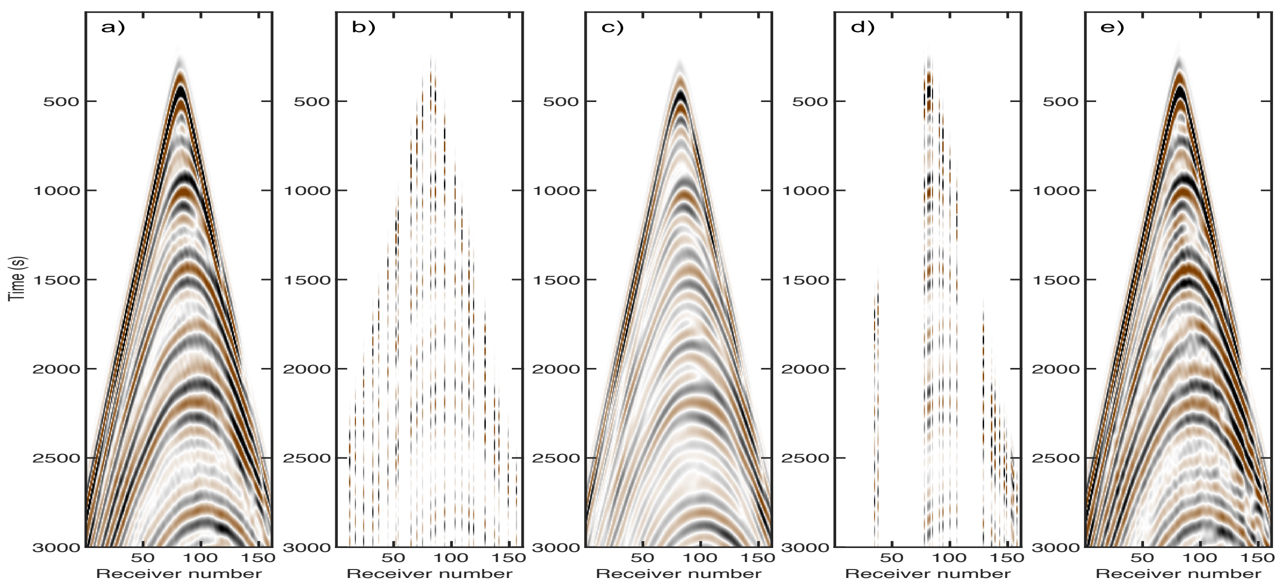

Figure 3: Comparison between the original shot gather and the reconstructed shot gather for randomly and optimally decimated data. a) The original shot gather. b) Randomly decimated data. c) Reconstruction from randomly decimated data. d) Optimally decimated data. e) Reconstruction from optimally decimated data.

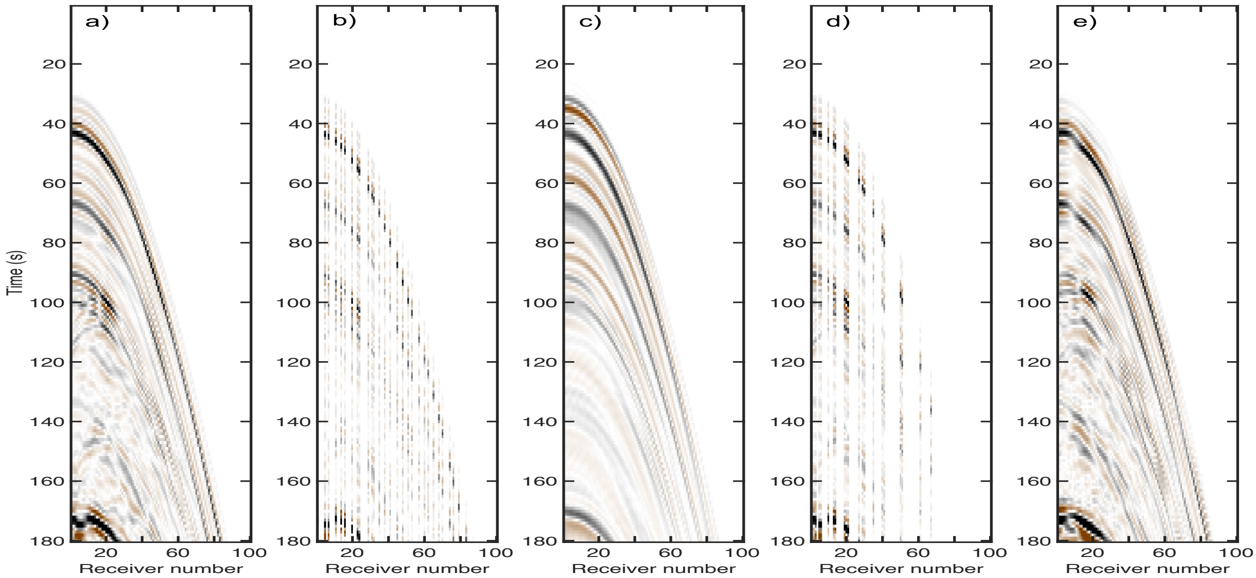

Figure 4: Comparison between original shot gather and reconstructed shot gather from randomly and optimally decimated data. a) The original shot gather. b) Randomly decimated data c) Reconstruction from randomly decimated data. d) Optimally decimated data. e) Reconstruction from optimally decimated data.

problem via POD. The results are promising. However, it is not clear to us how one can use this method to design practical algorithms for realistic seismic data reconstruction scenarios.

Acknowledgements

The authors wish to thank the sponsors of the Signal Analysis and Imaging Group at the Department of Physics of the University of Alberta. The first author also would like to thank the China Scholarship Council for their financial support.

References

Baraniuk, R., 2007, Compressive sensing [lecture notes]: IEEE Signal Processing Magazine, **24**, 118–121.
Gavish, M., and D. L. Donoho, 2013, The optimal hard threshold for singular values is 4/sqrt(3).
Hennenfent, G., and F. J. Herrmann, 2008, Simply denoise: Wavefield reconstruction via jittered undersampling: Geophysics, **73**, V19–V28.
Manohar, K., B. W. Brunton, J. N. Kutz, and S. L. Brunton, 2018, Data-driven sparse sensor placement for reconstruction: Demonstrating the benefits of exploiting known patterns: IEEE Control Systems Magazine, **38**, 63–86.
Stanton, A., and M. D. Sacchi, 2016, Efficient geophysical research in julia: CSEG GeoConvention 2016, 1–3.