Improved Ensemble Empirical Mode Decomposition and its Applications to Gearbox Fault Signal Processing

Jinshan Lin

School of Mechatronics and Vehicle Engineering, Weifang University
Weifang, 261061, China

Abstract

Ensemble empirical mode decomposition (EEMD) is a noise-assisted method and also a significant improvement on empirical mode decomposition (EMD). However, the EEMD method lacks a guide to choosing the appropriate amplitude of added noise and its computation efficiency is fairly low. To alleviate the problems of the EEMD method, the improved complementary EEMD method (ICEEMD) was proposed. Furthermore, the ICEEMD method was used to analyze realistic gearbox faulty signals. The results indicate that the ICEEMD method has some advantages over the EEMD method in alleviating the mode mixing and splitting as well as reducing the time cost and also outperforms the CEEMD method in alleviating the mode mixing and splitting. The paper also indicates that the ICEEMD method seems to be an effective and efficient method for processing gearbox fault signals.

Keywords: *Complementary Ensemble Empirical Mode Decomposition(CEEMD), Improved Complementary Ensemble Empirical Mode Decomposition(ICEEMD), Gearbox, Signal Processing.*

1. Introduction

It is a challenging task to develop signal processing techniques for non-stationary and noisy signals, which has attracted considerable attentions recently [1]. Many methods, such as short time frequency transform [2] and wavelet transform [3] , have been proposed for solving the problem and proved useful in some applications. However, because these methods usually need a priori knowledge about the researched signals, they naturally lack the self-adaption for the researched signals. The Wigner-Ville distribution has high time-frequency resolution, but its cross terms is unbearable [4]. Empirical mode decomposition (EMD) is a self-adaptive method and suitable to analyzing the non-stationary and nonlinear signals [4], which has been successfully applied to various fields [4]. Nevertheless, when the EMD algorithm is used to deal with a signal with intermittency, the mode mixing often emerges as an annoying problem [5-7]. To overcome the mode mixing problem, ensemble empirical mode decomposition (EEMD) is presented in place of EMD [8]. The EEMD method adds some white noise with limited amplitude to the researched signals, sufficiently taking advantage of the statistical characteristics of white noise whose energy density is uniformly distributed throughout the frequency domain, then projects the signal components onto the proper frequency bands and, finally, the added white noise can been counteracted by ensemble mean of enough corresponding components [8]. Therefore, the EEMD method is considered as a significant improvement on the EMD method and recommended as a substitute for the EMD method [8]. Indeed, the EEMD method has shown its superiority over the EMD method in some applications [9].

However, the EEMD method lacks a guide to how to choose the appropriate amplitude of the added noise and its computation efficiency is fairly low. As a result, the inappropriate amplitude of the added white noise for the EEMD method is going to cause the mode mixing and splitting that often exists in the EMD method [10, 11]. Although the reference [8] suggested that the amplitude of the added white noise should be about 0.2 times of the standard deviation of the investigated signal, unfortunately, with the suggested value, the decomposition results from the EEMD method often deviate from the realistic contents of the signals [11]. In addition, to further remove the residual of the added white noise and reduce a waste of time for the EEMD method, the complementary ensemble empirical mode decomposition (CEEMD) has been addressed to replace the EEMD method as a standard version of the EMD method [10]. Notwithstanding, the CEEMD method only partly enhances the computation efficiency of the EEMD method, and the above first problem regarding the EEMD method still remains untouched. Additionally, if the researched signal is a noisy signal in itself, its intrinsic noise will inevitably interact with the noise added through the EEMD method, which may further complicate the above first problem regarding the EEMD method. In particular, when the researched signals become very noisy, the above first problem regarding the EEMD method leaves a gap.

This paper explores the above two problems concerning the EEMD method. Then, the improved CEEMD (ICEEMD) method was proposed. Applications to analysis of defective gearbox signals proved the superiority of the ICEEMD method over the EEMD method.

2. The EMD and its Several Variations

2.1 The EMD method

The EMD method can self-adaptively decompose any a non-stationary and nonlinear signal into a set of intrinsic mode functions (IMFs) from high frequency to low frequency, which may be written as

$$x(t)=\sum_{i=1}^{N} c_i(t)+r_N(t) \tag{1}$$

where $c_i(t)$ indicates the ith IMF and $r_N(t)$ represents the residual of the signal $x(t)$. An IMF is a function which must satisfy the following two conditions: (1) the number of extrema and the number of zero crossings either equal to each other or differ at most by one, and (2) at any point, the local average of the upper envelope and the lower envelope is zero [8]. The residual $r_N(t)$ usually is a monotonic function or a constant.

2.2 The Ensemble EMD

An annoying problem associated with the EMD method is the mode mixing due to intermittency, defined as either a single IMF consisting of widely disparate scales, or a signal residing in different IMF components. To alleviate the imperfection of the EMD method, the ensemble EMD (EEMD), a noise-assisted method, is proposed. The EEMD method can be stated as follows:

$$x_m(t)=x(t)+w_m(t), \quad m=1,2,\cdots,N \tag{2}$$

$$x_m(t)=\sum_{i=1}^{L} c_{m,i}(t)+r_{m,L}(t), \quad m=1,2,\cdots,N \tag{3}$$

$$x_m(t)=\sum_{i=1}^{L} c_{m,i}(t)+r_{m,L}(t), \quad m=1,2,\cdots,N \tag{4}$$

where $x(t)$ is the original signal, $w_\mathrm{m}(t)$ is the mth added white noise, $x_m(t)$ is the noisy signal of the mth trial, $c_{m,i}(t)$ is the ith IMF of the mth trial, L is the number of IMFs from the EMD method, and N is the ensemble number of the EEMD method.

The EEMD method adds white noise with the finite amplitude to the signal, sufficiently taking advantage of the uniform statistic characteristics of white noise in the frequency domain, projects the different frequency signal components onto the corresponding frequency banks and, as a result, effectively overcomes the mode mixing due to the existence of intermittency [8]. Nonetheless, to totally clear the residual of the added white noise from the IMFs of the EEMD method, a large ensemble number is usually demanded, which will cause a tremendous waste of time.

2.3 The Complementary EEMD method

To better eliminate the residual of added white noise persisting in the IMFs of the EEMD method and raise the computation efficiency of the EEMD method, the complementary ensemble EMD (CEEMD) [10], a novel noise-enhanced method, is presented. The CEEMD method adds white noise in pairs with one positive and another negative to the original signal and then produces two sets of ensemble IMFs. Hence, two different combinations of the original signal and the added white noise can be obtained, i.e.

$$\begin{bmatrix} x_m^{+}(t) \\ x_m^{-}(t) \end{bmatrix}=\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} x(t) \\ w_m(t) \end{bmatrix}, \quad m=1,2,\cdots,N \tag{5}$$

where $x(t)$ is the original signal, $w_m(t)$ is the mth added white noise, $x_m^{+}(t)$ is the sum of the original $x(t)$ and the mth added white noise $w_m(t)$, and $x_m^{-}(t)$ is the difference of the original $x(t)$ and the mth added white noise $w_m(t)$. In light of (3) and (4), the original signal $x(t)$ can be expressed as

$$x(t)=\frac{1}{2N}\sum_{i=1}^{L}\sum_{m=1}^{N}\left(c^{+}_{m,i}(t)+c^{-}_{m,i}(t)\right)+\frac{1}{2N}\sum_{m=1}^{N}\left(r^{+}_{m,L}(t)+r^{-}_{m,L}(t)\right) \tag{6}$$

where $c^{+}_{m,i}(t)$ is the ith IMF of $x_m^{+}(t)$ and $c^{-}_{m,i}(t)$ is the ith IMF of $x_m^{-}(t)$. Although the CEEMD method appears to be able to completely remove the residual of the added white noise persisting in the IMFs of the EEMD method, it still remains unsolved how to choose the appropriate amplitude of the added white noise for the EEMD method.

3. The Improved CEEMD Method

3.1 The Choice of Amplitude of the Added Noise

The amplitude of the added white noise is a key parameter of the EEMD method, which will exert a decisive impact on whether or not the EEMD method can yield the reasonable decomposition results. If the added noise is too weak to bring the changes of extrema of the original signal, the EEMD method will degenerate into the EMD method. Conversely, if the added noise is too strong to reveal the original signal, the EEMD method will derive meaningless results which are mainly controlled by the added noise and scarcely associated with the original signal [11], whether or not the ensemble number is large enough. The reference [11] demonstrated that the decomposition results of the EEMD method varied with the different amplitude of the added noise and considered it appropriate for the EEMD method to set SNR in the range of 50-60 dB. In fact, for the simple simulated example by which the conclusions were drawn in [11], when the amplitude of the added noise is set as 0.01, which is considered an optimal choice by [11], the SNR between the original signal and the added noise is approximately 37 dB which is outside the range of 50-60 dB. Accordingly, the conclusions given by [11], relating to the choice of the amplitude of the added white noise of the EEMD method, is not entirely dependable. Then, the problem is further investigated in this paper using two simulated signals. Thus, the Pearson's correlation coefficient (PCC) is used as a parameter to measure the performance of the EEMD method with different amplitude of the added white noise. Here, the ensemble number of the EEMD method is set as 100.

First, a simple noiseless simulated signal was used to examine the choice of the amplitude of the added white noise for the EEMD method. The signal is a combination of a low-frequency sinusoid component $x_1(t)$ and a high-frequency damped transient component $x_2(t)$, shown in Fig. 1, and its formula can refer to [11]. The relationship between the PCCs and the amplitude of the added white noise is illustrated in Fig. 2. As shown in Fig. 2, when the amplitude of the added white noise is 0.0063, the two PCCs almost simultaneously reach their maximum values. Table 1 exhibits the average powers of two components of the signal and the square roots of the average powers. As seen in Table 1, the value of 0.0063 just equals to the square root of the average power of the weak transient component $x_1(t)$. As a result, the square root of the average power of the weak transient component apparently approximates the optimal amplitude of the added white noise.

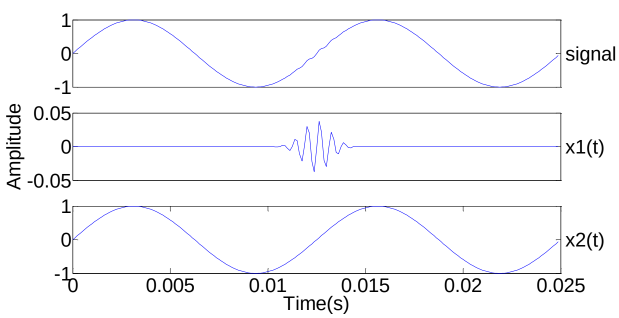

Fig. 1 A simple noiseless simulated signal and two components.

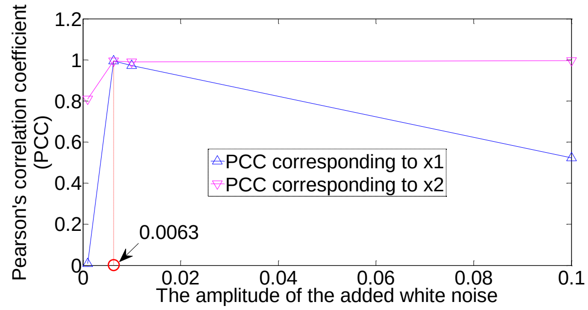

Fig. 2 The relationship between the Pearson's correlation coefficient (PCC) and the amplitude of the added white noise for the simple signal.

Table 1: The average powers of two components of the simple noiseless simulated signal and the square roots of the average powers

Parameter	***The two components of the simple signal***	
	$x_1(t)$	$x_2(t)$
The mean power	4.0106×10^{-5}	0.5
The square root of the mean power	0.0063	0.7071

Subsequently, a complex noiseless simulated signal was utilized to further verify the conclusion. The signal consisting of four components imitates realistic vibration signals of a rolling bearing, shown in Fig. 3, and its formula can refer to [11]. The relationship between the PCCs and the amplitude of the added white noise is illustrated in Fig. 4. As shown in Fig. 4, when the amplitude of the added white noise lies in the range of 0.0085-0.0138, the four PCCs almost simultaneously reach their maximum values. Table 2 depicts the average powers of four components of the signal and the square roots of the average powers. As shown in Table 2, the value of 0.0085 is just equal to the square root of the average power of the weak sinusoid component $x_4(t)$ and the value of 0.0138 is just equal to the square root of the average power of the weak transient component $x_1(t)$. More generally, Fig. 4 indicates that an optimal internal of the amplitude of the added white noise for the EEMD method may lie between the square root of the average power of the weak sinusoid component and the square root of the average power of the weak transient component, which is in accordance with the conclusions drawn from the above simple example.

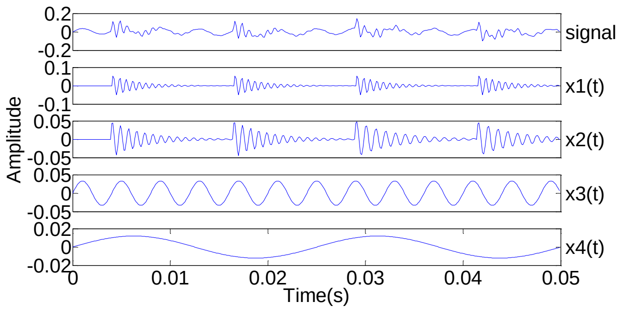

Fig. 3 A complexly-simulated signal and its four components.

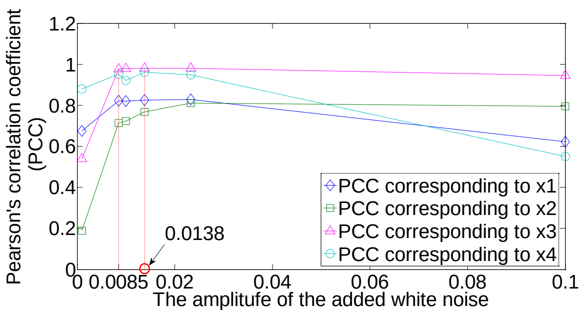

Fig. 4 The relationship between the Pearson's correlation coefficient (PCC) and the amplitude of the added white noise in a complex noiseless simulated signal.

Table 2: The average powers of four components of the complex noiseless simulated signal and the square roots of the average powers

Parameter	***The four components of the complex signal***			
	$x_1(t)$	$x_2(t)$	$x_3(t)$	$x_4(t)$
The mean power	1.9032	1.9377	5.4450	0.7200
The square root of the mean power	0.0138	0.0139	0.0233	0.0085

3.2 The improved CEEMD

On the premise that the noise is neglected, the above section gives an optimal interval of the amplitude of the added white noise for the EEMD method. Actually, a realistic signal is usually contaminated by strong or weak noise. When the EEMD method is applied to a noisy signal, the intrinsic noise will inevitably interact with the added noise, which can make a great impact on how much extrinsic noise should be added. Although the reference [8] has suggested that the amplitude of the added white noise should be about 0.2 times of the standard deviation of the investigated signal, the conclusion is pretty rough and inappropriate in many cases [11]. To alleviate the problem existing in the EEMD method for analyzing a noisy signal, the improved CEEMD (ICEEMD) is proposed. First, the noisy signal is roughly decomposed using the CEEMD algorithm. Then, both the weak transient component and the weak sinusoid component are obtained. As a result, the optimal interval of the amplitude of the added white noise can be determined. In the end, with the amplitude of the added white noise lying in the optimal interval obtained in the previous step, the CEEMD method is again performed.

4. Experiment verification

To further assess its performance, the ICEEMD method was exploited to examine the gearbox vibration data provided by Kayvan J. Rafiee [12]. The gearbox was running by the driving gear meshing with the driven one. The rotation speed of the input shaft is 24.05Hz, the rotation speed of the output shaft is 29.06Hz and the meshing frequency is 697.5Hz [12]. The signals were measured from the driving gear. The normal gearbox signal and the broken-tooth gearbox signal are shown in Fig. 5. Subsequently, the EEMD method, the CEEMD method and the ICEEMD method were utilized to explore the two signals, and the corresponding HHT spectra are shown in Fig. 6, Fig. 7 and Fig. 8, respectively. As shown in Fig. 6(a), Fig. 7(a) and Fig. 8(a), there are no obvious periodic characteristics in the HHT spectra of the normal gearbox signal; conversely, as shown in Fig. 6(b), Fig. 7(b) and Fig. 8(b), there are obvious periodic characteristics in the HHT spectra of the broken-tooth gearbox signal. However, as seen in Fig. 6(b) and Fig. 7(b), none of other explicit information can be found, in addition to the frequency bands scattered nearly throughout the frequency range only at the instant when the shocks happen, which implies that there occurs the mode mixing or splitting in the two methods because of the inappropriate amplitude of the added noise. Instead, as seen in Fig. 8(b), in addition to the frequency bands scattered nearly throughout the frequency range only at the instant when the shocks happen, there is another an instantaneous frequency curve similar to a cosine curve (highlighted with the red curve) with the modulation frequency of 24Hz and the carrier frequency of about 4800Hz, where the frequency 24Hz almost equals to the rotation speed of the input shaft and the frequency 4800Hz approaches the twelve times of the meshing frequency, which denotes that there occurs a frequency modulation phenomenon in the broken-tooth signal. Consequently, the comparisons between the three HHT spectra from Fig. 6(b), Fig. 7(b) and Fig. 8(b) prove that the ICEEMD method greatly alleviates the mode mixing and splitting of the EEMD/CEEMD method and can extract more and useful information from the faulty signals, which is essential for fault diagnosis of gearboxes. In addition, Fig. 9 presents that ICEEMD, comparable to CEEMD, can reduce a waste of time by 80% compared with the EEMD method. This indicates that the ICEEMD method is seemingly an effective and efficient method for gearbox fault signal processing.

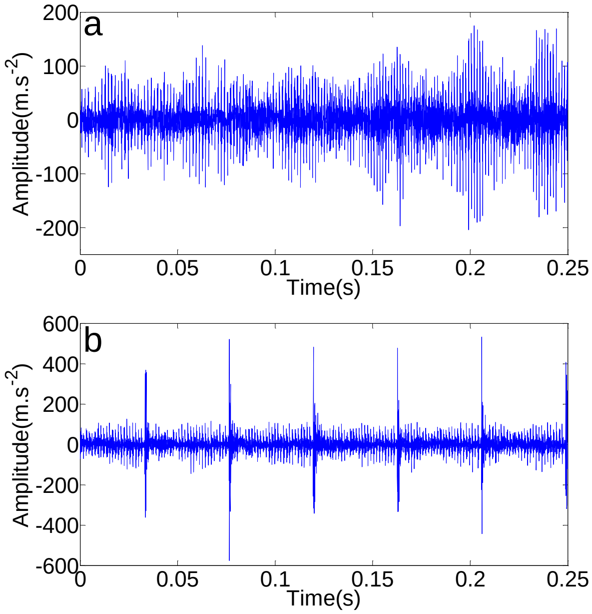

Fig. 5 The two gearbox vibration signals: (a) The normal gearbox signal; (b) The broken-tooth gearbox signal.

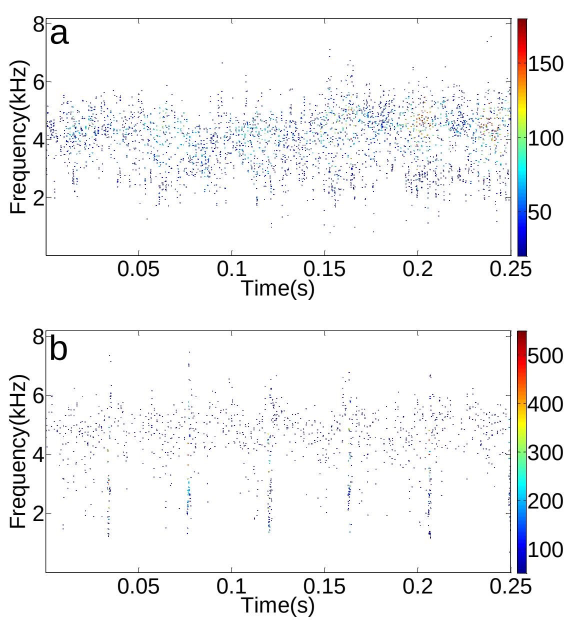

Fig. 6 The HHT spectra of the two vibration signals using the EEMD method: (a) The normal gearbox signal; (b) The broken-tooth gearbox signal.

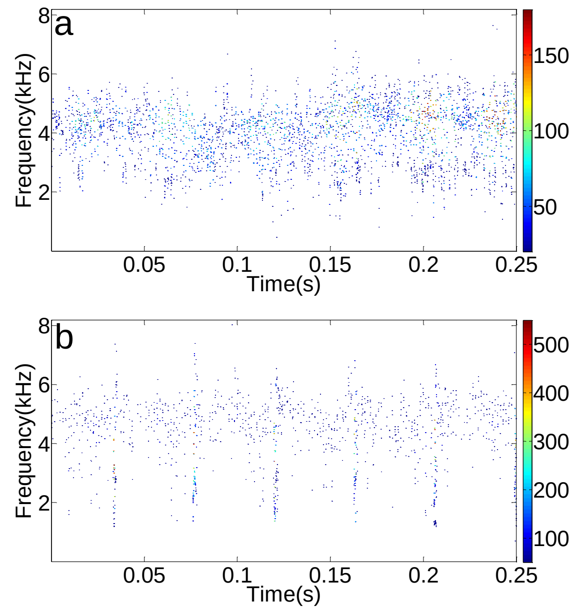

Fig. 7 The HHT spectra of the two vibration signals using the CEEMD method: (a) The normal gearbox signal; (b) The broken-tooth gearbox signal.

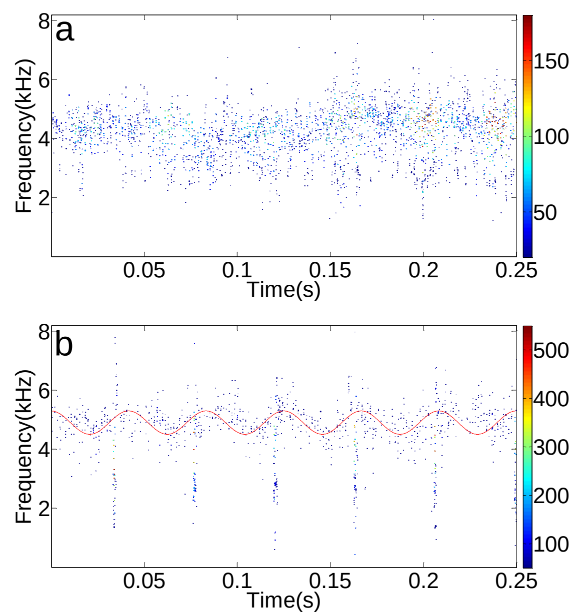

Fig. 8 The HHT spectra of the two vibration signals using the ICEEMD method (The red cosine curve is added to highlight the instantaneous frequency curve.): (a) The normal gearbox signal; (b) The broken-tooth gearbox signal.

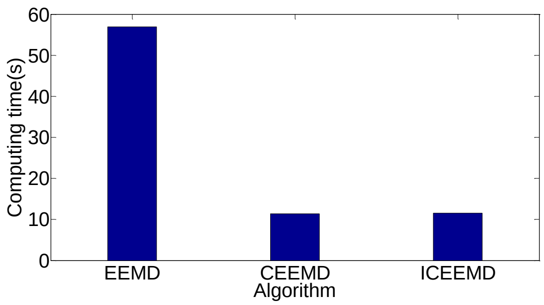

Fig. 9 Comparisons of computing time between the three different methods for the broken-tooth signal.

5. Conclusions

The paper aims to provide guidance on choosing the appropriate amplitude of the added white noise for the EEMD method and reduce the tremendous time waste occurring in the EEMD method. To solve those problems, based on the CEEMD method, the ICEEMD method is addressed in this paper. Besides, the numerical examples and the experimental examples have tested the ability of the ICEEMD method. The comparisons with the EEMD and CEEMD methods show that the ICEEMD method outperforms the EEMD method in alleviating the mode mixing and splitting as well as reducing the time waste and also outperforms the CEEMD method in alleviating the mode mixing and splitting. This paper indicates that the ICEEMD method is seemingly an effective and efficient method for gearbox fault signal processing. In addition, combined with some other methods[13, 14], the ICEEMD method may achieve better results in analyzing gearbox faulty signals.

Acknowledgments

The author thanks Kayvan J. Rafiee for his generous providing free gearbox vibration dataset on his personal official webpage. The work is supported by Natural Science Fund of Shandong Province (ZR2012EEL07), Development Program of Science and Technology of Weifang City (201104050, 201104063) and Young Science Fund of Weifang University (2012Z20).

References

[1] W. Bartelmus, R. Zimroz, "A new feature for monitoring the condition of gearboxes in non-stationary operating conditions", Mechanical Systems and Signal Processing, Vol. 23, No. 5, 2009, pp. 1528-1534.
[2] L. Satish, "Short-time Fourier and wavelet transforms for fault detection in power transformers during impulse tests", Proceedings of the Institute of Electrical Engneering–Science, Measurement and Technology, Vol. 145, No 2. 2002, pp. 77-84.
[3] X. Jiang, S. Mahadevan, "Wavelet spectrum analysis approach to model validation of dynamic systems", Mechanical Systems and Signal Processing, Vol. 25, No 2, 2010, pp. 575-590.
[4] N.E. Huang, Z. Shen, S.R. Long, et al, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis", Proceedings of the Royal Society of London Series A - Mathematical Physical and Engineering Sciences, Vol. 454, No. 1971, 1998, pp. 903-995.
[5] R. Ricci, P. Pennacchi, "Diagnostics of gear faults based on EMD and automatic selection of intrinsic mode functions", Mechanical Systems and Signal Processing, Vol. 25, No. 3, 2011, pp. 821-838.
[6] J. Cheng, D. Yu, J. Tang, Y. Yang, "Application of frequency family separation method based upon EMD and local Hilbert energy spectrum method to gear fault diagnosis", Mechanism and Machine Theory, Vol. 43, No. 6, 2008, pp. 712-723.
[7] L. Lin, J. Hongbing, "Signal feature extraction based on an improved EMD method", Measurement, Vol. 42, No. 5, 2009, pp. 796-803.
[8] Z.H. Wu, N.E. Huang, "Ensemble empirical mode decomposition: A noise-assisted data analysis method", Advances in Adaptive Data Analysis, Vol. 1, No. 1, 2009, pp. 1-41.
[9] Y. Lei, Z. He, Y. Zi, "Application of the EEMD method to rotor fault diagnosis of rotating machinery", Mechanical Systems and Signal Processing, Vol. 23, No. 4, 2009, pp. 1327-1338.
[10] J.I.A.R. Yeh, J.S. Shieh, N.E. Huang, et al, "Complementary ensemble empirical mode decomposition: A novel noise enhanced data analysis method", Advances in Adaptive Data Analysis, Vol. 2, No. 2, 2010, pp. 135-156.
[11] J. Zhang, R. Yan, R.X. Gao, Z. Feng, "Performance enhancement of ensemble empirical mode decomposition", Mechanical Systems and Signal Processing, Vol. 24, No. 7, 2010, pp. 2104-2123.
[12] J. Rafiee, P. Tse, "Use of autocorrelation of wavelet coefficients for fault diagnosis", Mechanical Systems and Signal Processing, Vol. 23, No. 5, 2009, pp. 1554-1572.
[13] Liyong Ma, Naizhang Feng, Qi Wang, Non-Negative Matrix Factorization and Support Vector Data Description Based One Class Classification, International Journal of Computer Science Issues, Vol. 9, No. 5, 2012, pp. 36-42.
[14] Hao Zhao, Application of Wavelet De-noising in Vibration Torque Measurement, International Journal of Computer Science Issues, Vol. 9, No. 5, 2012, pp. 29-33.

Jinshan Lin received his BE degree from Shandong University in Mechanical Engineering and Automation in 2000. Next, he received his ME degree from University of Jinan in Mechanical and Electronic Engineering in 2006. Since 2006, he have taught at Weifang University. Currently, his research interests include fault detection, pattern classification and signal processing.