Noisy blind source separation based on CEEMD and Savitzky-Golay filter

Hua-Fu Peng and Gao-Ming Huang
College of Electronic Engineering, Naval University of Engineering, Wuhan, China

Abstract. The standard independent component analysis (ICA) algorithm is difficult to extract signals in noise condition, a blind separation algorithm based on denoising pretreatment was proposed. Mixed signals firstly were decomposed into several stationary intrinsic mode components (IMF) using complementary ensemble empirical mode decomposition (CEEMD), and high frequency IMF components were filtered with Savitzky-Golay filtering, then using the whole components reconstructed the mixed signals, finally applying the fast independent component analysis(FastICA) to separate the reconstructed signals. Simulation results showed that the proposed method improved the effect of blind signal separation under low signal-to-noise ratio.

1 Introduction
Blind source separation refers to the process of recovering source signals based on only observed mixed signals in the case of unknown source signals and their mixing methods, which are widely used in the fields of speech processing, image processing, signal reception and communication jamming. As an important method of blind source separation, ICA has good separation effect under the linear mixing condition of the signal.

Under the actual conditions, the nonlinear mixing intensity of the signal is increased due to the influence of various noises, and the performance of the linear hybrid blind separation algorithm is weakened. The blind separation of noisy mixed speech signals is realized by approximating maximum likelihood estimation method in [1], and the joint maximum likelihood estimation algorithm based on mixed matrix and independent component analysis is proposed in [2] [3]. Because of the influence of noise, the non-linear relationship between the observational data and the source signal is difficult to direct the blind source separation. The paper [4] first utilizes the non-sampling wavelet transform (UWT) to denoise the mixed image, then the FastICA processing is obtained, and the separation effect is better. The signal denoising pretreatment based on particle filter is applied to blind source separation in [5], and the mixed signal separation performance is enhanced, but the computational quantity is large. For the noisy mixed signals, the pretreatment of denoising can effectively improve the subsequent blind source separation.

2 Blind separation algorithm based on denoising preprocessing

2.1 Denoising based on CEEMD
As a time-frequency domain analysis method, EMD can decompose the signal $x(k)$ into $M$ linear steady state intrinsic modal components [6]
The denoising function is realized by selecting the appropriate IMF component reconfiguration signal, and it has an adaptive advantage compared with the wavelet threshold Denoising method. However, EMD decomposition is unstable, there are modal aliasing phenomena, i.e. the signals with similar scales are decomposed into different IMF components.

In order to overcome the modal aliasing phenomena in EMD decomposition, Wu proposed Ensemble Empirical mode Decomposition (EEMD) in [7], using the white noise power spectral density evenly distributed characteristics, adding a certain white noise to the original signal, so that the signals on different scales were continuous, and denoising by summation. In practical application, in order to achieve better anti-noise performance, the number of superimposed noises is increased and the computational quantity is improved. For this, Yeh in [8] proposed the Complementary Ensemble Empirical mode Decomposition (CEEMD) [9], by limiting the addition of auxiliary noise to positive and negative pairs, eliminating the effects of superimposed noise, with high computational efficiency. The calculation flow is as follows:

1. Generates L pairs of positive and negative white noises with the signal \( x(k) \), respectively, superimposed with the original signal, obtaining 2L noise signals \( x_i(k) \) 
   \[ x_i(k) = x(k) + N_i(k) \quad i = 1, 2, \cdots, 2L \]  

Where \( N_i(k) \) is the white noise.

2. Respectively, each EMD decomposition is obtained, and the corresponding IMF component is \( I_{i1}(k), I_{i2}(k), \cdots, I_{iM}(k) \).

3. The IMF component, which sums the corresponding IMF components on average, after CEEMD decomposition
   \[ I_i(k) = \sum_{j=1}^{M} I_{ij}(k) \]  

2.2 Savitzky-Golay filter

Savitzky-Golay filter which is proposed by Savitzky A and Golay m in 1964, is a moving window average weighted algorithm [10], with simple, fast advantages [11], compared with other similar average methods, it retains the maximum value, minimum and width distribution characteristics better.

Take \( M \) sample points data before and after \( x_i \) in the original data, and \( x_i \) is set as the origin, that is to construct a \( 2M+1 \) sampling points of the window array, constructing a \( P \)-order polynomial
   \[ q(n) = \sum_{i=0}^{P} a_n^i \]  
fitting this array, where \(-M \leq n \leq M\), \( P \geq 2M + 1 \),
   \[ C = \sum_{n=-M}^{M} \left[q(n) - x(n)\right] \]  
   \[ = \sum_{n=-M}^{M} \sum_{i=0}^{P} a_n^i - x(n) \]  

When the minimum value is obtained, the fitting effect is best, and the smoothing filter is realized by fitting the original data of all sample points.

2.3 Fast Independent component Analysis

ICA is an important method of blind source separation, which can be described as follows [12] :
   \[ X = AS \]  

\[ x(k) = \sum_{i=1}^{M} I_i(k) \]  

(1)
Where \( \mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_m]^T \) is a m-dimensional observational signal matrices, \( \mathbf{x}_i = [x_{i1}, x_{i2}, \cdots, x_{im}] \) is the observational signal, \( \mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_n]^T \) is the n-dimensional independent source signal component matrices, \( \mathbf{s}_j = [s_{j1}, s_{j2}, \cdots, s_{jn}] \) is the source signal, \( \mathbf{A} \) is a mixed matrices. The main task of blind source separation is to solve the mixed matrix \( \mathbf{A} \) and the source signal \( \mathbf{S} \) based on the observation matrix \( \mathbf{X} \).

Fast independent component Analysis (FastICA) is an iterative optimization algorithm for independent component analysis by Hyvarinen [13], with negative entropy as the objective function to measure the independence of each component, Newton iterative algorithm is used to batch the sampling points of the observational signals, which has the advantages of robust and fast convergence. Negative entropy is defined as:

\[
J(\mathbf{w}, \mathbf{x}) = \left\{ E \left[ \mathbf{G}(\mathbf{w}, \mathbf{x}) \right] - E \left[ \mathbf{G}(\mathbf{v}_j) \right] \right\}^2
\]

Where \( \mathbf{w}_i \) is the \( i \) row of the separating matrix \( \mathbf{W} \), \( \mathbf{v}_j \) is a Gaussian random signal with zero mean and unit variance, \( \mathbf{G}(\cdot) \) is a nonlinear function. The delivery of \( \mathbf{W} \) can be obtained under orthogonal constraint conditions:

\[
\mathbf{w}_{i+1} = \mathbf{E} \left[ \mathbf{xg} \left( \mathbf{w}_i^T \mathbf{x} \right) \right] - \mathbf{E} \left[ \mathbf{g}^I \left( \mathbf{w}_i^T \mathbf{x} \right) \right] \mathbf{w}_i
\]

### 3 Blind signal separation algorithm based on denoising preprocessing

Because of the influence of noise, the nonlinear characteristic of mixed signal is enhanced, thus decreasing the effect of signal separation. Therefore, the pretreatment of mixed signal denoising can effectively improve the blind separation effect of subsequent signals. The paper [6] proposes a denoising method based on EEMD and Savitzky-Golay filtering, which has better ability to denoise, but EEMD decomposition is influenced by superposition noise. Based on the improvement of the denoising algorithm based on EEMD, this paper combines the preprocessing of the SG filter to complete the signal denoising, reduces the effect of the noise on the mixed signal, and improves the blind separation effect of the signal. The specific process of the algorithm is shown in Figure 1.

Firstly, the mixed signal is decomposed into multiple IMF components by using CEEMD, then the high frequency component is processed by the SG filter; Secondly, the signal denoising is realized by using the filtering component signal and residual component signal to reconstruct the mixed signal. Finally, the FastICA algorithm is used to complete the blind separation of mixed signals.

Since multiple mixed signals are separated by blind source, the order of the signals is disturbed, the traditional evaluation indices based on mean square error (MSE) cannot be applied directly, need to sort the restored signals. Aiming at this problem, this paper defines the evaluation index of separating effect

\[
E_s = \frac{1}{MT} \sum_{i=1}^{M} \sum_{j=1}^{T} \left( y_{ij}^2 - s_{ij}^2 \right)
\]

Where \( y_{ij} \) is the signal after separation, \( s_{ij} \) is the source signal. The evaluation index does not need to reorder the separated signals, the smaller the \( E_s \) value, the closer the signal is to the original signal.
4 Simulation

4.1 Analysis of EMD-SG denoising performance

The emulation parameters are set as follows: \( f_s = 1 \text{GHz} \) is the sampling rate. 
\( s(k) = \sin(2\pi f_1/k) \sin(2\pi f_2/k) \) is the source signal, where \( f_1 = 2 \text{MHz} \), \( f_2 = 200 \text{MHz} \). 
\( T = 1000 \) is the signal length, \( \sigma = 0.5 \) is the noise standard deviation, \( 2L = 20 \) is the number of noises.

Separately, EEMD and CEEMD have been used to decompose the signal into multiple IMF components \( I_i(k) \), and then add all the IMF components to obtain the restored signal \( \tilde{s}(k) = \sum_{i=1}^{2L} I_i(k) \), the following illustration shows the signal \( d(k) = s(k) - \tilde{s}(k) \) which is the difference signal of the original signal and the restored signal. It can be seen that because of the influence of superposition noise, and the superposition number is limited, the signal can not be restored after EEMD decomposition, because the auxiliary noise is positive and negative pairs in CEEMD, so that the signal can be completely restored using CEEMD.

Figure 1. Algorithm Flow
To verify the outperformance of the proposed approach, we compare the filter results of two methods which are EEMD-SG in [6] and CEEMD-SG proposed in this paper. The results are taken average over 100 Monte Carlo simulations with SNR ranging from 0-25dB. $E = \sum_{i=1}^{r} (\hat{x}(i) - x(i))^2$ is used to evaluate separation performance.
4.2 Algorithm performance analysis

Considering the following original signals

\[ s_1(t) = \sin(2\pi f_c t), \text{where } f_c = 20\text{MHz}. \]
\[ s_2(t) = \sin(2\pi f_1 t)\sin(2\pi f_2 t), \text{where } f_1 = 2\text{MHz}, f_2 = 200\text{MHz}. \]
\[ s_3(t) \] is a linear FM signal with start frequency is 50MHz and bandwith is 100MHz.

Mixed matrices

\[
A = \begin{bmatrix}
0.1589 & 0.517 & 0.3654 \\
0.3206 & 0.7965 & 0.0457 \\
0.8656 & 0.2555 & 0.9452
\end{bmatrix}
\]

is randomly generated.

In order to validate the performance of the algorithm, the separation performance of the algorithm and the standard FastICA algorithm is compared. 100 times Monte Carlo experiments, the result is shown in the figure 4. It can be seen that the algorithm is better than the standard FastICA algorithm.

5 Conclusion

Because the noise can enhance the nonlinearity of the signal mixing, thus weakening the separation effect of the traditional ICA, a blind separation algorithm based on denoising preprocessing is presented in this paper. By adopting CEEMD eliminating the influence of superposition noise on signal EEMD decomposition, and decreasing the computational quantity, the effect of the algorithm based on EEMD...
denoising is improved, the noise effect is suppressed by the preprocessing of signal denoising using CEEMD and SG filter, and the blind separation effect of mixed signal is improved. In addition, a blind separation effect evaluation index is given, which does not need to reorder the separated signals.

References

[1] Moulines E, Cardoso J F, Gassiat E. (1998). Maximum likelihood for blind separation and deconvolution of noisy signals using mixture models. *in Proc of ICASSP, Washington*, 3617-3620.

[2] Hyvärinen A. (1998). Noisy independent component analysis, maximum likelihood estimate, and competitive learning. *in proc of IJCNN, Alaska*, 2282-2286.

[3] Hyvärinen A. (1999). Gaussian moments for noisy independent component analysis. *IEEE Signal Processing Lett*, 145-147.

[4] CAI Weihua, HE Xuansen. (2016). Noisy blind source separation based on undecimated wavelet transform and independent component analysis. *Computer Engineering and Applications*, 52(16), 180-185.

[5] Qiu Hao, Huang Gaoming & Gao Jun. (2014). RAO-BLACKWELLIZED PARTICLE FORWARD FILTERING BACKWARD SMOOTHERING WITH APPLICATION TO BLIND SOURCE SEPARATION. *ICSP*.

[6] XINF Hong-yan, ZHU Qing-qing. (2016). The Sea Culter De-noising Based on Ensemble Empirical Mode Decomposition. *ACTA ELECTRONICA SINICA*, 44, 1-7.

[7] Wu Z H, Huang N E. (2009). Ensemble empirical mode decomposition: a noise assisted date analysis method. *Advances in Adaptive Data Analysis*, 1, 1-41.

[8] Yeh J R, Shieh J S, Huang N E. (2010). Complementary ensemble empirical mode decomposition: a novel noise enhanced data analysis method. *Advances in Adaptive Data Analysis*, 2, 135-156.

[9] Colominas MA, Schlotthauer G, Torres ME. (2014). Improve complete ensemble EMD: A suitable tool for biomedical signal processing. *Biomedical Signal Processing and Control*, 14, 19-29.

[10] WEI Xiu-lei, LIN Rui-lin, LIU Shu-yong et al. (2015). Hybrid wavelet-SG-EEMD algorithm and its application in chaotic de-noising. *Journal of Vibration and Shock*, 34, 100-110.

[11] Bruno M, Schettino, Carlos A, Duque and Paulo M, Silveira. (2016). Current-Transformer Saturation Detection Using Savitzky-Golay Filter. *IEEE TRANSACTIONS ON POWER DELIVERY*, 31, 1400-1401.

[12] JIANG shao-fei, CHEN Zhi-gang, SHEN Qing-hua, et al. (2016). Damage detection and locating based on EEMD-FastICA. *Journal of Vibration and Shock*, 35, 203-209.

[13] A. Hyvärinen. (1999). Fast and robust fixed-point algorithms for independent component analysis. *IEEE Transactions on Neural Networks*, 10, 626–634.