DEBLENDING OF SIMULTANEOUS-SOURCE SEISMIC DATA VIA PERIODICITY-CODED NONNEGATIVE MATRIX FACTORIZATION

Tzu-Hao Lin   *Yu Tsao*

Research Center for Information Technology Innovation,
Academia Sinica

ABSTRACT

To increase the efficiency of seismic acquisition, one needs to break down the simultaneous-source seismic data into single source responses by a procedure called deblending. In this study, we employed the periodicity-coded non-negative matrix factorization (PC-NMF) to separate the primary and the secondary sources recorded in Petroleum Geo-Services dataset. Due to the application of random delay times between consecutive shots, the two sources displayed different patterns among various shots. By combining the response matrices of successive receivers, the PC-NMF can learn the basic components and group the basic components into two clusters according to the periodicity among shot index. Therefore, the deblending of simultaneous-source seismic data can be effectively achieved in an unsupervised manner.

1. INTRODUCTION

In the procedure of seismic acquisition, high-intensity sounds are produced by air guns towed behind a seismic vessel. Therefore, the geophysical structure of submarine environment can be studied by recording the energy of reflected seismic waves [1]. To increase the acquisition efficiency, the time interval between two consecutive shots needs to be reduced. Shorter shot intervals not only reduce the total amount of time of seismic acquisition but also diminish the potential noise impacts on marine animals [2]. However, the analysis of interested source may be interfered by simultaneous sources if shot intervals are less than the traveling time for all reflected energy to reach the final receiver.

In this regard, the data challenge of deblending of simultaneous-source seismic data is organized [3]. Figure 1 shows the response of 256 receivers over time for the first shot in the data provided by Petroleum Geo-Services (PGS). Interferences by simultaneous sources can be observed in the first 200 receivers. According to the response of a receiver to all shots, the secondary source was highly unpredictable among different receivers (Fig. 2). Besides, some unwanted noise was also recorded in the period of the primary source. Therefore, an effective seismic acquisition requires a blind source separation algorithm which can recognize the primary source and the other noise sources in an unsupervised manner.
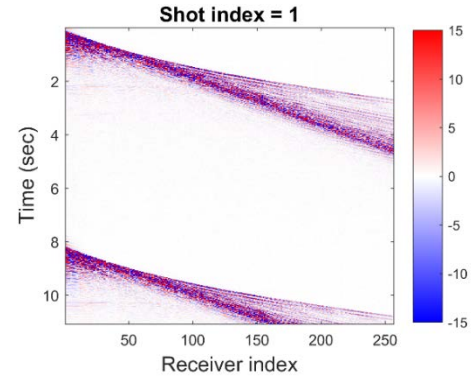

Fig. 1. The response of 256 receivers for the first shot in the PGS dataset. Note that all values higher than 15 or lower than -15 have been clipped to visualize the detail.
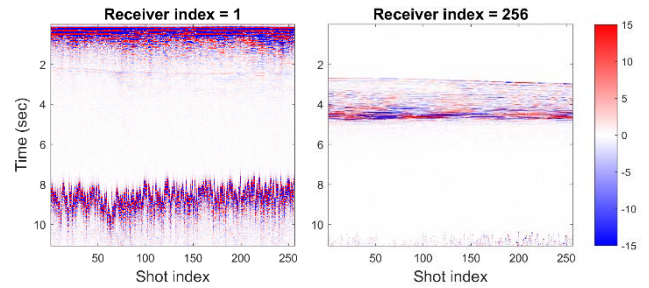

Fig. 2. The response of the 1st and 256th receivers to all shots in the PGS dataset. The secondary source is visible after 8 s in the 1st receiver, but it is only visible after 10 s in the 256th receiver.

To separate different sources, it is important to understand the behavior of each source. In Figure 2, the intensity response of the primary source gradually changes with successive shots. The gradual change of intensity response is due to the short distance traveled by the seismic vessel during the shot interval. However, the response observed in the secondary source is not following the same pattern due to the application of random delay times between consecutive shots. Therefore, a blind source separation may be applicable if the algorithm can effectively learn the structure of intensity response over time.

Non-negative matrix factorization (NMF) is a self-learning algorithm which can learn parts-based representation from an input matrix [4]. NMF has been extensively applied in many source separation techniques [5]. Lin et al. developed the periodicity-coded NMF (PC-NMF) to learn the temporal occurrence of multiple sound sources from long-duration recordings [6]. On the basis of PC-NMF, sounds with strong periodical occurrence can be effectively enhanced. In this challenge, we employ the PC-NMF in the deblending of simultaneous-source seismic data by learning the periodical occurrence of intensity response from consecutive receivers over all shots. Our results show that the PC-NMF can effectively separate the primary source and the other noise sources without any supervised training.

2. METHOD AND SYSTEM SETUP

In this study, the receiver-based response matrix was analyzed by the PC-NMF (Fig. 3). Due to the non-negative constraint of NMF, absolute values of response matrix were obtained, and their signs were kept separately to reconstruct the output matrix after source separation procedures. The PC-NMF is a two stage NMF, which conduct matrix decomposition by a traditional NMF and basis clustering by a clustering NMF [6]. At the first stage, a response matrix was decomposed into a basis matrix and an encoding matrix based on a user-defined basis number. After iterative updates, the basis matrix describes the basic components of the input data and the encoding matrix provides the intensity of each basis for each shot.

The result of matrix decomposition shows that bases associated with the primary source gradually change their encoding values among shots. However, the other bases change their encoding values in a transient pattern. Therefore, two basis clusters can be easily discriminated according to the periodicity among shots measured by discrete Fourier transform. In this study, we employ the semi-NMF [7] as the clustering algorithm, instead of the sparse NMF in [8]. Once the clustering assignment of basis has been decided, different sources can be reconstructed based on its basis cluster and the associated encoding information.

To improve the periodicity of the primary source, response matrices of two consecutive receivers were combined together. Then, we can start the system by conducting the blind source separation on the response matrices of the first and the second receivers. Then, the basis and encoding matrices were saved to initiate the next source separation on the response matrices of the second and the third receivers. This procedure can be finalized until all receivers were processed.

3. EXPERIMENTAL SETUP AND RESULT

The basis number represents the key parameter in our system. To evaluate the performance of our algorithm, we qualitatively investigated the separation performance among different basis numbers. Figure 4 shows the separation result of the primary source in the response matrix of 256 receivers over time for the first shot. The secondary source had been effectively removed by using the PC-NMF. However, it seems that the PC-NMF recognized part of the primary source as the unwanted noise if the secondary source is not very prominent.
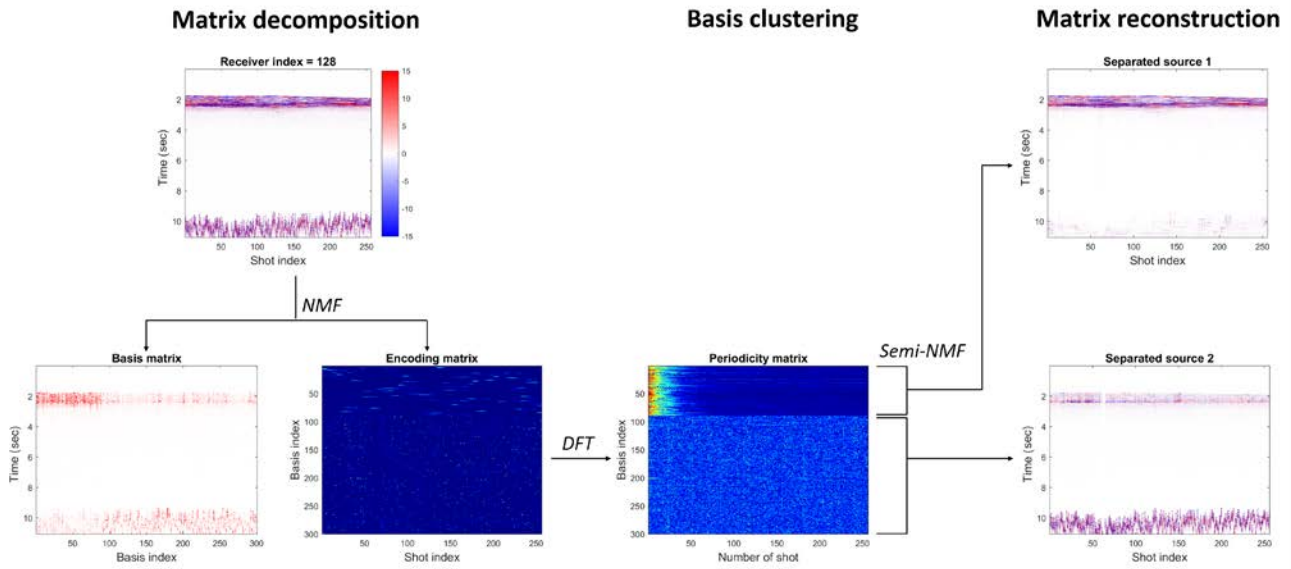

Fig. 3. Procedures of using the PC-NMF in separating the primary and the secondary sources from a response matrix of single receiver.
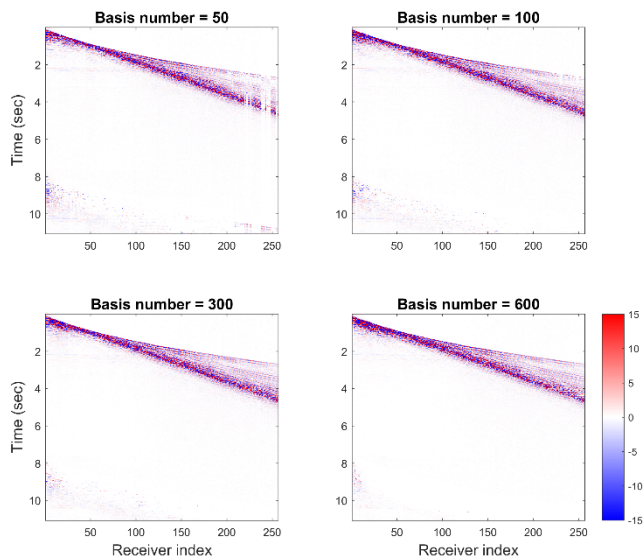
Fig. 4. Primary sources separated by using the PC-NMF with different basis numbers. Only the data for the first shot was presented. Please refer to Figure 1 for the original data.

Therefore, severe distortion of the primary source can be observed in those receivers close to the end of an array when a small basis number is chosen.

Figure 5 shows the differences between the original data and the data presented in Figure 4. A larger basis number performs better in preventing the distortion of the primary source. For a complex data, such as the response matrix of a receiver close to the air gun, multiple sources may be recorded. A small basis number may not be sufficient to learn the different structure between the primary and the secondary sources. Therefore, part of the primary source may also be removed by the PC-NMF during the reconstruction stage.

By PC-NMF, the deblending of simultaneous-source seismic data can be effectively achieved in an unsupervised manner. There is no need to provide time delay information, and the PC-NMF can automatically learn the structure of the primary and the secondary sources. The PC-NMF achieves the blind source separation due to the variation of periodicity among shot index. The transient nature of the secondary source caused by the application of random delay times represents the key point of successful source separation. Therefore, the performance of PC-NMF will not be substantially affected if the shot interval is reduced. In the future, this technique may be applied to facilitate the efficiency of seismic acquisition.
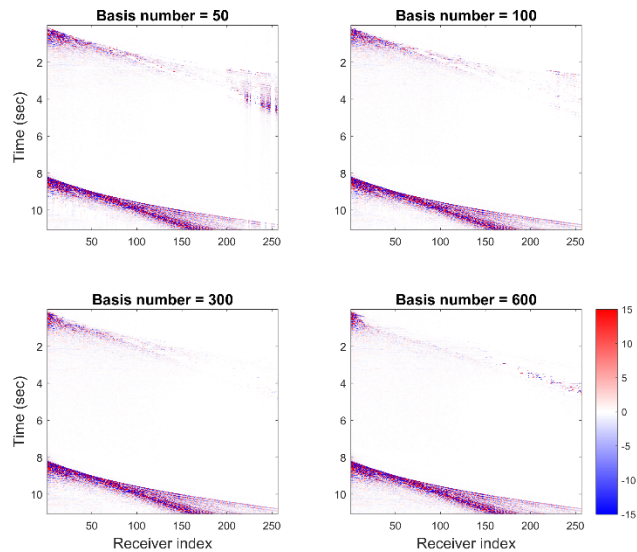

Fig. 5. Difference maps between the original data presented in Figure 1 and the separation result shown in Figure 4.

4. REFERENCES

[1]   Ö. Yilmaz, "Seismic data processing," *Society of Exploration Geophysicists*, 1987.

[2]   L. Weilgart, "A review of the impacts of seismic airgun surveys on marine life," *CBD Expert Workshop on Underwater Noise and its Impacts on Marine and Coastal Biodiversity*, pp. 1–9, 2013.

[3]   IEEE MLSP Technical Committee, "Source Separation for Simultaneous Seismic Data Acquisition," *IEEE Dataport*, 2017. [Online]. Available: http://dx.doi.org/10.21227/H2TP46.

[4]   D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, 1999.

[5]   P. Smaragdis et al., "Static and dynamic source separation using nonnegative factorizations: A unified view," *IEEE Signal Proc Mag*, vol. 31, no. 3, pp. 66–75, 2014.

[6]   T.-H. Lin et al., "Improving biodiversity assessment via unsupervised separation of biological sounds from long-duration recordings," *Sci Rep*, vol. 7, no. 4547, 2017.

[7]   C. Ding, T. Li, and M. I. Jordan, "Convex and semi-nonnegative matrix factorizations," *IEEE Trans Pattern Anal Mach Intell*, vol. 32, no. 1, pp. 45–55, 2010.

[8]   P. O. Hoyer, "Non-negative matrix factorization with sparseness constraints," *J Mach Learn Res*, vol. 5, pp. 1457–1469, 2004.