Improved Ensemble Empirical Mode Decomposition for Rolling Bearing Fault Diagnosis

Youpeng Zhang*, Ting Zhang, Jie Teng, Hongsheng Su

School of Automatic & Electrical Engineering, Lanzhou Jiaotong University, Lanzhou, 730070, China Address: 28 mailbox, Lanzhou Jiaotong University, An-ning district, Lanzhou, 730070 Gansu Province, China *Corresponding author, E-mail address: zhangtingylj@163.com

Abstract

Rolling bearing is an important part in mechanical system and faults occur frequently with vibration noise. Empirical mode decomposition (EMD) is a tool for nonlinear and non-stationary signals analysis. However, the major drawbacks of EMD are mode mixing problem, ensemble empirical mode decomposition (EEMD) provides a new tool for signal analysis, and it is an improved technique of EMD. In order to alleviate the mode mixing problem and choose useful IMFs, a method called EEMD and distributing fitting testing is proposed in this paper, and it is used in rolling bearing fault diagnosis. Firstly, using it for rolling bearing fault diagnosis, the fault signal is decomposed by EEMD. Then applying distributing fitting testing to choose components with truly physical meaning and the de-noised signal can be obtained. Finally, utilizing envelope spectrum to distinguish different faults. The results demonstrate the proposed method can sift useful IMFs and diagnose faults effectively, such as inner race fault, outer race fault. The advantage of the proposed method is suitable for rolling bearing diagnosis.

Keywords: Ensemble empirical mode decomposition (EEMD); Empirical mode decomposition (EMD); Distributing fitting testing; Fault diagnosis; Envelope spectrum

Copyright © 2014 Institute of Advanced Engineering and Science. All rights reserved.

1. Introduction

Rolling bearing is an important part in large rotating machinery. Because of running frequently, rolling bearing's potential faults will appear, and become more difficult to find. If failures occur, it always accompany with the vibration noise, so useful fault characteristic information is submerged by noise. Vibration signals always carry a lot of information, so vibration-based signal processing technique is needed.

Traditional signal processing techniques include Fourier transform and short-time Fourier transform (STFT). However, these methods are based on the assumption that the process generating signals is stationary and linear [1]. Wavelet transform is used in machinery as a popular signal processing technique extensively, because it possesses good time and frequency localization properties [2-4]. But Wavelet transform also has some drawbacks, its essence is Fourier transform, the mechanical fault signals are non-stationary and non-linear mostly, so it can't deal with these questions perfectly [5]. On the other hand, Wavelet transform has wavelet base function, different functions lead to different analysis results. Therefore, Wavelet transform is not a self-adaptive signal processing technique.

These methods are suitable for stationary and linear signal processing. They are still based on the Fourier transform and lack the self-adaption for investigated signals. But rolling bearing vibration signal is always non-stationary and nonlinear. In order to deal with this type of signal, empirical mode decomposition, a new time-frequency analysis technique, has been widely used in fault diagnosis of rotating machinery. Because its good performance, it has been applied in gearbox fault diagnosis [6, 7], rotor fault diagnosis [8, 9], rolling bearing diagnosis [10-12]. EMD combined with Wigner-Ville Distribution (WVD) method is used in signal processing and has got a reasonable result [13]. Improved EMD algorithm is used in low frequency oscillations in power system [14]. EMD is based on the local characteristic time scales of a signal and could decompose the complicated signals into a finite and small number of intrinsic mode function (IMF) [15]. However, EMD also has drawbacks, the major drawbacks of EMD is

539

mode mixing problem, which is defined as a IMF consisting of disparate time scales, or a similar time scale exist in different IMFs. Mode mixing problem is caused by intermittency [16].

Ensemble empirical mode decomposition (EEMD), an improved technique of EMD, can alleviate the mode mixing problem, is proposed [16]. As a new noise assisted data analysis (NADA) method, EEMD can restrain mode mixing problem effectively. Because of the superiority, EEMD is applied in rotor fault diagnosis [1], gearbox fault diagnosis [5, 17] and so on. However, if rotating machinery faults occur and fault information is submerged by noise, using EEMD individually is hard to analyze the fault pattern and lead to useless results of diagnosis.

In this paper, a method called EEMD and distributing fitting testing for rolling bearing fault diagnosis is proposed. Fault signals always accompany with noise, in order to achieve denoised signals, after using EEMD method, distributing fitting testing is applied to choose useful IMFs and abandon noise IMFs. Then we choose denoised signals to obtain envelope spectrum, it can diagnose different failure frequency. Rolling bearing with inner-race faults and out-race faults are analyzed by the proposed method. The result is shown the method based on EEMD and distributing fitting testing obtain an ideal result.

2. Research Method

2.1. EMD Algorithm and Mode Mixing Analysis

EMD decompose complicated signals self-adaptively into a series of IMFs. An IMF is a function that satisfies two conditions [15]:

(1) In the whole data set, the number of extrema and the number of zero crossings must either equal or differ at most by one;

(2) At any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero.

We need to confirm all of the local extrema, then local maxima are connected by a cubic spline as the upper envelope, local minima are connected as the lower envelope. The major drawback of EMD is mode mixing problem. Mode mixing is caused by intermittency, when it occurs, IMF will lost its physical meaning. It not only leads to serious aliasing in the time-frequency distribution, but also renders the physical meaning of individual IMF unclear [16].

2.2. EEMD Algorithm and Anti-Aliasing Analysis

To alleviate mode mixing problem in EMD, EEMD is proposed. EEMD is a NADA method, which defines the true IMF components as the mean of an ensemble of trials plus white noise of finite amplitude to original signal; it can separate scale clearly [16]. The statistical properties of white noise demonstrated that noise could help data analysis in the EMD [18].

The principle of the EEMD is simple: the added white noise constitutes components of different scales and it would populate the whole time-frequency space uniformly. Of course, a noisy results for each individual trial would be produced. Noise is canceled out in the ensemble mean of enough trials, because each trial is different in separate trials. A large number of trials are added in the ensemble result in ensemble mean is considered as true answer. The EEMD principle advanced here is based on the following observations [16]:

(1) A collection of white noise cancels each other out in an ensemble; therefore, only the signal can survive and persist in the final noise-added signal ensemble mean.

(2) White noise is necessary to force the ensemble to discuss all possible solutions; the finite amplitude noise makes the different scale signals reside in the corresponding IMFs, and render the resulting ensemble mean more meaningful.

(3) The ensemble mean of lots of trials consisting of the noise-added signal is true and physically meaning of the EMD, not the one without noise.

Based on above mentioned, the EEMD algorithm can be described as follows [16,19]: (1) Given x(t) is an original signal, add a white noise signal $n_i(t)$ to x(t)

$$x_j(t) = x(t) + n_j(t) \tag{1}$$

 $x_i(t)$ is the noise-added signal, j = 1, 2, 3, ..., M, *M* is the number of ensemble.

(2) Decompose the $x_i(t)$ into a series of IMFs $c_{i,i}$ using the EMD method as follows:

$$x_{j}(t) = \sum_{i=1}^{N} c_{i,j} + r$$
(2)

 $c_{i,i}$ denotes the *i*th IMF of the *j*th trial, and N is the number of IMFs.

(3) If j < M, then repeat steps (1) and (2), but with different white noise series each time.

(4) Calculate the ensemble mean $\overline{c_i}$ of the *M* trials for each IMF

$$\overline{c_i} = \left(\sum_{j=1}^{M} c_{i,j}\right) / M \tag{3}$$

where i = 1, 2, 3, ..., N.

(5) Report the mean $\overline{c_i}$ of each of the *N* IMFs as the final IMFs.

From EMD algorithm and EEMD algorithm, we know EEMD is an improved technique of EMD, it not only reserves the advantages of EMD, but also alleviates mode mixing problem.

2.3. EEMD and Distributing Fitting Testing

As discussed by Wu and Huang et al. [19, 20], after white noise using the EMD method, the probability density function for each IMF is approximately normally distributed. To choose appropriate IMFs, normal probability plot is proposed based on distributing fitting testing. Normal probability plot theory can be stated as follows:

Supposing that distribution function F(x), variable X, testing H_0 :

$$X \sim N(\mu, \sigma^2), \ -\infty < \mu < +\infty, \ \sigma^2 > 0$$
(4)

If H_0 pass the testing, then

$$\frac{X-\mu}{\sigma} = U \sim N(0,1) \tag{5}$$

Standard normal distribution N(0,1) has its distribution function $\Phi(x)$, $\Phi(x)$ can replace

F(x)

$$F(x) = \Phi\left(\frac{X-\mu}{\sigma}\right) = \Phi(y)$$
(6)

and

$$y = \frac{1}{\sigma}(x - \mu) \tag{7}$$

In the *xoy* rectangular coordinate system, assuming that unit length between x-axis and y-axis are equal, equation $y = \frac{1}{\sigma}(x - \mu)$ represent a straight line, pass the point of $(\mu, 0)$, the slope in $\frac{1}{\sigma}$

is $\frac{1}{\sigma}$.

Then we update the coordinate system: x-axis don't change, F(x) become new y-axis, x-axis can be divided into x_1, x_2, \dots, x_n equally, y-axis is divided into $F(x_1), F(x_2), \dots, F(x_n)$ unequally, as result, to get normal probability plot, a straight line is formed because of connecting these points $[x_1, F(x_1)], [x_2, F(x_2)], \dots, [x_n, F(x_n)]$.

According to the distribution of the sample points, normal probability plot can judge whether these points follow normal distribution. If the distribution of the sample points is similar to straight line, they could be considered following normal distribution; otherwise, if the

Improved Ensemble Empirical Mode Decomposition for Rolling Bearing ... (Youpeng Zhang)

distribution of the sample points are same as curve or irregularity, they don't obey normal distribution.

In order to improve the validity of test, the Jarque-Bera test also is proposed based on method of test for skewness and kurtosis, it is the most universal normality test method now. Jarque-Bera test theory can be seen as follows [21, 22]:

The J-B test is define as

$$JB = \frac{n}{6} \left[S^2 + \frac{\left(K - 3\right)^2}{4} \right] \approx a\chi_2^2$$
(8)

Where *S*, *K*, n, χ_2^2 denotes skewness, kurtosis, the sample size and chi-square distribution with degrees of freedom of 2, respectively. For a normal distribution, sample skewness should be near 0 and sample kurtosis should be near 3. Jarque-Bera test often use the chi-square distribution to estimate critical values for large samples [23].

The noise IMFs test procedure can be summarized as follows:

(1) Each IMF should be sampled into a large sample size.

(2) Getting the significance level of 0.05 and utilizing the Jarque-Bera test.

(3) The test returns the value H=1 if it rejects samples is normal distribution, otherwise, the value H=0 if it accepts.

In this paper, combing normal probability plot and Jarque-Bera test to distinguish between noise IMFs and useful IMFs.

2.4. EEMD and Distributing Fitting Testing Algorithm

The EEMD and distributing fitting testing algorithm can be given as follows.

(1) EEMD decompose original signals into a series of IMFs.

(2) To obtain enough sample points, each IMF should be sampled, normal probability plot and Jarque-Bera test are used to examine them whether obey normal distribution.

(3) Abandon the IMF which conform to white noise features and plus the remainder IMFs to get de-noised signal.

(4) Get envelope spectrum from which the faults in a rolling bearing can be diagnosed.

The algorithm flow chat is given in Figure 1.

3. Results and Discussion

3.1. Simulation Experiment

3.1.1. Mode Mixing Simulation Experiment

To illustrate the mode mixing problem, a simulation signal s(t) is given in this section. The simulation signal, shown in Figure 2(c), is constituted by a sine wave of 7 Hz and Gauspuls impulses, the two functions are shown in Figure 2(a) and Figure 2(b), the signal has 1024 points. Applying the EMD to decompose signal s(t), the decomposition result is shown in Figure 3. It can be seen from Figure 3 that s(t) is decomposed into three IMFs and a residue. It is obvious that IMF1 and IMF2 are distorted seriously, mode mixing problem exist in these two IMFs. Gauspuls impulses can't be decomposed individually. IMF3 just represents the sine wave of 7 Hz, so the decomposed IMFs fail to represent the essence of s(t), this is mode mixing problem.

To expound the advantage of EEMD better, the simulation signal also is s(t), the decomposition result with EEMD is shown in Figure 4. Two parameters are set in EEMD, the number of ensemble and the amplitude of the added white noise. Because of the large amplitude of Gauspuls impulses, signal is decomposed using EEMD with ensemble number 100 and the added white noise 0.06 time standard deviation of the signal. It can be seen from Figure 4 that superposition of components c1 and c2 is Gauspuls impulses. The fact that the impulses signal actually resides in two EEMD components is due to the average spectra of neighboring IMFs of white noise overlapping, as revealed by Wu and Huang [15]. The sine wave of 7 Hz (c4) is represented nearly prefect. Therefore, the EEMD method can solve the mode mixing problem and obtain characteristic of signal s(t) accurately.



Figure 1. The flow chat of algorithm



Figure 2. The simulation signal s(t): (a) sine wave of 7 Hz, (b) Gauspuls impulses, (c) the simulation signal



Improved Ensemble Empirical Mode Decomposition for Rolling Bearing ... (Youpeng Zhang)

3.1.2. EEMD and Distributing Fitting Testing Simulation Experiment

To demonstrate the algorithm proposed former, a simulation signal $s_noise(t)$ is composed by s(t) and white noise which variance is 0.1. As can be seen from Figure 5, EEMD decomposition can get six IMFs and a remainder, to obtain the de-noised signal, each IMF should be sampled firstly, then distributing fitting testing is used to choose useful IMFs and abandon IMFs which represent white noise signal, at last the de-noised signal is available. The amplitude of reminder is small, so it can be abandoned. In the whole distributing fitting testing, significance level is $\alpha = 0.05$, H = 1 represents refusing hypothesis, H = 0 represents accepting hypothesis, the result for each IMF's normal probability plot can be seen from Figure 6. Under the significance level of 0.05, the *H* value of each IMF can be seen from Table 1.



Figure 5. The decomposition result of s_noise(t) with EEMD

Figure 6. The normal probability plots of IMFs of *s_noise*(*t*)



Figure 7. The de-noised signal of $s_noise(t)$

These results show that components c1, c2, c3 and c4 accept the hypothesis, they can be abandoned because of the characteristic of white noise, then c5 and c6 are added together, denoised signal can be got. The denoised signal is displayed in Figure 7, it indicates the sine wave of 7 Hz accurately; thus, the denoised method has achieved a perfect result.

3.2. Application to the Fault Diagnosis of Rolling Bearing 3.2.1. Inner Race Fault Diagnosis of Rolling Bearing

In this paper, the data of Case Western Reserve University Bearing Data Center is selected as a simulation object. The rolling bearing experimental setup is shown in Figure 8, it consists of motor, a torque transducer and a dynamometer. The test bearings support the motor shaft, it is located in motor driver end. Data sampling rate is 12000 Hz, motor speed is 1750 r/min, so rotation frequency is f = 29.17 Hz, according to the parameters of 6205-2RS SKF deep groove ball bearing, inner race failure frequency is $f_i = 157.9$ Hz. The proposed method based on EEMD and distributing fitting testing are applied to inner race fault diagnosis for rolling bearing. EEMD decomposition can be seen from Figure 9, there are seven IMFs and a remainder, the remainder can be abandoned because of small amplitude, the inner race fault signal $x_i(t)$ also is contaminated by white noise, so we need to select proper IMFs. There are a series of impulses in c1 and c2, it is close to 6.56 ms between the two impulses, which approximately equals inner race failure frequency (1/157.9 = 6.33 ms). Every IMF's normal probability plot and H value are presented in Figure 10 and Table. 2, it can be seen the distribution of the sample points in c1, c2, c3 and c7 are not similar to a straight line, and the value of H for them are equal to 1, combining with them can get the result, components c1, c2, c3 and c7 are chosen to recombine signal, it is shown in Figure 11. The de-noised signal not only reserves the fault features, but also the fault features are more obvious. There are abundant failure frequencies in envelope spectrum, which is shown in Figure 12. There are many relevant frequencies, $f_1 = 29.3$ Hz, $f_2 = 58.6$ Hz, $f_3 = 158.2$ Hz, $f_4 = 316.5$ Hz. Therefore, f_1 is rotation frequency, f_2 is twice of rotation frequency, f_3 has the largest amplitude, it is the inner race failure frequency, f_4 is the twice of inner race failure frequency.



Figure 8. The rolling bearing worktable

Table 2. The <i>H</i> value of each IMF of inner race fault signal							
Ci	c1	c2	c3	c4	c5	c6	c7
Н	1	1	1	0	0	0	1



Figure 9. The decomposition result with EEMD of inner race fault signal



Figure 11. The de-noised signal of inner race fault signal



Figure 10. The normal probability plots of IMFs of inner race fault signal



Figure 12. The envelope spectrum of inner race fault signal

3.2.2. Outer Race Fault Diagnosis of Rolling Bearing

The data also comes from Case Western Reserve University Bearing Database. Data sampling rate is 12000 Hz, motor speed is 1750 r/min, outer race failure frequency is $f_o = 104.57$ Hz. EEMD decomposition can be seen from Figure 13 and a series of impulses exist in IMFs which the period is 9.48 ms, it equals outer race failure frequency (1/104.57=9.56 ms). Every IMF's normal probability plot and *H* value are presented in Figure 14 and Table. 3, components c1, c4, c6 and c7 are chosen, the de-noised signal is shown in Figure 15. Then, the envelope spectrum is obtained and shown in Fig. 16. It also contain many frequencies, $f_1 = 11.7$ Hz, $f_2 = 29.3$ Hz, $f_3 = 105.5$ Hz, $f_4 = 211$ Hz. Thus, f_1 approximately equals cage fault frequency (11.6 Hz), f_2 is rotation frequency, f_3 is outer race failure frequency with the largest amplitude, f_4 is the twice of outer race failure frequency. Therefore, the proposed method is able to diagnose the faults in rolling bearing effectively.





Figure 13. The decomposition result with EEMD of outer race fault signal

Figure 14. The normal probability plots of IMFs of outer race fault signal



Figure 15. The de-noised signal of outer race fault signal



4. Conclusion

A satisfactory result has been got, the method called EEMD and distributing fitting testing are applied in rolling bearing diagnosis. It can recognize different rolling bearing failures effectively, such as inner race failure, outer race failure. Because of stochastic vibration process, EEMD can alleviate the mode mixing problem perfectly. In this paper, a simulation signal with white noise is decomposed by EEMD, in order to choose the appropriate IMFs, we need to distinguish the white noise components from real signal components, the distributing fitting testing is proposed. The satisfactory result has been got because of normal probability plot and Jarque-Bera test, using these two tests method can improve the correctness of the test, utilizing the IMFs distribution probability characteristic can choose useful IMFs effectively and

reserve the fault information very well. This method provide a thinking that how to choose suitable IMFs to denoise signal. [11] has got a correct result of rolling bearing diagnosis. Because of the shortcoming of EMD, the envelope spectrum is hard to recognize the rolling bearing fault frequency obviously. The ensemble empirical mode decomposition (EEMD) method regards as a tool for nonlinear and non-stationary signal analysis, it can alleviate the mode mixing problem. Aiming at choice of IMFs, the distributing fitting testing is used to recognize which IMF is needed. EEMD and distributing fitting testing method is used to diagnose rolling bearing faults in this paper. The analyzed results demonstrate that the proposed method is able to identify rolling bearing multi-fault diagnosis effectively.

Acknowledgments

This work is supported by Science and technology support program of Gansu Province (No. 1011JKCA172) and Science and technology program of Lanzhou (2011-1-106).

References

- [1] YG Lei, ZJ He, YY Zi. Application of the EEMD method to rotor fault diagnosis of rotating machinery. Mechanical Systems and Signal Processing. 2009; 23: 1327-1338.
- [2] NG Nikolaou, IA Antoniadis. Rolling element bearing fault diagnosis using wavelet packets. NDT&E International. 2002; 35: 197-205.
- [3] CJ Sheng, YD Jie, Y Yu. Application of an impulse response wavelet to fault diagnosis of rolling bearings. Mechanical Systems and Signal Processing. 2007; 21: 920-929.
- [4] WS Su, FT Wang, H Zhu, ZX Zhang, ZG Guo. Rolling element bearing faults diagnosis based on optimal Morlet wavelet filter and autocorrelation enhancement. Mechanical Systems and Signal Processing. 2010; 24: 1458-1472.
- [5] JS Lin, Q Chen. Application of the EEMD Method to Multiple Faults Diagnosis of Gearbox. Advanced Computer Control (ICACC), 2010 2nd International Conference on, Shenyang. 2010: 395-399.
- [6] SJ Loutridis. Damage detection in gear systems using empirical mode decomposition. Engineering Structures. 2004; 26: 1833-1841.
- [7] B Liu, S Riemenschneider, Y Xu. Gearbox fault diagnosis using empirical mode decomposition and Hilbert spectrum. Mechanical Systems and Signal Processing. 2006; 20: 718-734.
- [8] Q Gao, C Duan, H Fan, et al. Rotating machine fault diagnosis using empirical mode decomposition. Mechanical Systems and Signal Processing. 2008; 22: 1072-1081.
- [9] X Zhao, TH Patel, MJ Zuo. Multivariate EMD and full spectrum based condition monitoring for rotating machinery. Mechanical Systems and Signal Processing. 2012; 27: 712-728.
- [10] Y Yu, YD Jie, CJ Sheng. A roller bearing fault diagnosis method based on EMD energy entropy and ANN. Journal of Sound and Vibration. 2006; 294: 269-277.
- [11] DJ Yu, JS Cheng, Y Yang. Application of EMD method and Hilbert spectrum to the fault diagnosis of the roller bearings. Mechanical Systems and Signal Processing. 2005; 19: 259-270.
- [12] H Li, YP Zhang, HQ Zheng. Hilbert-Huang transform and marginal spectrum for detection and diagnosis of localized defects in roller bearings. Journal of Mechanical Science and Technology. 2009; 23: 291-301.
- [13] MZ Sun, HT Tang, YL Xu. Mixed Programming Realization of the EMD-WVD Combined Method. TELKOMNIKA Indonesian Journal of Electrical Engineering. 2013; 11(4): 1902-1908.
- [14] JB Yi, Q Huang, S Jing, LJ Ding. An Anti Mode Mode Mixing EMD Algorithm for Detecting the Characteristic of Low Frequency Oscillations in Power System. TELKOMNIKA Indonesia Journal of Electrical Engineering. 2013; 11(8): 4742-4750.
- [15] NE Huang, Z Shen, SR Long. The Empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings of the Royal Society of London. 1998; 454: 903-995.
- [16] ZH Wu, NE Huang. Ensemble empirical mode decomposition: a noise assisted data analysis method. Advances in Adaptive Data Analysis. 2009; 1: 1-41.
- [17] DC Cao, JS Kang, JM Zhao, XH Zhang. Fault Diagnosis of Gearbox Based on EEMD and HMM. Prognostics & System Health Management Conference, Beijing. 2012: MU3032.
- [18] P Flandrin, G Rilling, P Goncalvés. Empirical mode decomposition as a filter bank. IEEE Signal Processing Letter. 2004; 11: 112-114.
- [19] NE Huang, SSP Shen. Hilbert-Huang Transform and Its Applications, 4. World Scientific, Singapore, 2005
- [20] ZH Wu, NE Huang. A study of the characteristics of white noise using the empirical mode decomposition method. Proceedings of the Royal Society of London. 2004; 460A: 1597-1611.
- [21] E Besdok, ME Yüksel. Impulsive noise suppression from images with Jargue-Bera test based median filter. Int. J. Electron. Commun. (AEÜ). 2005; 59: 105-110.

548

- [22] CM JARQUE, AK BERA. Efficient tests for normality homoscedasticity and serial independence of regression residuals. *Economics Letters*. 1980; 6: 255-259.
- [23] MathWorks. MATLAB the language of technical computing. MATLAB function reference. New York: The MathWorks, Inc.; 2002.