Scientific Research Publishing

Int. J. Communications, Network and System Sciences, 2017, 10, 138-147
http://www.scirp.org/journal/ijcns
ISSN Online: 1913-3723
ISSN Print: 1913-3715

Application of Single Channel Blind Separation Algorithm Based on EEMD-PCA-RobustICA in Bearing Fault Diagnosis

Wei Xu, Xiangzhou Yan

College of Information and Communication Engineering, Harbin Engineering University, Harbin, China
Email: xuwei@hrbeu.edu.cn, yanxiangzhou@hrbeu.edu.cn

How to cite this paper: Xu, W. and Yan, X.Z. (2017) Application of Single Channel Blind Separation Algorithm Based on EEMD-PCA-RobustICA in Bearing Fault Diagnosis. *Int. J. Communications, Network and System Sciences*, **10**, 138-147.
https://doi.org/10.4236/ijcns.2017.108B015

Received: May 25, 2017
Accepted: August 11, 2017
Published: August 14, 2017

Abstract

Aiming at the problem that ICA can only be confined to the condition that the number of observed signals is larger than the number of source signals; a single channel blind source separation method combining EEMD, PCA and RobustICA is proposed. Through the eemd decomposition of the single-channel mechanical vibration observation signal the multidimensional IMF components are obtained, and the principal component analysis (PCA) is performed on the matrix of these IMF components. The number of principal components is determined and a new matrix is generated to satisfy the over-determined blind source separation conditions, the new matrix input RobustICA, to achieve the separation of the source signal. Finally, the isolated signals are respectively analyzed by the envelope spectrum, the fault frequency is extracted, and the fault type is judged according to the prior knowledge. The experiment was carried out by using the simulation signal and the mechanical signal. The results show that the algorithm is effective and can accurately diagnose the location of mechanical fault.

Keywords

EEMD, PCA, RobustICA, Envelope Spectrum, Fault Diagnose

1. Introduction

In the traditional signal processing, it is generally necessary to know in advance some of the prior knowledge of the signal or the mathematical model of the signal mixing matrix, and then estimate the source signal by filtering or transforming. In practice, the prior knowledge of the signal is not easy to obtain, signal processing cannot solve the problem. The advantage of blind source separation is that it can use the least prior knowledge to obtain the greatest information, the problem originally comes from the cocktail effect, in the cocktail meeting, noisy voices and background music is great, but if people talk to a conversation Interested, still can hear this voices.

The classical blind separation algorithm can recover the source signal better when the number of observed signals is larger than the number of source signals, FastICA algorithm, RobustICA algorithm, second order blind identification algorithm, joint approximation diagonalization and so on. In case of the underdetermined condition, the above algorithm cannot solve the problem. For the underdetermined model, new algorithms are proposed. Based on the blind separation of time-frequency distribution, sparse signal analysis is the main method to solve this problem.

Single channel blind separation is an extreme condition of the underdetermined condition, that is, only through single channel observation signal to estimate the multichannel source signals, in real life, due to environmental or cost constraints, often encountered such extreme problems. In this case, some scholars decompose the signal with wavelet, the resulting signal component is subjected to ICA processing, finally get the source signals [1]; Some scholars put forward the method of space-time, the method is to delay the mixed signal collected multi-channel signals, and then use the independent component analysis algorithm of multiple mixed signal separation, thus realize the rotating machinery fault diagnosis [2].

Based on the existing methods, EEMD, PCA and RobustICA are applied to the mechanical fault diagnosis. Firstly, the EEMD algorithm is used to increase the dimension of the single observation signal, and the number of the source signal is estimated by PCA algorithm. Then, the signal is processed by RobustICA. The experimental results show that the method can effectively isolate the mechanical fault of each part.

2. Methodology

2.1. Ensemble Empirical Mode Decomposition

1998, Norden E. Huang, first proposed the intrinsic Mode Function (Instrinc Mode Function, the IMF) and its Decomposition method, the concept of Empirical Mode Decomposition (EMD) [3].

EMD is a new algorithm in the field of modern signal processing. With the popularity of the algorithm and a wide range of applications, its inherent defects are inevitably exposed. Modal aliasing is the most important and deadly defect in empirical mode decomposition. It refers to the fact that the signal is interrupted during the decomposition process due to the weak interference of the signal to be decomposed, so that the adjacent natural modal components are superimposed together, Masking the instantaneous nature of the source signal. In response to this problem, Huang *et al.* proposed a collective empirical mode decomposition (EEMD) method [4]. The empirical empirical modal decomposition algorithm is a further sublimation of the empirical modal decomposition algorithm, which adds the normal distribution of Gaussian white noise on the basis of the empirical mode decomposition processing signal data, and simultaneously eliminates the random scale of the uniform distribution of the signal to be decomposed The interference can effectively suppress the aliasing phenomenon, so that the decomposition of the inherent modal component with its proper physical meaning [4].

EEMD decomposition principle is as follows:

1) In the single channel mixed signal $x(t)$ which needs to be decomposed, add Gaussian white noise with the mean value of zero and the standard deviation σ (usually 0.1 - 0.4), which is $n(t)$, that is, (subscript represents the first decomposition):

$$x_1(t) = x(t) + n_1(t) \tag{1}$$

2) Decomposes the signal $X_1(t)$ into a series of IMF components with EMD, we can get:

$$x_1(t) = \sum_{i=1}^{m} c_i^1(t) + r_1(t) \tag{2}$$

3) Repeat the above steps to continue adding the same Gaussian white noise (eg. Kth decomposition) during the repetition process:

$$x_k(t) = x(t) + n_k(t) \tag{3}$$

Perform EMD decomposition:

$$x_k(t) = \sum_{i=1}^{m} c_i^k(t) + r_k(t) \tag{4}$$

4) Repeat the N times to find the average value of the decomposed groups of intrinsic modal components and the signal margin:

$$\mathrm{c}_i(t) = \frac{1}{N}\sum_{k=1}^{N} c_i^k(t) \tag{5}$$

$$r_m(t) = \frac{1}{N}\sum_{k=1}^{N} r_k(t) \tag{6}$$

EEMD decomposition results were obtained:

$$x(t) = \sum_{i=1}^{m} c_i(t) + r_m(t) \tag{7}$$

Since the Gaussian white noise with a mean value of zero is added to the single channel mixed signal that needs to be decomposed, after the EEMD decomposition, the intrinsic modal component are averaged, and the white noise in the result will eventually cancel each other, and the noise is eliminated at the same time, and the phenomenon of modal aliasing is avoided.

2.2. Principal Component Analysis

In order to achieve single channel blind separation, first estimate the number of source signals, that is, after EEMD decomposition with principal component analysis (PCA) to estimate the number of source signals. The purpose of principal component analysis is to find r (r less than n) new vectors, which are used to represent the main features of the entire n-dimensional vectors, thus reducing the dimensionality of the original vector and compressing the entire matrix. Each of the new variables is a linear combination of the original variables, that is, extracted "principal components", each principal component is uncorrelated, and orthogonal, with practical physical meaning, can represent the original n-dimensional vector of the whole feature. By PCA processing, the n-dimensional vector is reduced to r dimension.

For PCA algorithm, how to solve the new vector number r is particularly important. Although the value of r in the algorithm is as small as possible, the smaller the r, the lower the dimension, making the result analysis simpler and less interfering, but this may make some of the key information lost. In order to solve the problem of data loss which may occur in the above algorithm, the solution is to analyze the contribution rate of any vector to information. Contribution rate refers to the proportion of the principal component used in all the data analysis. In determining the r principal component, unless there is a special requirement, the general requirements of the contribution rate to reach more than 85%, because the contribution rate to a certain extent, reflects the size of the reliability, that is, the greater the contribution of the principal component, the greater the reliability.

2.3. RobustICA Algorithm

The blind source separation model is as follows: Assuming that N independent source signals are received by M sensors, the source signal is randomly mixed through an unknown mixing system to form an observation signal, that is:

$$X(t) = AS(t) + n(t) \tag{8}$$

where $X = (x_1, x_2, \cdots, x_m)^T$ represents the observed m mixed signals; $S = (s_1, s_2, \cdots, s_n)^T$ is n unknown source signal vector; $A = (a_{ij})_{m\times n}$ is $m \times n$ dimensional mixed matrix; $n = (n_1, n_2, \cdots, n_m)^T$ is the noise vector received by m sensors, and must satisfy the condition $m \le n$.

In the modern signal processing process, in most cases the noise can also be considered a class of source signals, or that through other methods to reduce the noise to a negligible level, so the content of this study does not take into account the noise problem, Equation (8) can be rewrite as:

$$X(t) = AS(t) \tag{9}$$

The goal of the blind source separation algorithm is to solve the separation matrix W, and then realize the estimation of the source signal. The estimated values of the source signals derived from the theory are shown below:

$$Y(t) = WX(t) = WAS(t) \tag{10}$$

$$W = A^{-1} \tag{11}$$

The key problem with blind source separation is to find the solution matrix W [5] [6].

RobustICA uses the independent component analysis algorithm based on kurtosis and optimal step length to search the global optimal step by using the sentinel as the control function, find the solution matrix W, and calculate the approximate value of the original signal [7].

2.4. Single Channel Blind Separation Algorithm Based on Ensemble Empirical Mode Decomposition and Principal Component Analysis

The method of single channel blind separation based on empirical mode decomposition and principal component analysis is the process of recovering the source signal from the observed signal. Combined with EEMD, PCA and RobustICA, the problem of single channel blind separation and the problem of source number estimation are solved. The process is as follows:

1) The single channel observation signal x is decomposed by EEMD to obtain the IMFs components.

2) Using pca to reduce the IMFs component dimension, Select several elements whose contribution rate is 95% to constitute a new signal. The dimension of this multidimensional signal is the number of source signals estimated by the PCA.

3) The new multi-dimensional signal is processed by RobustICA, and the blind source is separated to obtain the separated source signal [8].

3. Matlab Simulation

3.1. Experimental Simulation Signal

Here, the experimental signals are selected as sinusoidal signal $s_1(t)$, cosine signal $s_2(t)$, AM modulation signal $s_3(t)$, its expression and specific parameters are as follows

$$s_1(t) = \sin(2\pi f_1 t + \theta_1) \tag{12}$$

$$s_2(t) = \cos(2\pi f_2 t + \theta_2) \tag{13}$$

$$s_3(t) = 2\sin(2\pi f_3 t)(1 + \cos(2\pi f_4 t + \theta_3)) \tag{14}$$

where $\theta_1 = \theta_2 = \theta_3 = 0°$, $f_1 = 15Hz$, $f_2 = 35Hz$, $f_3 = 120Hz$, $f_4 = 20Hz$. The signal sampling frequency is 1024 Hz, the signal length is 512, then the source signal time domain diagram as "**Figure 1**".

The source signal $[s_1, s_2, s_3]$ is mixed by 1×3 random matrix A, which is taken as $[0.5, 0.5, 0.3]$ to obtain a single channel mixed signal $s(t)$. The mixed signal $s(t)$ time domain waveform is shown in "**Figure 1**". Use EEMD to decompose the mixed signal, obtain a series of IMF components, the decomposition results is shown in "**Figure 2**".

PCA on this series of IMF components to reduce the dimension, select the main component whose contribution rate is 95% to constitute a new signal. The dimension of this multidimensional signal was selected 3 by MATLAB, and the reconstruction signal is shown in "**Figure 3**".

Regard the reconstructed signals as new observation signals, then use RobustICA to isolate the source signals, The results are shown in "**Figure 4**".

Calculated by MATLAB, the correlation coefficient between each signal in

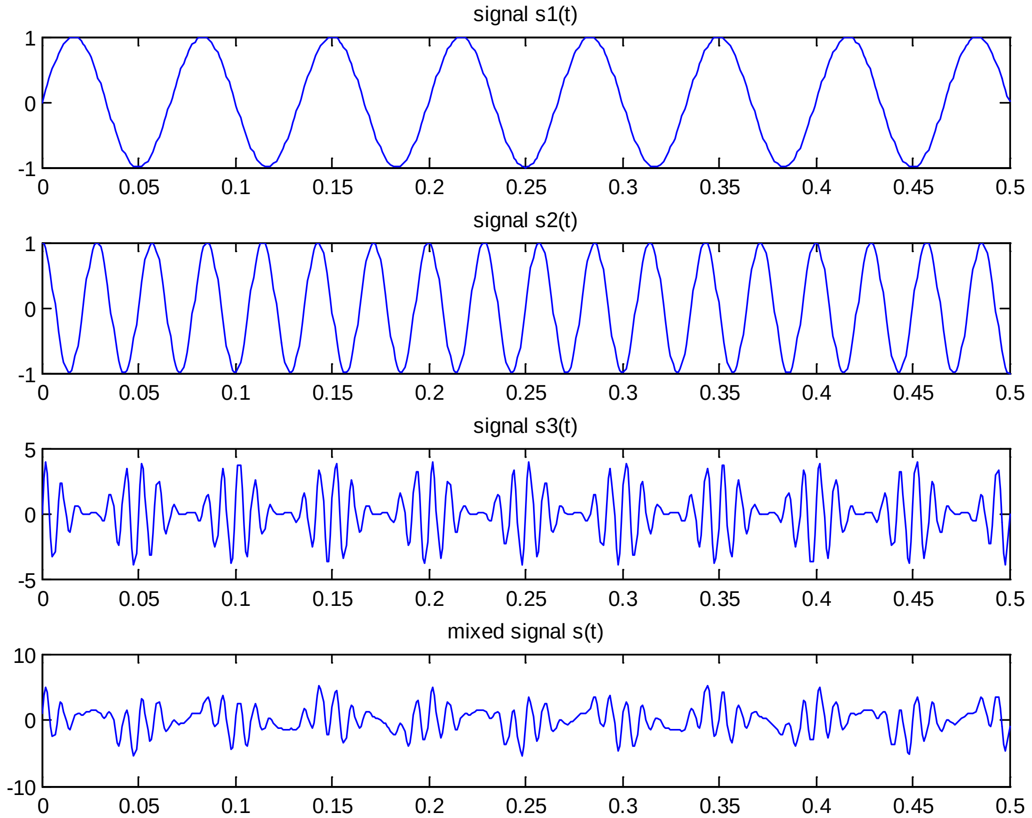

Figure 1. Source signals and mixed signal.

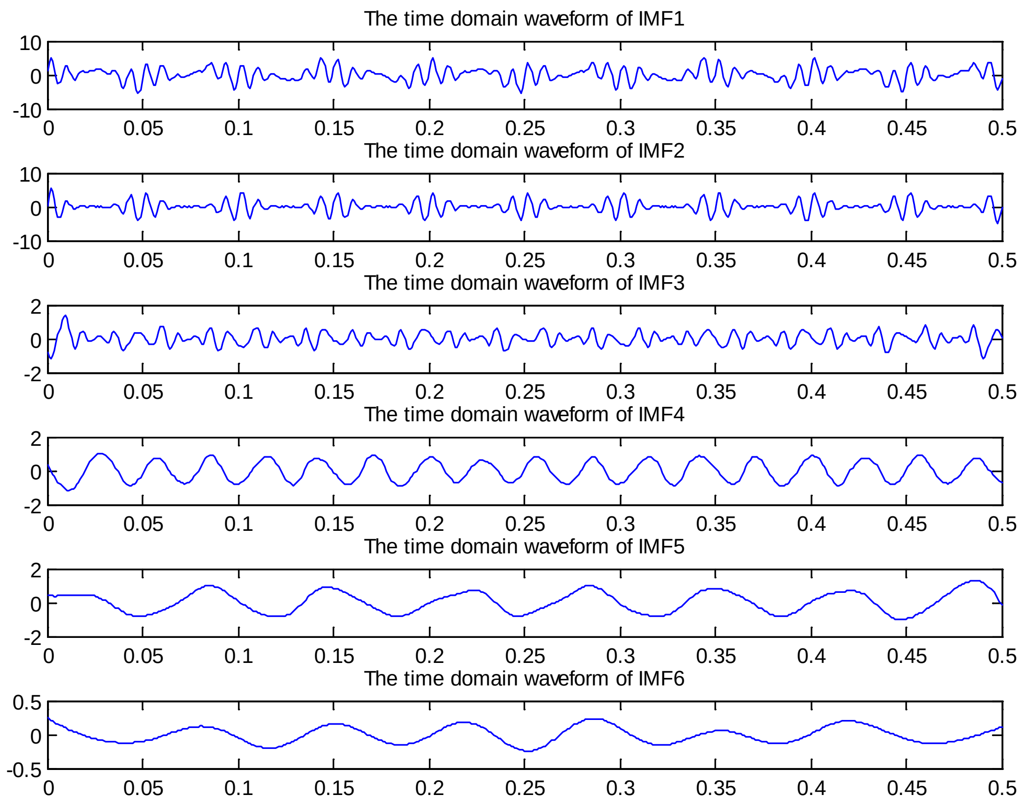

Figure 2. EEMD decomposition results of mixed signal $s(t)$.

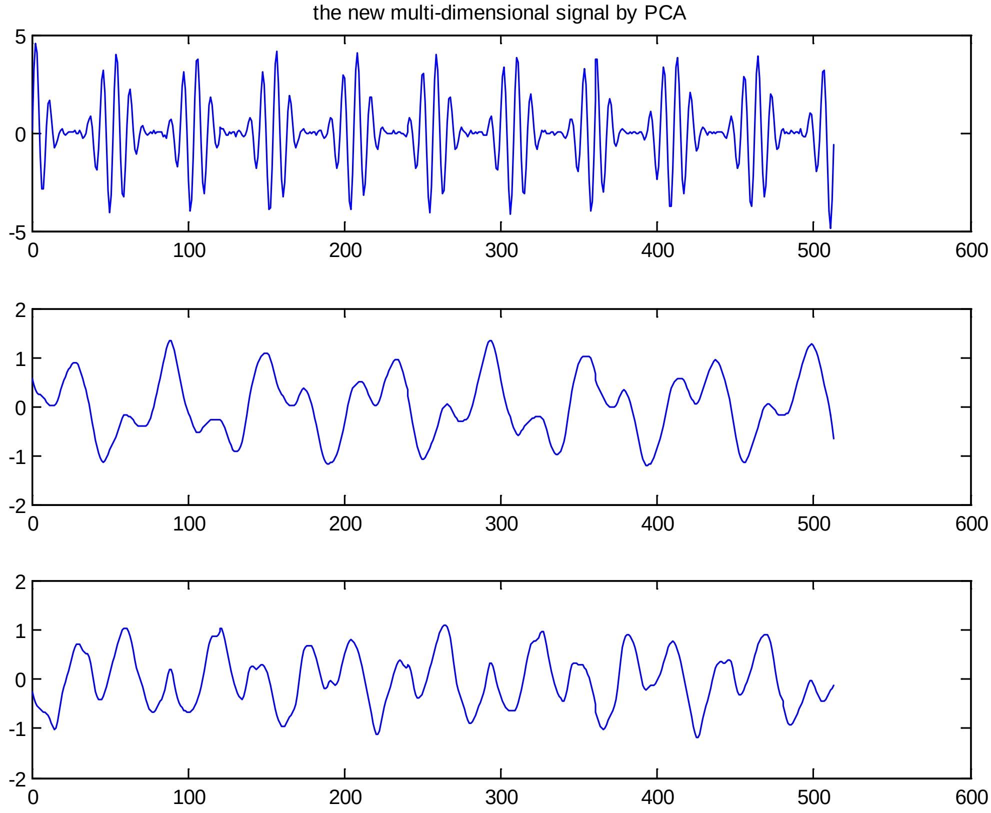

Figure 3. Input signal LFM1 frequency domain waveform.

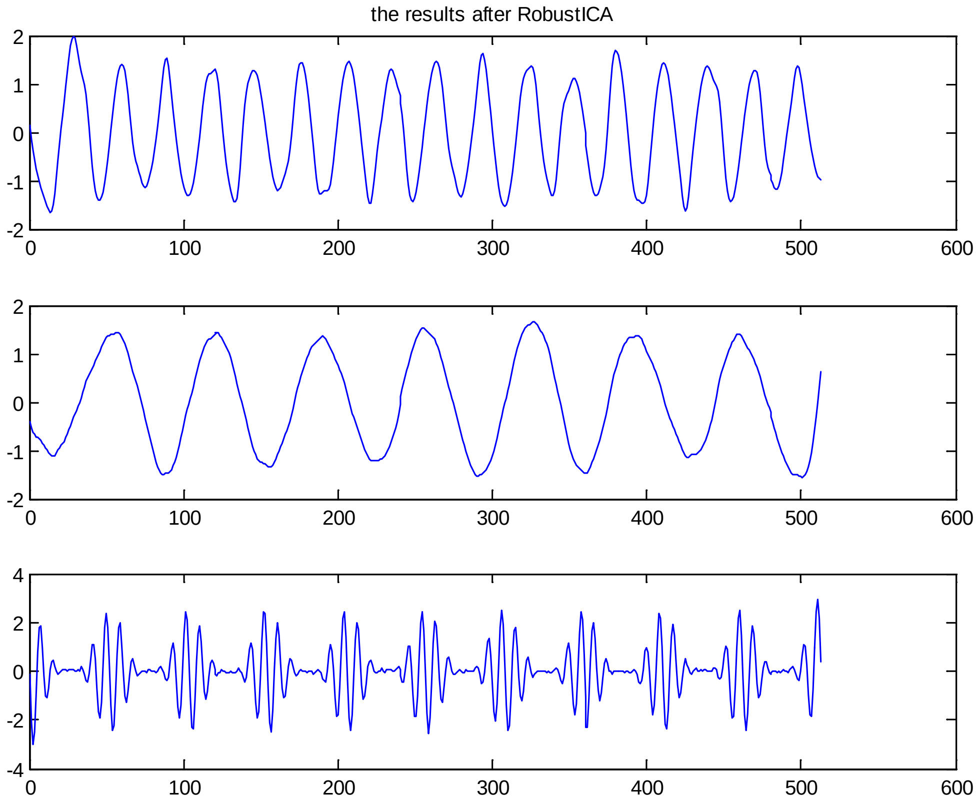

Figure 4. Input signal LFM2 frequency domain waveform.

"**Figure 4**" and the source signals are $0.9609, 0.9824, 0.9769$, indicating that the separated signal is very similar to the source signal.

3.2. Actual Mechanical Fault Signal Simulation

The experimental data are from the electrical engineering laboratory of Case Western Reserve University. The bearing fault types include inner ring fault, outer ring fault and rolling element fault. The inner ring fault and outer ring fault are selected as the fault source signal from the bearing state. The source signal is linearly mixed with a 1 × 2 random matrix A to obtain a mixed fault signal, *i.e.* a single channel observation signal. In **Figure 5**, $s_1(t)$ is the inner ring fault signal time domain waveform, $s_2(t)$ is the outer ring fault signal time domain waveform, $s(t)$ is the single composite fault observation signal time domain waveform.

The single-channel fault observation signal is processed by the EEMD-PCA-RobustICA method described in the previous chapter, and the resulting separation signal is shown in **Figure 6**. The correlation coefficient between the separated signal 1 and the outer ring signal is 0.96, and the correlation coefficient between the separated signal 2 and the inner ring signal is 0.88, which proves that the single channel blind separation algorithm of EEMD-PCA-RobustICA is in the actual signal Equally effective.

Figure 7 shows the envelope of the separation signal 1, in the figure can be seen 104.7 Hz peak, and the characteristics of the outer ring fault coincides with the map, in the figure you can see the peak of the separation signal 2 is 157.5 Hz, and the characteristics of the inner ring fault coincides with the frequency. So

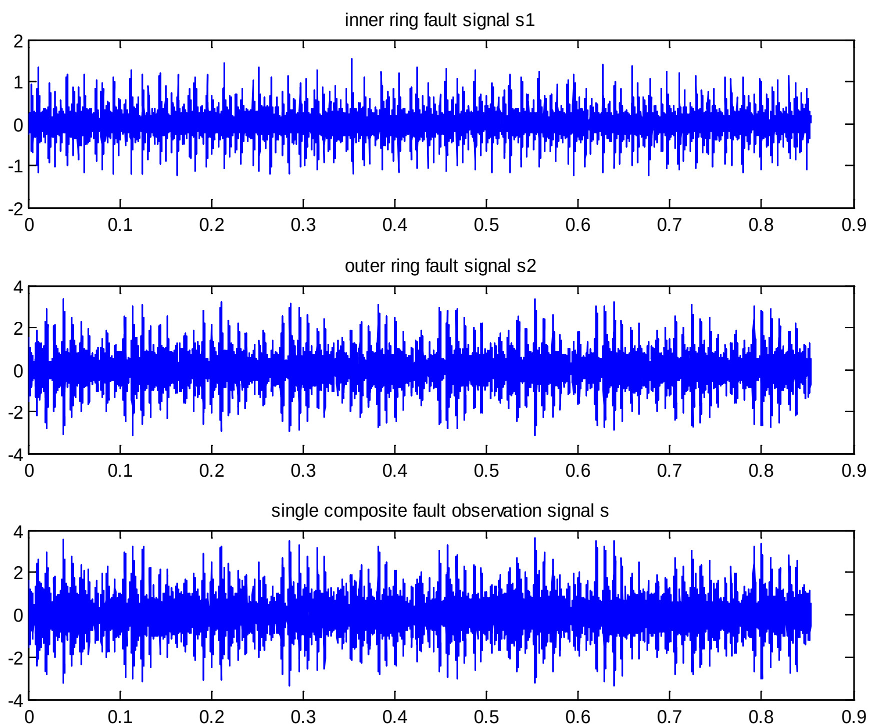

Figure 5. Bearing fault signal.

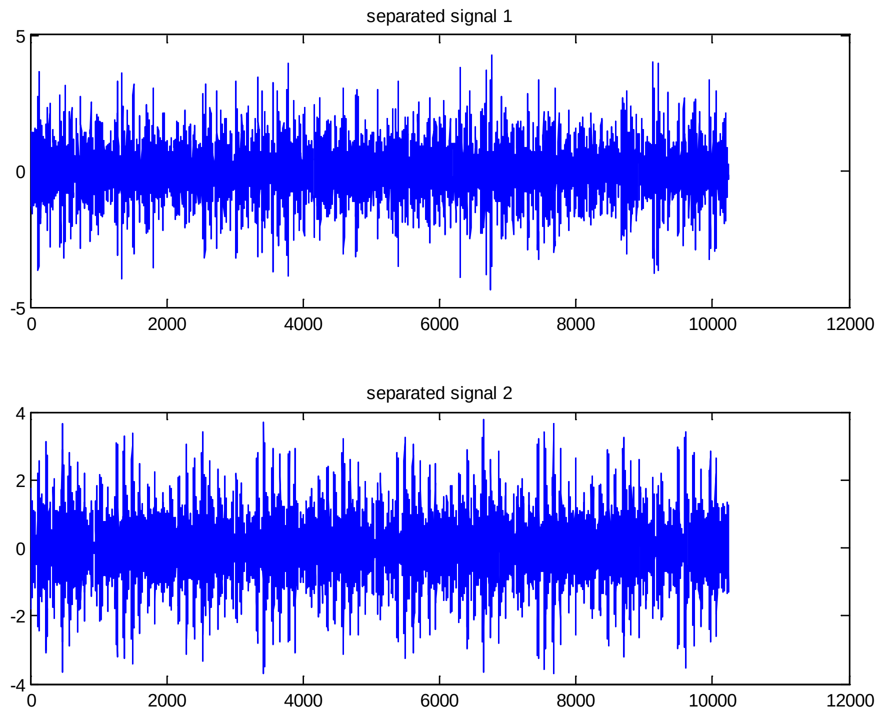

Figure 6. The separation signal.

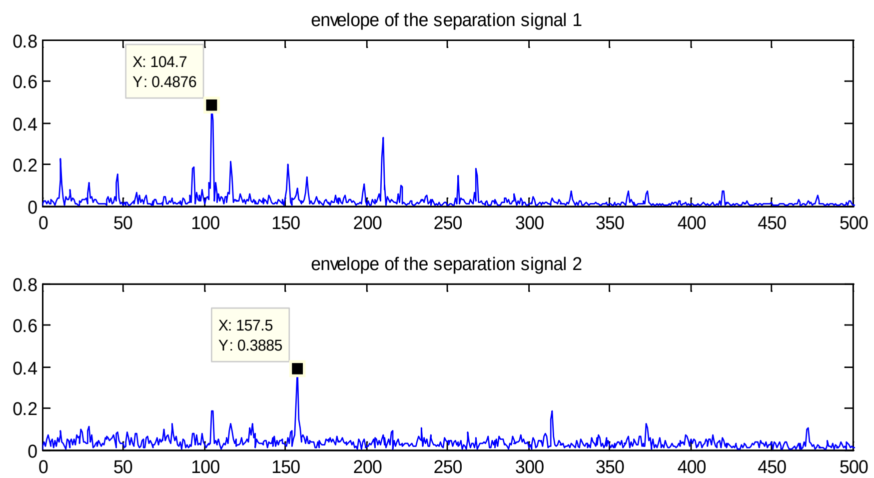

Figure 7. Envelope of the separation signal.

the separation signal 1 can be diagnose as outer fault, and the separation signal 2 can be diagnose as inner fault, which verifies that the algorithm can accurately determine the type of mechanical fault [9].

4. Conclusion

In this paper, the EEMD-PCA-RobustICA method is proposed, and the simulation signal and the actual mechanical fault signal are used to experiment. The source signal is well separated from the single channel observation signal, which proves the effectiveness of the method. At the same time, the method used in mechanical fault diagnosis, can effectively determine the type of mechanical fault.

References

[1] Shao, H., Shi, X.H. and Li, L. (2011) Power Signal Separation in Milling Process Based on Wavelet Transform and Independent Component Analysis. *International Journal of Machine Tools and Manufacture*, **51**, 701-710. https://doi.org/10.1016/j.ijmachtools.2011.05.006

[2] Davies, M.E. and James, C.J. (2007) Source Separation Using Single Channel ICA. *Signal Processing*, **87**, 1819-1832. https://doi.org/10.1016/j.sigpro.2007.01.011

[3] Huang, N.E., Shen, Z., Long, S.R., *et al.* (1998) The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Non-Stationary Time Series Analysis. *Proceedings of the Royal Society of London A*: *Mathematical*, *Physical and Engineering Sciences*, The Royal Society, 903-995. https://doi.org/10.1098/rspa.1998.0193

[4] Wu, Z.H. and Huang, N.E. (2009) Ensemble Empirical Mode Decomposition: A Noise Assisted Data Analysis Method. *Advances in Adaptive Data Analysis*, **1**, 1-41. https://doi.org/10.1142/S1793536909000047

[5] Comon, P. (1994) Independent Component Analysis: A New Concept. *Signal Processing*, **36**, 287-314. https://doi.org/10.1016/0165-1684(94)90029-9

[6] Hyvärinen, A. (1999) Fast and Robust Fixed-Point Algorithms for Independent Component Analysis. *IEEE Transactions on Neural Networks*, **10**, 626-634. https://doi.org/10.1109/72.761722

[7] Vicente, Z. and Pierre, C. (2010) Robust Independent Component Analysis by Iterative Maximization of the Kurtosis Contrast with Algebraic Optimal Step Size. *IEEE Trans Neural Networks*, **21**, 248-261. https://doi.org/10.1109/TNN.2009.2035920

[8] Mijovic, B., De Vos, M., Gligo Rjevic, I., *et al.* (2010) Combining EMD with ICA for Extracting Independent Sources from Single Channel and Two-Channel Data. *Proceedings of the* 32*nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 5387-5390. https://doi.org/10.1109/IEMBS.2010.5626482

[9] Konar, P. and Chattopadhyay, P. (2011) Bearing Fault Detection of Induction Motor Using Wavelet and Support Vector Machines (SVMs). *Applied Soft Computing*, **11**, 4203-4211. https://doi.org/10.1016/j.asoc.2011.03.014