The Study of Blind Source Separation Based on Sparsity and Decorrelation

Jiahui Li
School of Business Management, Shanghai University, Shangda Road 99, China
1055078250@qq.com

Abstract. The main research of blind source separation algorithm is based on sparsity and decorrelation. The traditional sparse algorithm is used to estimate the mixed matrix by sparse signal then the source signal is reconstructed by optimizing function in time-frequency domain. But for some sparse signals, the separation effect is not ideal. The alternate decomposition algorithm can be used as the optimization method for the signal that is less sparse. Alternate decomposition approach is to update the mixed matrix in each iteration and estimate the source signals so that the parameters can be updated in each iteration, the sparse components can be obtained, and then the source signal is reconstructed. If a decorrelation step has been joined in every iteration, it can reduce the correlation between signals effectively, and get sparse and independent of composition in the separation process. It can greatly improve separation performance by this way.

1. Introduction

1.1. Basic Model and Optimization Algorithm
Blind source separation algorithm based on sparsity is the basic step to reconstitute the reconstruction error between the real source signal and the estimated source signal, mathematically, and constantly updated iterations, finally get the minimum error, then estimate the source signals.

The instantaneous linear mixture model of blind source separation:

\[ x = As + e \]  \hspace{1cm} (1)

1.2. Notations
\( X \in \mathbb{R}^{M \times T} \) are the matrices of mixture channels, and \( s \in \mathbb{R}^{N \times T} \) are the matrices of source signals. \( A \in \mathbb{R}^{M \times N} \) is the mixing matrix and \( e \in \mathbb{R}^{M \times T} \) is the background noise.

Represent by \( \Phi \in \mathbb{C}^{T \times B} \) the matrix STFT (Short-Time Fourier Transform), \( s \) can be resynthesized from their synthesis coefficients \( a \in \mathbb{C}^{T \times B} \) by:

\[ s = a\Phi^* \] \hspace{1cm} (2)
Φ* is the adjoint operator of Φ, which also called Hermitian transpose. The model based on the sparse blind source separation problem can be expressed as an optimization problem, and its mathematical expression is as follows:

\[
\min_{\mathbf{A}, \mathbf{a}} f(\mathbf{A}, \mathbf{a}) + \Psi(\mathbf{a}) + g(\mathbf{A})
\]  

This model is the norm of \( \ell_1 \). The model can be divided into three parts: For the first part, \( f \) is the loss between the mixture \( x \) and the synthesis coefficients \( \alpha \) of sources by the mixing matrix \( \mathbf{A} \). The optimization of the cost function is to find the global error minimum in the local error minimum to isolate the source signal. For the second part, \( \Psi \) models the sparsity assumption of the sources coefficients \( \alpha \). That is, you can represent all the source signals with a linear combination of fewer eigenvectors. For the third part, \( g(\mathbf{A}) \) is mainly used to avoid the separation ambiguity [1] such as the scaling and permutation problem. Permutation problem refers to the different order of signal and the source signal after the separation, and the scaling problem refers to the different amplitude size after separation. The reasons for the uncertainty are the lack of prior information and the unknown \( S \) and \( A \). The uncertainty of the scaling and permutation of source signals can be reduced in the process of \( g(\mathbf{A}) \). If the norm \( \ell_2 \) is used, the model becomes:

\[
\min_{\mathbf{A}, \mathbf{a}} L(\mathbf{A}, \mathbf{a}) = \frac{1}{2} \| \mathbf{x} - \mathbf{A} \mathbf{a} \Phi^\dagger \|_2^2 + \lambda \| \mathbf{a} \|_1 + t_B(\mathbf{A})
\]  

\( t_B(\mathbf{A}) \) is a convex set of closed indicator functions which is:

\[
t_B(\mathbf{A}) = \begin{cases} 0 & \text{if } \| a_n \| < 1 \\ +\infty & \text{otherwise} \end{cases} \quad \forall n
\]  

\( a_n \) is the \( n \)-th column of \( \mathbf{A} \).

This functional normally appears in BSS. Also, a generalized morphological component analysis algorithm (Generalized Morphological Component Analysis, GMCA) is proposed for image processing [2].

2. The Related Generalized Morphological Component Analysis Algorithm

The correlation algorithm mainly uses the correlation of signals, which is a fast convergence adaptive algorithm. This algorithm has the advantages of fast convergence and small computation. This paper studies a kind of relevant steps to implement each iteration algorithm to ensure the independence of the source, this will make decorrelation be applied in other situations.

In practical applications, the correlation of random variables is difficult to express directly. Therefore, the covariance matrix of random vectors is used to represent the correlation between random variables. It is known from the knowledge of linear algebra that the covariance matrix is semi-positive and symmetric. When the random variables are not related, the covariance matrix is the diagonal matrix. So in general, if the covariance matrix formed by a random variable is not a diagonal matrix, then there is a correlation between the random variables which means the redundancy between the random variables is higher. If you can use some methods to correlate random variables, you can make the covariance matrix into a diagonal matrix.

First, the method of Minimized mean-squared Error (MMSE) [3] is introduced:
Assume \( s \in \mathbb{R}^{N \times T} \) is a positive definite covariance matrix. \( \Sigma_s = ss^T \) is a zero-mean signal matrix. \( W \) is the optimal decorrelation transformation. In a diagonal matrix, \( W \) can minimize the Mean-Squared Error between \( S \) and \( y = Ws \) with covariance \( \Sigma_y \):

\[
\min_{y=Ws} \| s - y \|_2^2
\]

(6)

In \( W = \text{diag}\left( \Sigma_s^{1/2} \right) \Sigma_s^{-1/2} \), \( \text{diag} \) (.) is the diagonal matrix which is consisted of its diagonal elements. So (6) can be decomposed in the following way:

\[
\min_{\hat{W}} \| s - \hat{W}y \|_2^2
\]

\[
\hat{W} \hat{W}^T = D
\]

(7)

\( y^* \) is the whitened signal. \( \bar{y} = \sum_s^{1/2} s \) And \( D \) is any diagonal matrix. Then the Mean-Squared Error can be written as:

\[
e = \| s \|_2^2 + \| \hat{W}y \|_2^2 - 2 \langle s, \hat{W}y \rangle
\]

(8)

According to Cauchy-Schwarz inequality, \( \langle s, \hat{W}y \rangle \leq \| s \|_2 \| \hat{W}y \|_2 \) will be equal if and only if \( \hat{W} \) is a diagonal matrix. Assume \( \hat{d}_i \) to be the i-th element on the diagonal of \( \hat{W} \), then \( e \) can be shown as:

\[
e = \sum_{i=1}^{N} (s_i - \hat{d}_i \bar{y}_i)(s_i - \hat{d}_i \bar{y}_i)^T
\]

(9)

\( s_i \) is the i-th row of \( s \) and \( \bar{y}_i \) is the i-th row of \( \bar{y} \). Now it is easy to minimize \( e \) with the condition of \( \hat{d}_i = s_i \bar{y}_i^T \) which is the same as \( \hat{W} = \text{diag}\left( s \bar{y}^T \right) \). 

---

3
We can obtain the decorrelated-GMCA by adding such a projection in the GMCA algorithm:

**Decorrelated-GMCA**

Initialization: \( a_1 \in \mathbb{C}^{N \times B}, A_1 \in \mathbb{R}^{M \times N}, L_k = \|A_k\|, k = 1 \)

repeat

1. \( a_{k+1} = \operatorname{prox} \left( \alpha - \frac{\nabla f(A, a_k)}{L_k} \right) \);
2. \( s_k = a_{k+1} \Phi^* \);
3. \( \Sigma = s s^T \);
4. \( W = \operatorname{diag} \left( \Sigma_{x}^{1/2} \right) \Sigma_{z}^{-1/2} \);
5. \( y = W s \);
6. \( A_{k+1/2} = y y^T \);
7. \( a_{k+1} = a_{k+1/2} / \|a_{k+1/2}\| \);
8. \( L_{k+1} = \|A_{k+1}\|, k = k + 1 \);

until convergence

Above all, minimum mean square error decorrelation method in every iteration (steps 2 to 5) is to ensure that the signals are decorrelated and can keep conservation. Furthermore, these extra steps do not add to the complexity of the operation. However, the convergence of the algorithm does not prove a good derivation, which is still a problem to be solved.

3. Numerical Experiments

In the simulation experiment, the linear instantaneous hybrid blind source separation model is considered. The related generalized morphological component analysis algorithm is used in the simulation experiment. For the number of the hybrid system, we choose three-way source signals and two-way observation signals. The signal sampling points are 50000 points. Linear instantaneous mixing is performed in the time domain. Let two-way observation signals which are mixed by the three-way source signals transform from time domain to frequency domain by STFT (Short-Time Fourier Transform). The frames length of STFT is 1024 points and the frames overlap covers 512 points. This experiment will evaluate the effectiveness of this algorithm through the separation performance of different signals.

3.1. Separation of Audio Signals

In this group of experiments, three different audio signals are selected from the phonetic library as the source signal. Figure 1 (a) is the three-way audio signals time domain waveform diagram. Figure 1 (b) is the time domain waveform diagram of the two-way observation signals obtained after the linear instantaneous mixing of the three-way source signals in the time domain. Figure 1 (c) is the time domain waveform diagram of three-way recover signals by the method in this paper.

In order to evaluate the performance of blind source separation algorithm, the experiment adopted widely used in Signal performance evaluation standard of the distortion of the Signal and noise ratio (Signal to Distortion Ratio, SDR), Signal distortion and interference ratio (Signal to Interference Ratio, SIR).
SIR) and the Signal with fake Signal distortion ratio (Signal to Artifact Ratio, SAR) these three indicators as the separation of Signal blind source separation algorithm performance evaluation standard [4]. The Signal to Distortion Ratio (SDR), the Signal to Interference Ratio (SIR) and the Signal to Artifact Ratio of the signals are respectively shown in table 1, table 2 and table 3 respectively.

![Signal Separation Results](image)

(a) Three-Way Source Signals  (b) Two-Way Observation Signals  (c) Three-Way Recover Signals

**Figure 1.** Blind Source Separation Results of Three-Way Source Signals

**Table 1.** SDR of Three-Way Source Signals

| Output Signal | $\tilde{s}_1(n)$ | $\tilde{s}_2(n)$ | $\tilde{s}_3(n)$ | Average SDR of output signal |
|---------------|-----------------|-----------------|-----------------|-----------------------------|
| SDR (dB)      | 10.3299         | 10.0041         | 13.6229         | 11.3190                     |

**Table 2.** SIR of Three-Way Source Signals

| Output Signal | $\tilde{s}_1(n)$ | $\tilde{s}_2(n)$ | $\tilde{s}_3(n)$ | Average SIR of output signal |
|---------------|-----------------|-----------------|-----------------|-----------------------------|
| SIR (dB)      | 13.4877         | 18.8570         | 16.8227         | 16.3891                     |

**Table 3.** SAR of Three-Way Source Signals

| Output Signal | $\tilde{s}_1(n)$ | $\tilde{s}_2(n)$ | $\tilde{s}_3(n)$ | Average SAR of output signal |
|---------------|-----------------|-----------------|-----------------|-----------------------------|
| SAR (dB)      | 13.3879         | 10.6662         | 16.5409         | 13.5317                     |

4. Conclusion

In this experiment, different types of three-way audio signals were used to test the method. Separation results as shown in figure 1. From figure 1 (a) and (c), it can be seen that the recovery performance of the research method is basically satisfactory, and the different speech signals are separated. By comparing the average value of SDRS in the evaluation criteria of separation signals, the average value of the SDR is 11.3190 dB, and the average of SIR is 16.3891 dB, and the SAR average is 13.5317 dB. The performance of this method is good for separating the audio signals. The method presented in this study has excellent performance relative to other methods: it reduces the requirement for the sparse signal of source signals, and solves the problem of blind separation of noise signals. Although the separation process has been added to the relevant steps, it does not increase the computational complexity of the algorithm.
References

[1] Boumaraf H, Separation aveugle de mélanges convolutifs de sources, Ph.D. thesis, Université Joseph-Fourier-Grenoble I, 2005.

[2] Bobin J, Starck J L, Fadili L, and Moudden Y, “Sparsity and morphological diversity in blind source separation,” Image Processing, IEEE Transactions on, vol.16, no. 11, pp. 2662 – 2674, 2007.

[3] Zhao Xiaowei. Image restoration algorithm based on sparse constraint regularization [D]. Hebei: Yanshan university school of information science and engineering, 2014.

[4] Song Jifei. Research on the separation algorithm of undetermined blind source under noise conditions [D]. Dalian: Dalian university of science and technology, xintong college, 2015.