

# Independent Component Analysis Method Based on Genetic Algorithm in Compound Fault Separation

Yansong Hao<sup>1</sup>, Liuyang Song<sup>2</sup>, Jingle Li<sup>3</sup>, Xin Xie<sup>4</sup>, Huaqing Wang<sup>1,\*</sup>

<sup>1,2,3,4</sup> *Beijing University of Chemical Technology, Beijing, 100029, China*

*hqwang@mail.buct.edu.cn*

*hys\_buct@163.com*

## ABSTRACT

Compound faults often exist in roller bearing, which increase the difficulty for the fault diagnosis. In order to extract the characteristics of compound fault signals, independent component analysis (ICA) method was used to separate fault signals. However, the selection of initial weight about ICA will affect the number of iterations and astringency. Therefore, a novel ICA method based on improved genetic algorithm was proposed. The kurtosis of signal was used as the optimization function, and then genetic algorithm was applied to find the separation matrix according to the maximum of the kurtosis. This method avoids the problem of redundant iteration and convergence, which is caused by randomness of initial weight vector. Finally, the proposed method was used to separate source signal from the mixed signals and achieve the roller bearings fault identification and separation. The results show that the proposed method is superior to traditional ICA method, and the compound fault can be separated based on the proposed method.

**Key words:** roller bearing; independent component analysis; genetic algorithm; compound faults separation;

## 1. INTRODUCTION

Roller bearing, whose failures may lead to the breakdown of whole mechanical system, is an integral component of rotating machinery [1-2]. For the fault diagnosis of rotating machinery, vibration of the observed signal contains a lot of noise. And how to accurately extract the fault signal from the observed signal is the key for equipment of fault diagnosis. Therefore, it is of great significance to carry out fault diagnosis of roller bearings, guaranteeing the rotating machinery to operate in a normal and stable condition. However, compound faults often occur in practical operation and the traditional fault diagnosis approaches, such as envelope spectrum, cannot extract the compound fault features of roller bearings.

To solve above problem, Independent Component Analysis (ICA) was proposed to separate source component [3-4]. ICA is a method based on high order statistics of signals, which can be used to find the independent components in the mixed signals. Therefore, ICA has been widely concerned in the field of mechanical fault signal processing. But the ICA algorithm mostly adopts the gradient method to optimize the separation matrix, and the initial value and the setting step will have great influence on the performance of the algorithm. Moreover, the algorithm is easy to fall into local optimum values, which greatly affect the separation ability of ICA and reduce the separation effect of the algorithm. Genetic algorithm is a random search and optimization method based on global mimic natural biological evolution [5-6]. Therefore, a new method of compound fault separation is proposed by combining ICA method with genetic algorithm to overcome the defects of the ICA algorithm.

## 2. FUNDAMENTAL THEORY

### 2.1. Independent Component Analysis [7-8]

Independent component analysis (ICA) is a kind of analysis method based on high order statistics of signals and its mathematical models are as follows:

$$x = As \quad (1)$$

$x$  is observation signal,  $A$  is mixed matrix, and  $s$  is source signal. The purpose of ICA is that the separation matrix can be obtained only based  $x$  when  $s$  and  $A$  is unknown, which make separation component is statistically independent as far as possible to estimate the source signal. And  $y$  is the best estimate of the source signal.

$$y = A^{-1}x = Wx \approx s \quad (2)$$

### 2.2. Genetic Algorithm [9-10]

Genetic algorithm is a novel randomized directed search algorithm, based on the natural selection and genetic mechanism. Its basic process is as follows:

- (1) Code. Code is the representation of the feasible solution in the solution space into a gene string structure.
- (2) Initialization. The size of the population, crossover probability, mutation probability and evolution termination criteria are determined.
- (3) Fitness evaluation. The fitness function is used to evaluate the quality of each chromosome.
- (4) Choice. The parent generation is selected randomly into the next generation of iterative process based on the fitness value of each chromosome, and the low adaptability individuals are abandoned.
- (5) Crossover. A new generation of offspring is generated by the crossover of the gene segment
- (6) Mutation. A single gene mutation was carried out based on the probability of mutation by randomly selecting some individuals.

**3. COMPOUND FAULTS SEPARATION BASED PROPOSED METHOD**

It is known that the separation of matrix is the key of ICA algorithm and the accuracy of the matrix directly affects the separation effect. The method was proposed to overcome the shortcoming of gradient method and avoid algorithm falling into local optimum values when optimize the separation matrix. Firstly, EEMD method was used to decompose the mixed signal into several channels, which meets the ICA condition. Then, Genetic algorithm was used to find the separation of matrix. Non Gauss property can be expressed by the four order cumulant that is kurtosis, so kurtosis can indirectly indicate the independence of the signal, it was used to evaluate the merits of chromosomes. In addition, the constraint of orthogonality was added to ensure that the vector of the separation matrix was not in the same direction in the process of finding the separation matrix. Finally, the results show that the compound faults signals can be separated successfully based on the proposed method, compared with the traditional ICA method. The flow chart of compound faults separation is shown in Figure 1.

**4. EXPERIMENTAL RESULTS AND DISCUSSION**

The vibration signals of roller bearings faults are collected through a couple of accelerometer. And the experimental system of bearing diagnosis, including a rotating machine, a roller bearing and acceleration sensors, is shown in Figure 2. The size of flaws are all 0.5\*0.15mm (width\*depth), artificially created by using a wire-cutting machine. In addition, the sampling frequency is 100 KHz, the sampling time is 10s and the rotating speed of a machine is 900 rpm. The fault passing frequency of each element of roller bearings can be calculated according to Eq. (3-4). The results are shown in Table 1. Where  $D$  is the pitch diameter,

$Z$  is the number of rollers,  $d$  is the diameter of rollers,  $\alpha$  is the contact angle of the rollers, and  $f_r$  is the rotating frequency.  $f_o$  and  $f_b$  also represent the fault characteristic frequency of outer race and rollers.

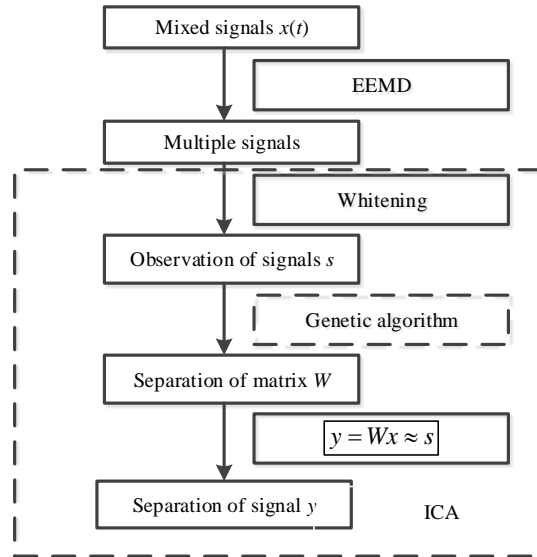


Figure 1. Flowchart of the proposed method.

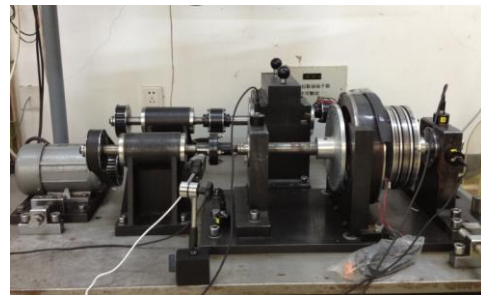


Figure 2. Experimental system of roller bearing diagnosis

$$f_o = \frac{Z}{2} (1 - \frac{d}{D} \cos \alpha) f_r \tag{3}$$

$$f_b = \frac{D}{2d} [1 - (\frac{d}{D} \cos \alpha)^2] f_r \tag{4}$$

Table 1. Fault characteristic frequency of roller bearings at 900rpm.

Fault location	Fault characteristic frequency
Outer race	59.8Hz
Rollers	71.8Hz

The waveforms and envelope spectrum of original signals collected by the experimental system at 900 rpm are shown in Figure 3 and Figure 4. Figure 3 shows that there are obvious impulses in the time domain. And in Figure 4, the frequency of outer race at 60.12Hz which is very close to the theoretical calculation  $f_o$ . It means that the fault exists

in the outer race. However, because only the outer race fault can be extracted, the compound faults cannot be separated based on the spectrum analysis. Therefore, the proposed method was used in the experiment.

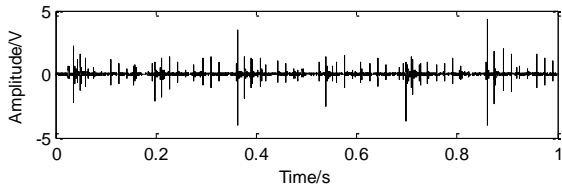


Figure 3. The waveform of original signals

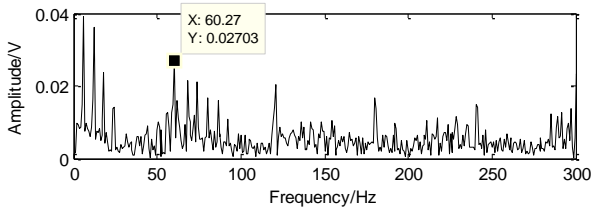


Figure 4. The envelope spectrum of original signals

Firstly, EEMD was used to decompose the original signal into multiple signals. Four high correlation signals are shown in Figure 5, which were selected based on the correlation coefficient. And the envelope spectrum of the four components is shown in Figure 6. Secondly, the decorrelation and whitening of the four signals were used as observation signals. Chromosome with the largest kurtosis value was used to be a good parent to cross, which based on the calculation of each component kurtosis. Then, separation of matrix which was shown in Eq. (5) was obtained based on genetic algorithm. Therefore, the separation of signal can be obtained based on the matrix and observation according to Eq. (2). The waveform of separation signal is shown in Figure 7, and envelope spectrum is shown in Figure 8. The frequency of 60.27Hz in the Figure 8(c) is very close to the theoretical calculation  $f_o$  and the frequency of 74.01Hz in the Figure 8(b) is close to the  $f_b$ . It can be seen that the compound faults which include the defect of outer race and rollers are extracted based on the proposed method. Figure 9 shows that only outer race defect can be recognized which means traditional ICA cannot perform well. As a result, the proposed method can be used to separate compound fault signals effectively which is superior to traditional ICA method.

$$A = \begin{bmatrix} -0.8081 & 0.7217 & 0.6960 & -0.4174 & 0.7880 \\ -0.3015 & -0.2513 & -0.3894 & 0.5279 & 0.0258 \\ -0.1869 & -0.2136 & -0.0692 & -0.0348 & -0.0764 \\ 0.3609 & 0.5542 & 0.4556 & 0.5170 & 0.6098 \end{bmatrix} \quad (5)$$

5. CONCLUSION

How to separate complex faults is an important problem in fault diagnosis. In this paper, a novel method of compound

fault separation is proposed by combining ICA method with genetic algorithm. The genetic algorithm is used to find the separation matrix to avoid the problem that the traditional ICA algorithm is trapped in the local extremum and improve algorithm efficiency in the process of finding the separation matrix. The experimental results show that the proposed method obviously extracts compound fault features which include the defects of outer race and roller. In this approach, the run time of algorithm needs to be further reduced. The future work will focus on the point of the method.

ACKNOWLEDGEMENT

This work is partially supported by the National Natural Science Foundation of China (Grant no.51675035).

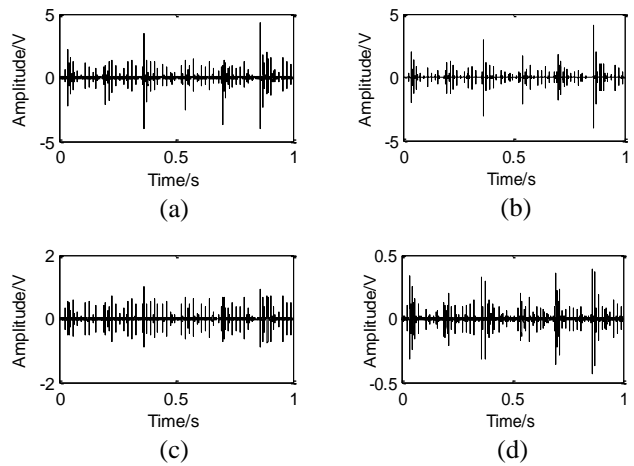


Figure 5. Four components decomposed by EEMD: (a) IMF1 (b) IMF2 (c) IMF3 (d) IMF4.

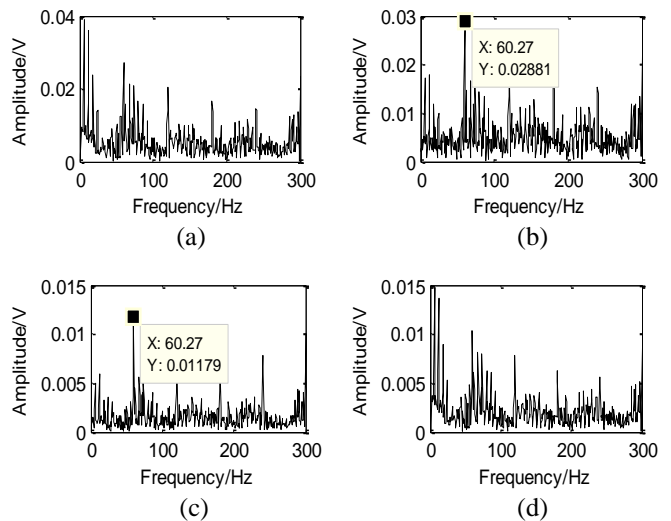


Figure 6. The envelope spectrum of four components: (a) IMF1 (b) IMF2 (c) IMF3 (d) IMF4.

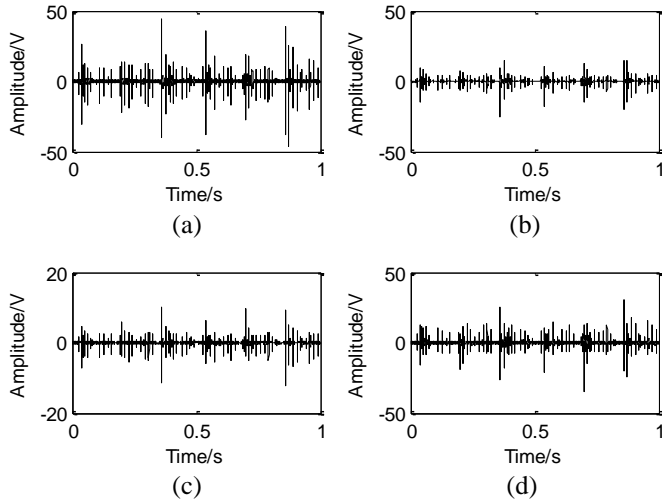


Figure 7. The waveform of separation signal:  
 (a) First separation signal (b) Second separation signal  
 (c) Third separation signal (d) Fourth separation signal

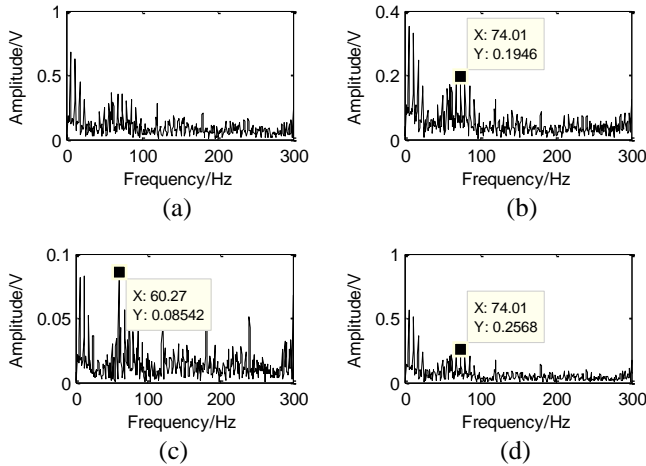


Figure 8. The envelope spectrum of separation signal:  
 (a) First separation signal (b) Second separation signal  
 (c) Third separation signal (d) Fourth separation signal

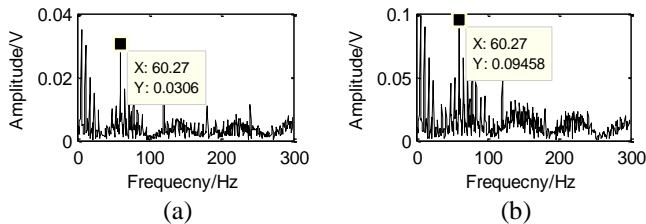


Figure 9. The envelope of separation signal by ICA  
 (a) First separation signal (b) Second separation signal

REFERENCES

Y Yang, D. J. Yu., & J. S. Cheng. (2007). A fault diagnosis approach for roller bearing based on IMF envelope spectrum and SVM. *Measurement*, 40(9-10):943-950.

Yau, H. T., Kuo, Y. C., Chen, C. L., & Li, Y. C. (2016). Ball bearing test-rig research and fault diagnosis investigation. *Iet Science Measurement & Technology*, 10(4), 259-265.

Mahvash, A., & Lakis, A. A. (2014). Independent component analysis as applied to vibration source separation and fault diagnosis. *Journal of Vibration & Control*, 22(6), 1682-1692.

Shlens, J. (2014). A tutorial on independent component analysis. *Computer Science*.

Delorme, A., & Makeig, S. (2004). Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1), 9.

De Lathauwer, L., De Moor. B., & Vandewalle, J. (2015). An introduction to independent component analysis. *Journal of Chemometrics*, 14(6), 123-149.

Goldberg, D. E. (1989). Genetic algorithm in search, optimization, and machine learning. *Xiii*, (7), 2104-2116.

Rlg, R., Fonseca, R. L., Schachettipereira, R., Peterson, A. T., & Lewinsohn, T. M. (2016). Native and exotic distributions of siamweed (*chromolaena odorata*) modeled using the genetic algorithm for rule-set production. *Weed Science*, 55(Jan 2007), 41-48.

Kobayashi, T., & Simon, D. L. (2015). Hybrid neural-network genetic-algorithm technique for aircraft engine performance diagnostics. *Journal of Propulsion & Power*, 21(4), 751-758.

Simpson, T., & D'Souza, B. (2013). Assessing variable levels of platform commonality within a product family using a multiobjective genetic algorithm. *Concurrent Engineering*, 12(2), 119-129.

BIOGRAPHIES

**Huaqing Wang** He is a doctor and doctoral tutor, born in March 1973. And now he is a Professor of Beijing University of Chemical Technology and College of mechanical and Electrical Engineering.

**Yansong Hao** He is a postgraduate, born in August 1992. And now he is studying in Beijing University of Chemical Technology and College of mechanical and Electrical Engineering.