Adaptive Multiple Subtraction Using Independent Component Analysis

Wenkai Lu, Institute of information processing, Department of automation, Tsinghua University, Beijing 100084,CHINA. Email: lwkmf@mail.tsinghua.edu.cn*
Yi Luo, Saudi Aramco, Dhahran, Saudi Arabia
Bo Zhao, Zhongping Qian, CNPC, Exploration Software Co. Ltd., Beijing, 100080, CHINA.

Summary

Adaptive subtraction is a critical step for the widely-used prediction plus subtraction multiple attenuation technique. In this paper, a new adaptive subtraction algorithm based on independent component analysis (ICA), which exploits the maximum kurtosis, is presented. The method has been applied on several synthetic datasets generated by simple convolution and finite-difference model (FDM) technique, and very encourage results has been obtained.

Introduction

Surface-related multiple attenuation (SRMA) has been shown to be an effective approach to remove multiples from seismic data that are difficult to be attenuated using other (e.g., Radon transform) methods. The SRMA method generally includes two steps: multiple prediction (or multiple modeling) and adaptive subtraction. It has been observed that adaptive subtraction is the main challenge for the success of the SRMA technique.

Adaptive subtraction is posed as a least-squares minimization problem that minimizes the energy difference between the original data traces and the multiple-model traces. It can be implemented as single/multi channel matching filter. In deed, the least square solution of the subtraction scalar is based on the second order statistics. The matching filtering method requires the primaries and multiples are orthogonal, meaning there are no cross-talks between multiples and primaries (Spitz, S, 1999). Since this orthogonal assumption can not be satisfied for most (if not all) real datasets, the matching filtering approach often cause over-subtraction of multiples and harm the primaries (Wang Y., 2003).

Independent component analysis (ICA) has been recently becoming an important tool for modeling and understanding empirical datasets as it offers an elegant and practical methodology for blind signal separation (BSS) and deconvolution. Multiple subtraction can be cast as a BSS problem with two sources and two mixtures (received signals). In this paper, we suppose there are no time delay between the true multiples and the predicted multiples. In other word, we treat the BSS for multiple subtraction as an instantaneous mixing problem. The kurtosis-maximization based ICA algorithm is adopted in our multiple subtraction approach.

In the following, we will first describe the theory of ICA-based multiple subtraction. Several datasets generated by simple convolution model and the Pluto 1.5 dataset from the Smart JV are used to evaluate our method. Some conclusions and future works are given at last.

Theory

Independent Component Analysis(ICA)

Originally, Independent component analysis was developed to solve problems that are related closely to the cocktail-party problem. Since the recent increase of interest in ICA, it has become clear that this principle has a lot of other interesting applications as well. It has been realized that ICA is very closely related to the method called blind source separation (BSS). Here, "source" means original signals, i.e. independent components, embedded in the data; "blind" means we know little about the mixing parameters and impose little assumptions on the source signals. ICA is probably the most widely used method for performing blind source separation.

Assume that n linear mixtures $x_1,...,x_n$ of n-independent sources $s_1,...,s_n$ are observed

$$x_j = a_{j1}s_1 + a_{j2}s_2 + ... + a_{jn}s_n \qquad j = 1,n \qquad (1)$$

Without loss of generality, we can assume that both the mixture variables and the independent components have zero means. Using a vector-matrix form, we rewrite (1):

$$\boldsymbol{x} = \boldsymbol{As} \qquad (2)$$

where $\boldsymbol{x}$ denote the vector whose elements are the mixtures $x_1,...,x_n$, $\boldsymbol{s}$ denote the vector whose elements are the independent components $s_1,...,s_n$, and $\boldsymbol{A}$ is the matrix with coefficients a_{ij} .

The statistical model in Eq. (2) is called independent component analysis, or ICA, model. The independent components are latent variables, meaning that they cannot be directly observed. Moreover, the mixing matrix $\boldsymbol{A}$ is assumed to be unknown. All we observe is the random vector $\boldsymbol{x}$, with which we must estimate both $\boldsymbol{A}$ and $\boldsymbol{s}$. This must be done under as general assumptions as possible.

A separating matrix $\boldsymbol{W}$, which inverts the mixing, will be sought so that $\boldsymbol{s} = \boldsymbol{W}\boldsymbol{x}$. To obtain the separating matrix, we express the separated signals as:

$$\boldsymbol{y} = \boldsymbol{W}\boldsymbol{x} \tag{3}$$

Assuming the source signals are non-Gaussian and statistically independent, we optimize $\boldsymbol{W}$ by maximizing the kurtosis of $\boldsymbol{y}$, which is defined as

$$kur(\boldsymbol{y}) = E(\boldsymbol{y}^4) - 3\left(E(\boldsymbol{y}^2)\right)^2 \tag{4}$$

ICA based Multiple Subtraction

In multiple subtraction problem, the main objective is to reconstruct the primary, $\boldsymbol{P}$, which is distorted by the multiples $\boldsymbol{M}$, from the seismic data $\boldsymbol{X}$ and predicted multiples $\boldsymbol{M}_p$. The relationship of these vectors can be written as:

$$\boldsymbol{X} = \boldsymbol{P} + \boldsymbol{M} \tag{5}$$

$$\boldsymbol{M}_p = \alpha \boldsymbol{M} \tag{6}$$

where α is a unknown scalar, note here that we simply assume the difference of predicted and true multiples is a scalar factor. The wavelet difference between the multiples in seismic data and the predicted multiples is ignored because it may be removed by other techniques. For example, the traditional multi-channel matching-filtering techniques are good candidates to remove the wavelet difference, this will be illustrated in next section in figure 2.

To estimate the subtraction scalar β, the traditional methods minimize the following cost function

$$E = \left\| \boldsymbol{X} - \beta \boldsymbol{M}_p \right\|_2 \tag{7}$$

Its least squares solution is

$$\beta = \frac{\boldsymbol{M}_p^{\boldsymbol{T}} \boldsymbol{X}}{\boldsymbol{M}_p^{\boldsymbol{T}} \boldsymbol{M}_p} \tag{8}$$

It can be approved that $\beta = \frac{1}{\alpha}$ only when the primary and multiple are orthogonal.

In stead of assuming primary and multiple are orthogonal (which is rarely valid in real datasets), we assume that the seismic signals (primaries and multiples) are non-Gaussian distributed. It is known that a mixture of two non-Gaussian signals tends to be Gaussian distributed. Therefore, our goal to optimize β is to make the separated primaries as far as possible away from Gaussian density. In our experiments, we exploit the zero-lag fourth order moment, kurtosis, as a measure of non-Gaussianity. The proposed ICA based method maximize the following cost function

$$E = kur(\boldsymbol{X} - \beta \boldsymbol{M}_p) \tag{9}$$

Example

To demonstrate the performance of the proposed method, we apply our method in several synthetic data generated by simple convolution and finite-difference model technique.

Fig. 1a shows a seismic data generated by convolution technique with three primary events and three multiple events. Fig. 1b is the predicted multiple, which is the same as the multiple events in the seismic data shown in Fig. 1a except for a scaling factor. Note that one primary event and a multiple event around 210ms (indicated by arrow) are almost overlapped. In our experiments, we use all the traces to obtain the subtraction scalar. Fig. 1c shows the demultiple result by method based on Eq. (8). Fig. 1d shows the demultiple result by our ICA based method. It is seen that our method can completely recover the primaries even if some primaries and multiples are seriously overlapped.

Applying these methods to the Pluto 1.5 dataset should produce a more realistic comparison. The dataset is a 2-D elastic finite-difference synthetic that is significantly more complicated than the previous examples. Fig 2a shows the common-offset data section, which contains significant multiple energy. The calculated multiples are seen in Fig 2b. Fig. 2c shows the demultiple result by method based on Eq. (8). Fig. 2d shows the demultiple result by our ICA-based method. In this example, the input for our method is the output after matching filtering based on Eq. (8), which is indeed the difference between figures 2a and 2c. The arrows in Fig. 2c and 2d indicate the residual multiples. It is clear that our method offers better demultiple result.

Conclusions

We proposed a new ICA-based method for multiple subtraction using higher order moment. We use kurtosis based cost function to calculate the subtraction scalar. The results obtained with synthetic data demonstrate the viability of the new method. The future work will focus on the multiple subtraction with time delay and apply the method to real datasets.

References

Ikelle, L. T., Roberts, G. and Weglein, A. B., 1997, Source signature estimation based on the removal of first-order multiples: GEOPHYSICS, Soc. of Expl. Geophys., **62,** 1904-1920.

Mendel, J. M., 1991, Tutorial on higher-order statistics in signal processing and system theory: Theoretical results and some applications, Proc. IEEE, **79**, 278–305.

Roberts, S., and Everson, R., 2001, Independent component analysis: principle and practice, Cambridge University Press.

Wang, Y., 2003, Multiple subtraction using an expanded multichannel matching filter, Geophysics, **68**, 1, 346–354..

Yoo, S. and Ikelle, L., 2001, An application of inverse Kirchhoff scattering multiple attenuation method to Pluto 1.5 dataset, 71st Ann. Internat. Mtg: Soc. of Expl. Geophys., 1289-1292.

Spitz, S., 1999, Pattern recognition, spatial predictivity, and subtraction of multiple events, The Leading Edge, No 1, 55-59

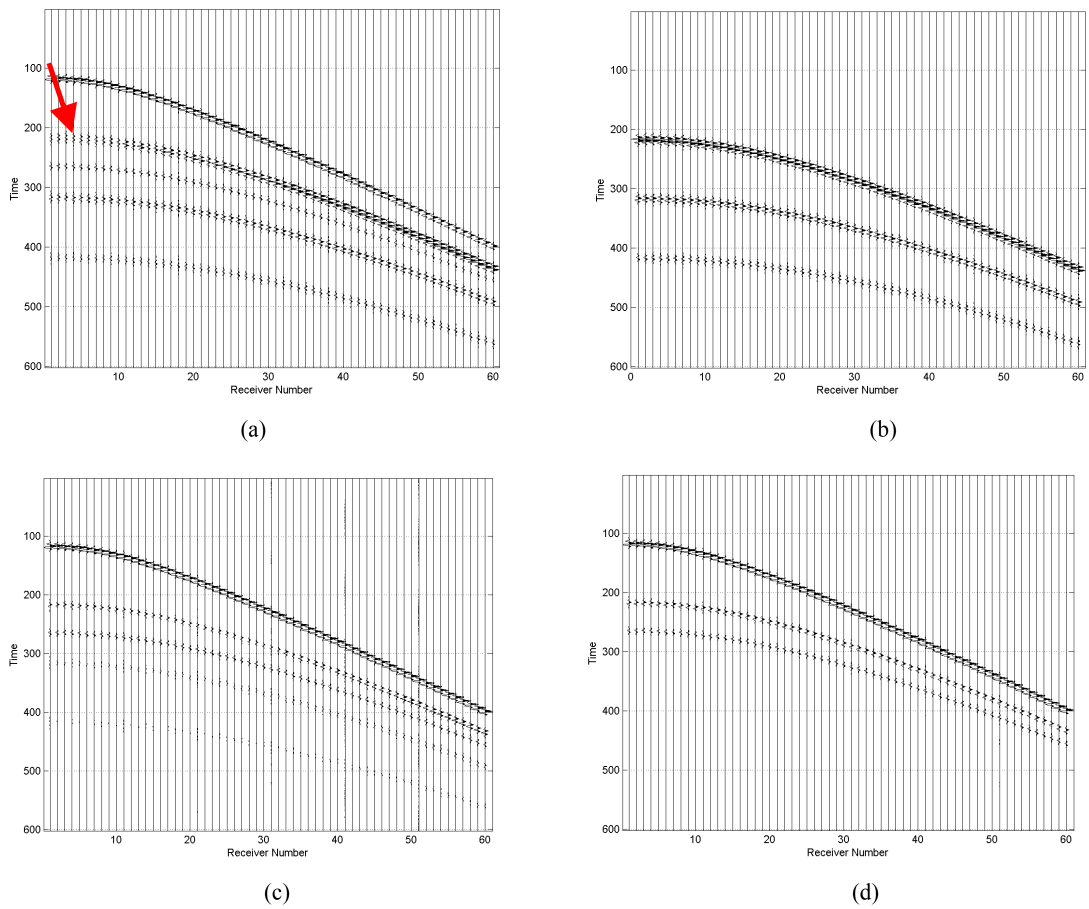

Fig. 1. A simple dataset generated by convolution model for demultiple, (a) the original seismic data with three primary events and three multiple events. Note that one primary event and a multiple event around 210ms (indicated by arrow) are almost overlapped; (b) the predicted multiple, which is the same as the multiple events in the seismic data shown in Fig. 1a except for a scaling factor; (c) the demultiple result by least squares based method; (d) the demultiple result by our ICA based method.

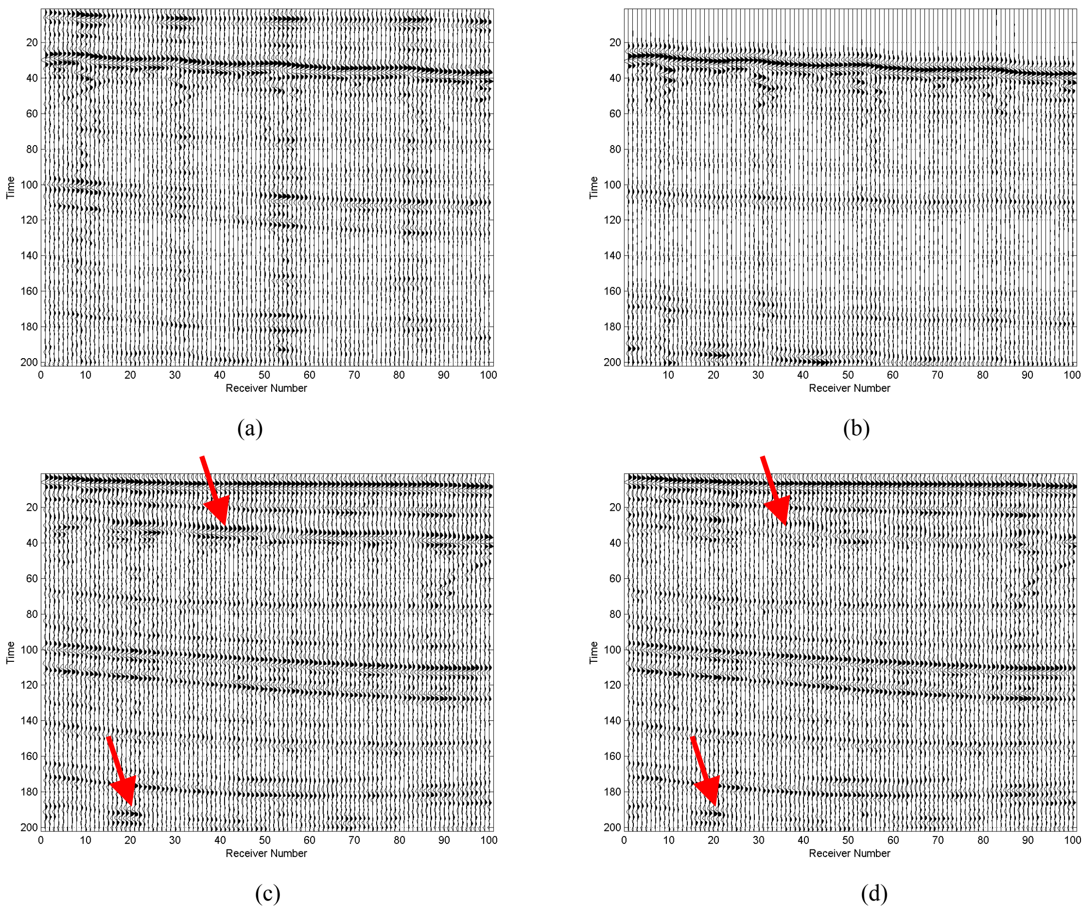

Fig. 2. The SMAART Pluto 1.5 demultiple example, (a) common-offset section; (b) the predicted multiple (c) the demultiple result by least squares based method; (d) the demultiple result by our new ICA based method.