Research Article

Gearbox Fault Diagnosis Using Complementary Ensemble Empirical Mode Decomposition and Permutation Entropy

Liye Zhao,[1,2] Wei Yu,[1,2] and Ruqiang Yan[1]

[1]*School of Instrument Science and Engineering, Southeast University, Nanjing 210096, China*
[2]*Key Laboratory of Micro Inertial Instrument and Advanced Navigation Technology, Ministry of Education, Nanjing 210096, China*

Correspondence should be addressed to Ruqiang Yan; ruqiang@seu.edu.cn

This paper presents an improved gearbox fault diagnosis approach by integrating complementary ensemble empirical mode decomposition (CEEMD) with permutation entropy (PE). The presented approach identifies faults appearing in a gearbox system based on PE values calculated from selected intrinsic mode functions (IMFs) of vibration signals decomposed by CEEMD. Specifically, CEEMD is first used to decompose vibration signals characterizing various defect severities into a series of IMFs. Then, filtered vibration signals are obtained from appropriate selection of IMFs, and correlation coefficients between the filtered signal and each IMF are used as the basis for useful IMFs selection. Subsequently, PE values of those selected IMFs are utilized as input features to a support vector machine (SVM) classifier for characterizing the defect severity of a gearbox. Case study conducted on a gearbox system indicates the effectiveness of the proposed approach for identifying the gearbox faults.

1. Introduction

Gears can be considered as significant subassembly in machines for power or rotation transmission from one shaft to another. Their fault may cause unexpected breakdown of the machine systems and lead to significant economic loss or even personnel casualties [1, 2]. Since structural defect-caused vibration signals often reflect changes of the dynamic characteristics related to the gearbox, many researches focus on transient feature extraction of the vibration signal and fault recognition of the defective gearboxes using vibration signal analysis [3]. Nonetheless, a number of factors related to structural transformation, friction, velocity shear, and strike affect the vibration-oriented signal study and reduce the effectiveness of defective diagnosis. Consequently, a number of conventional linear approaches might not operate well in detection of dynamic changes [4, 5].

Aiming at avoiding restrictions of conventional techniques, permutation entropy (PE) is used to characterize vibration signals for the purpose of fault diagnosis. The PE only uses the order of entropy for signal characterization and can overcome nonlinear distortion which existed in the signal. It has been applied in various applications. For instance, permutation entropy is proved to offer an efficient evaluation to monitor rolling bearings [5]. By integrating support vector machine (SVM) with multiscale PE, the operating condition of rolling bearing can be identified [6, 7]. Another study combined optimized SVM, ensemble empirical mode decomposition, and PE to detect and classify motor bearing faults [8]. The effectiveness of the PE has also been proved in detecting dynamic changes in rotating machines when comparing with that of other features, like Lyapunov exponent and fractal dimensions [9]. Furthermore, background noise which existed in real world applications always disturbs the result of the fault diagnosis. Therefore, performing noise reduction in the vibration signals is necessary before the PE method is executed.

Empirical mode decomposition (EMD), as an approach of adaptive signal treatment in the field of time frequency analysis, can decompose a signal into sets of intrinsic mode functions (IMFs) based on its features [10, 11]. The approach of EMD could be applied in pretreatment of the signals of vibration. For instance, a novel approach for extracting fault feature with combined AR model and EMD algorithm has

been applied in processing bearing vibration signals [5]. Nonetheless, when the EMD is used to process a signal with intermittent components, the signal could not be fully decomposed because of the existence of mode mixture phenomenon [12]. Further development with ensemble empirical mode decomposition (EEMD) was proposed by Huang et al. for EMD performance improvement. EEMD becomes more precise and efficient for decomposition of signals in comparison with the original EMD by adding noise to the original signal and continuously calculating the IMF means [12]. Though the approach of EEMD has efficiently solved the issue of mode-mixing, it takes lots of time to implement the large amount of ensemble mean. In other words, the efficiency of algorithm will be decreased. In order to resolve this issue, the complementary approach of EEMD (CEEMD) has been put forward [13]. Through complementary integration of IMFs and both positive and negative added white noises in the CEEMD, residual of the noises could be extracted out from the combination of white noises and data. The approach of CEEMD demonstrates similar effectiveness to that of the EEMD with improved computational efficiency.

By making full use of characteristics of the PE and CEEMD, this paper proposes a hybrid approach to diagnose gearbox faults. The CEEMD is utilized as the preprocessing to filter signals and extract IMFs that are closely associated with the filtered signal. Subsequently, PE value of each chosen IMF would be calculated. The PE value of the chosen IMFs is utilized as the feature vector to a classifier in which the support vector machine (SVM) is applied for identifying gearbox defect. The remaining parts of this paper are arranged as follows. Overview of the gearbox fault diagnosis approach is shown in Section 2. Experimental verification is conducted on automobile transmission gearbox system in Section 3. Last but not least, the last section presents the summary and comment.

2. Theoretical Framework

2.1. Complementary Ensemble Empirical Mode Decomposition. CEEMD is developed based upon EEMD. Originally, the EMD approach deals with a given signal $x_i(t)$ into the form presented in (1) through recursive elimination of the mean of the lower and upper envelope related to the maximum and minimum of the signal [14]:

$$x_i(t) = \sum_{i=1}^{N} C_{i,N}(t) + r_{i,n}(t),\tag{1}$$

where N refers to the number of IMFs, $C_{i,N}(t)$ refers to the component of IMF which covers a certain frequency band, and $r_{i,n}(t)$ refers to the mean trend of the signal residue. The EMD can be considered as adaptive local analysis approach for processing both nonlinear and nonstationary signals. However, the decomposition of EMD would generally undergo mixture of modes, which is defined as either a single IMF covering widely disparate scales or a signal existing in different IMF components.

Later, Huang et al. have proposed a noise-guided statistical approach to resolve the mode mixture issue, which is
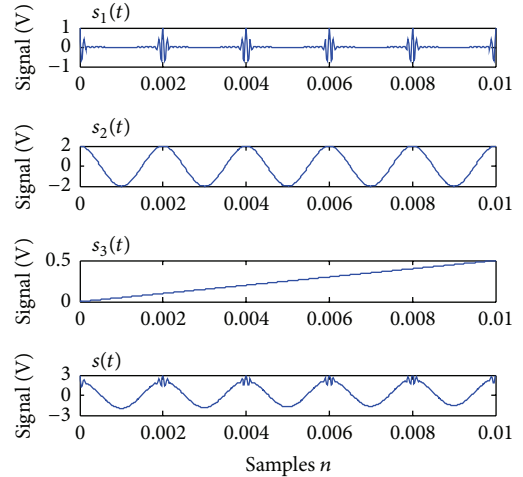

Figure 1: The simulated signal.

the ensemble empirical mode decomposition. However, the effect of the additional noise could only be restricted by a large amount of ensemble mean computation, causing high computational load.

Complementary ensemble mode decomposition, as an improved and noise enhanced data analysis approach, has been developed for reducing computational burden [13]. The procedure of CEEMD for the signal $x(t)$ is illustrated in the following steps.

Step 1. A pair of white Gaussian noises with the same amplitude is added to $x(t)$. Thus, two signals, $x_1 = x + x_n$ and $x_2 = x - x_n$, are generated.

Step 2. Decompose x_1 and x_2 by EMD for a number of times; then IMF_{x_1} referring to ensemble means of IMF from x_1 and IMF_{x_2} referring to those from x_2 are obtained.

Step 3. The final IMF which is the ensemble of IMF_{x_1} and IMF_{x_2} is calculated as the decomposition results of CEEMD as follows:

$$\text{IMF} = \frac{\left(\text{IMF}_{x_1} + \text{IMF}_{x_2}\right)}{2}.\tag{2}$$

Specifically, a simulated signal $s(t)$ composed of $s_1(t)$, $s_2(t)$, and $s_3(t)$ has been adopted as an instance. $s_1(t)$ is a Gaussian impulse interference signal, $s_2(t)$ is a cosine signal with the frequency of 500 Hz, and $s_3(t)$ is a trend term. Figure 1 illustrates the waveform of the simulated signal and Figure 2 illustrates the decomposed results by CEEMD.

Through comparing the result in Figure 2 with the signal waveforms in Figure 1, it is shown that there is no mode mixture. That is to say, CEEMD is more suitable for the study of signal.

2.2. Permutation Entropy. PE is a nonlinear dynamic parameter that characterizes a signal's complexity. Based on the
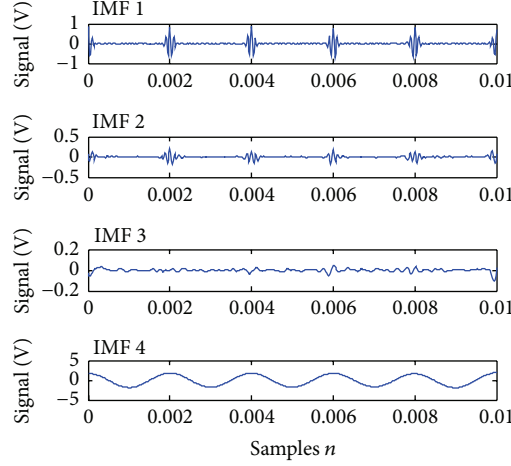
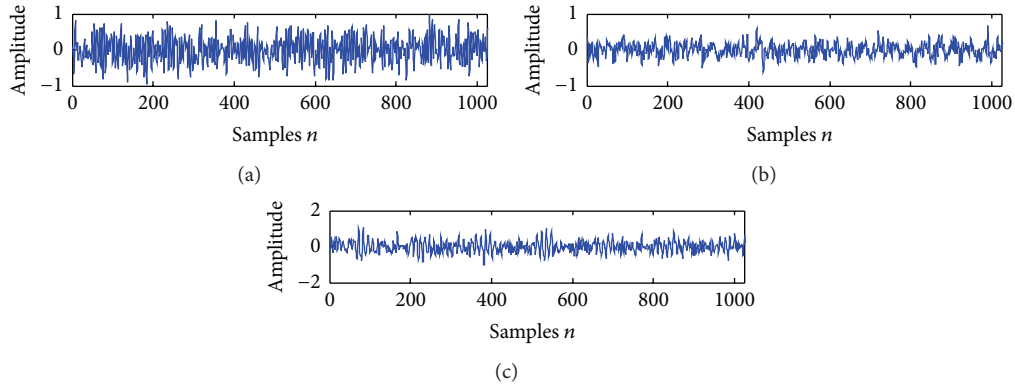Figure 2: The decomposition result by CEEMD.

(a)

(b)

(c)

Figure 3: Gearbox vibration signals under various operating conditions.

principle of Takens-Maine, the phase space of time series $\{y(i), i = 1, 2, \ldots, N\}$ can be expressed as

$$Y(i) = \{y(i), y(i+\tau), \ldots, y(i+(m-1)\tau)\}, \\ i = 1, 2, \ldots, N - (m-1)\tau, \tag{3}$$

where m refers to the embedded dimension, while τ refers to the delay of time. Furthermore, m sample points of data contained in every $Y(i)$ could be sorted in an incremental order as

$$\{y(i+(j_1-1)\tau) \le y(i+(j_2-1)\tau) \le \cdots \\ \le y(i+(j_m-1)\tau)\}. \tag{4}$$

If $y(i+(j_1-1)\tau) = y(i+(j_2-1)\tau)$, the original positions could be classified as $j_1 \le j_2$, $y(i+(j_1-1)\tau) \le y(i+(j_2-1)\tau)$. Thus, vector $Y(i)$ could be shown in a set of symbols as [15, 16]

$$T(l) = (j_1, j_2, \ldots, j_m), \tag{5}$$

where $l = 1, 2, \ldots, k$ and $k \le m!$. $T(l)$ refers to $m!$ symbol permutation which has been shown in m number symbols $(j_1, j_2, \ldots, j_m)$. If $P_1, P_2, \ldots, P_k$ are applied in denoting the possibility distribution of each symbol sequence and

$\sum_{j=1}^{J} P_j = 1$, the permutation entropy of m for the time series of $\{y(i), i = 1, 2, \ldots, N\}$ could be considered as the entropy of Shannon for k symbol sequence as follows:

$$H_{\text{PE}}(m) = -\sum_{j=1}^{J} P_j \ln P_j. \tag{6}$$

If all the symbol sequences appear with the same possibility distribution as $P_j = 1/m!$, the maximum value of $H_{\text{PE}}(m)$ could be described as $\ln(m!)$. Thus, the permutation entropy of order m can be standardized as

$$0 \le H_{\text{PE}} = \frac{H_{\text{PE}}(m)}{\ln(m!)} \le 1. \tag{7}$$

H_{PE} value shows the randomness level of the time series. A large value of H_{PE} indicates high randomness of the time series. On the contrary, a small value of H_{PE} means the time series has more regular characteristics.

To demonstrate the validity of the PE algorithm, sample vibration signals of a gearbox under three different conditions are shown in Figure 3, and the corresponding single factor analysis result is shown in Figure 4. Figure 4 shows that defect severity of the gearbox could be efficiently recognized by the value of PE.
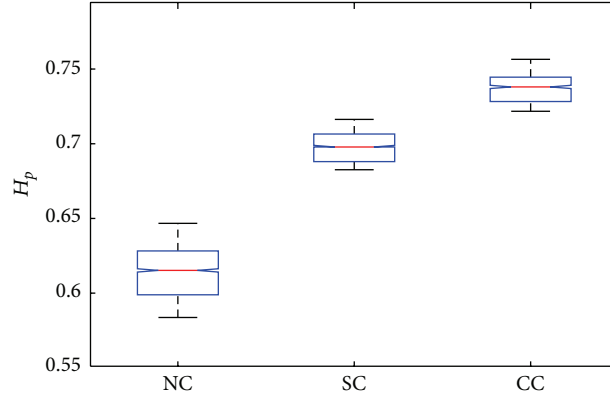
FIGURE 4: Boxplot of PE values on normal condition (NC), slight fault condition (SC), and catastrophe fault condition (CC).
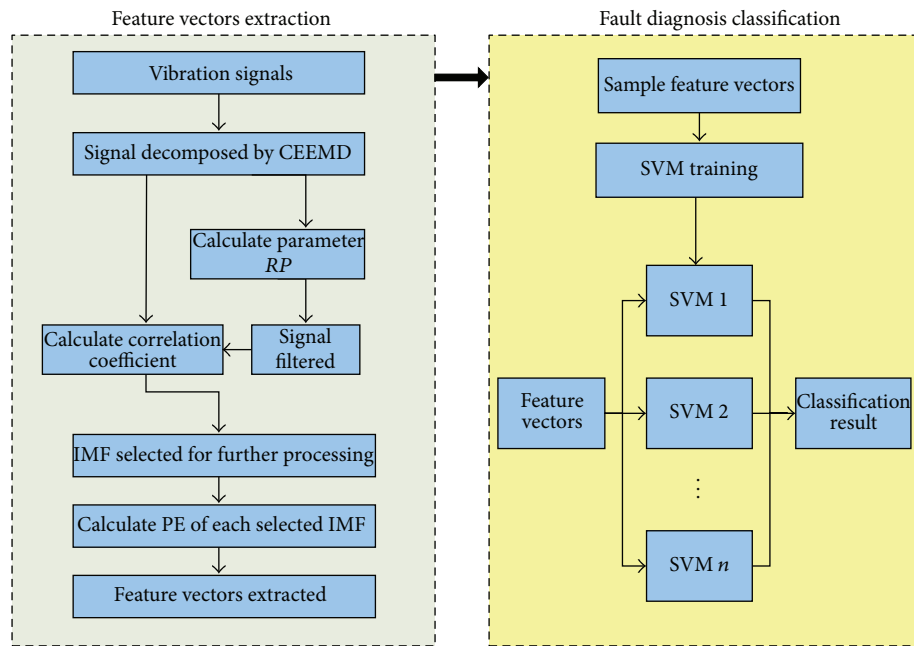

FIGURE 5: Flow chart of the proposed method.

2.3. Fault Diagnosis Based on CEEMD and Permutation Entropy. In this study, a gearbox fault diagnosis method has been developed using the CEEMD and PE, and Figure 5 shows the flow chart of the method. Particularly, the procedure to implement the proposed fault diagnosis method is as follows.

Step 1. The sampled vibration signal measured on gearbox is decomposed using CEEMD.

Step 2. The product P_j is calculated using (8), and the parameter RP_j is calculated by (9). The signal is filtered through comparison of the proposed threshold value and the parameter RP_j [17]. In other words, when $RP_j \geqslant 1$, it can be assured that P_j of the jth IMF can be enhanced for a number of times in comparison with the mean value of P_j which can be calculated based on the former $j - 1$ IMFs. Thus, the previous $j - 1$ IMFs with the term of trend can be eliminated

as noise and the residue IMFs can be considered as filtered signal:

$$P_j = E_j \times T_j, \tag{8}$$

$$RP_j = \left| \frac{P_j - (1/(j-1)) \sum_{i=1}^{j-1} P_j}{(1/(j-1)) \sum_{i=1}^{j-1} P_j} \right| \quad (j \geq 2), \tag{9}$$

where $E_j = (1/N) \sum_{i=1}^{N} [A_j(i)]^2$ refers to the jth IMF's energy density, $T_j = 2N/O_j$ refers to the mean period of the jth IMF, N represents the length of each IMF, A_j denotes the jth IMF's amplitude, and O_j refers to the overall number of extreme points in the jth IMF.

Step 3. The correlation coefficients between each IMF and filtered signal are calculated by (10). IMFs closely associated
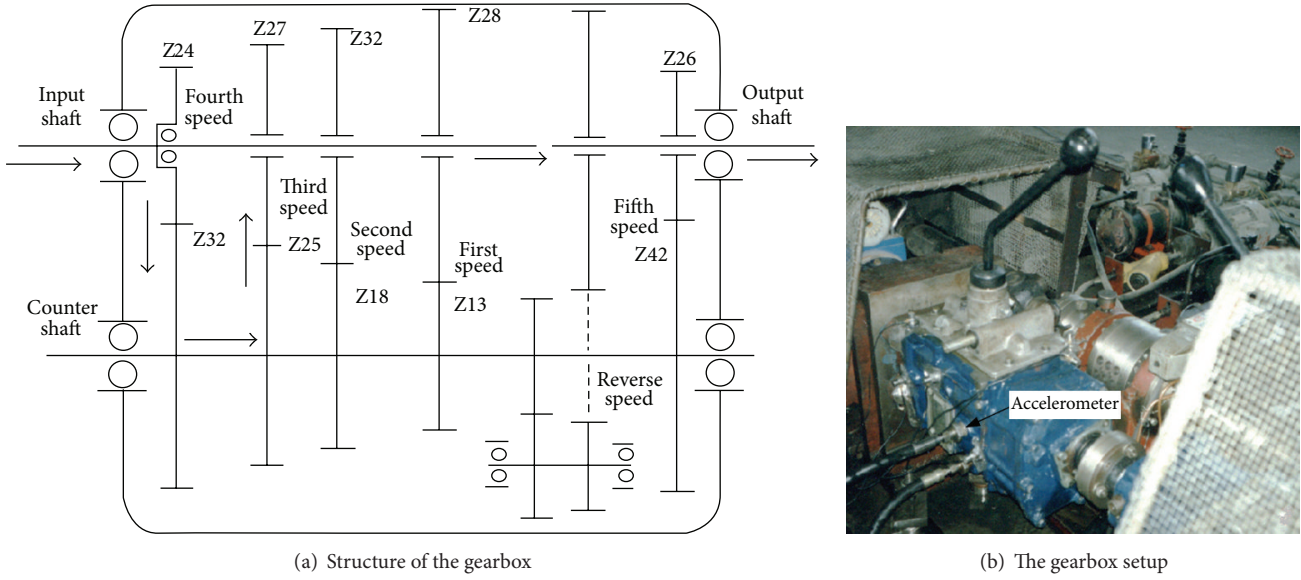
(a) Structure of the gearbox

(b) The gearbox setup

FIGURE 6: The automobile transmission gearbox.

(a) Normal condition

(b) Light fault condition
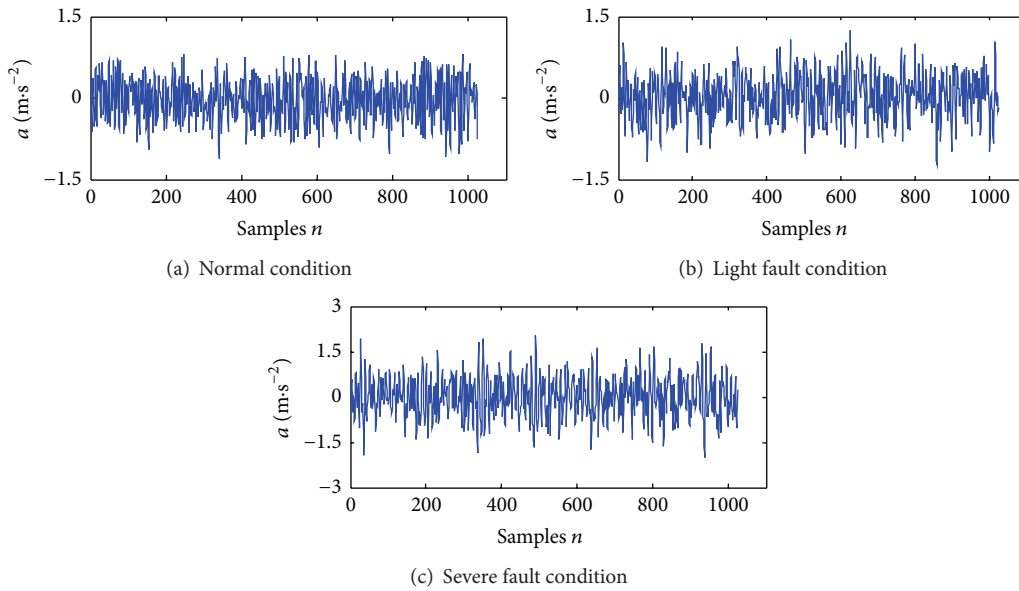

(c) Severe fault condition

FIGURE 7: Vibration signal waveforms of the gearbox under different conditions.

with the filtered signal are chosen to calculate the PE value [18]:

$$\rho_{xy} = \frac{\sum_{k=1}^{N} x(k) y(k)}{\left[\sum_{k=1}^{N} x(k)^2 \sum_{k=1}^{N} y(k)^2\right]^{1/2}}. \tag{10}$$

Step 4. The PE values of all the chosen IMFs are calculated to generate a feature vector which can be utilized to train the SVM for identification of gearbox operating condition.

Step 5. The PE feature vector from test gearbox vibration signal is extracted and utilized as input to the well-trained SVMs. In this way, the result of classification can be realized [19, 20].

3. Experimental Evaluation

A series of gearbox fault signals acquired from LC5T81 type transmission were used to verify the effectiveness of the presented approach. The data was measured from the testbed presented in Figure 6. One backward speed and five forward speeds could be load on the tested gearbox. The vibration signals were collected at 3000 samples per second using the accelerometer fixed on the gearbox case. The tested gearbox is operated with the third speed of 1600 rpm and the meshing frequency of 500 Hz.

The waveforms of the vibration signals collected from the test gearbox under three conditions are shown in Figure 7.

Table 1: Correlation coefficients between filtered signals and each IMF.

	Correlation coefficient									
	IMF 1	IMF 2	IMF 3	IMF 4	IMF 5	IMF 6	IMF 7	IMF 8	IMF 9	IMF 10
Normal condition	0.5235	0.7648	0.4783	0.4284	0.4036	0.1092	0.0971	−0.0056	−0.0012	−0.0078
Light fault condition	0.8594	0.4635	0.3582	0.2347	0.1921	0.0925	−0.0008	−0.0074	−0.0056	−0.0033
Severe fault condition	0.9219	0.2578	0.2215	0.1937	0.1625	0.0859	−0.0014	0.0008	−0.0015	0.0007

Table 2: Permutation entropy values of IMF 1~IMF 5.

	Permutation entropy value				
	IMF 1	IMF 2	IMF 3	IMF 4	IMF 5
Normal condition	0.958	1.239	0.782	0.837	0.5036
Light fault condition	0.710	0.859	0.518	0.421	0.8421
Severe fault condition	0.539	0.709	0.876	0.659	0.2625

Table 3: Fault diagnosis using improved approach based on CEEMD and PE.

Fault type	Test sample	Classification results			Classification rate [%]	Overall classification rate [%]
		Normal condition	Minor fault condition	Serious fault condition		
Normal condition	20	20	0	0	100	
Light fault condition	20	1	18	1	90	95
Severe fault condition	20	0	1	19	95	

Table 4: Fault diagnosis using the approach based on CEEMD and ApEn.

Fault type	Test sample	Classification results			Classification rate [%]	Overall classification rate [%]
		Normal condition	Minor fault condition	Serious fault condition		
Normal condition	20	19	0	1	95	
Light fault condition	20	1	17	2	85	88.3
Severe fault condition	20	1	2	17	85	

Figure 7(a) shows the signal under the normal condition, Figure 7(b) shows the signal under the light fault condition, and Figure 7(c) shows the signal from the severe fault condition.

Figure 8 illustrates the decomposed IMFs of these signals and Table 1 shows the correlation coefficients between the filtered signal and each of the IMFs.

It can be seen from the table that correlation coefficients for the first 5 IMFs are all more than 0.1. They can describe the main features of the signal and thus are selected for further analysis. According to the main steps of the presented fault diagnosis approach, the permutation entropy values of these IMFs are calculated, as listed in Table 2.

In the experiment, 120 feature vectors in total were gained from three different circumstances. 50% of the feature vectors were applied into classifier training, while the rest of them were used in classification of fault. Table 3 shows the results of classification. It shows that various working conditions can be efficiently identified. Among all the 60 groups of feature vectors, 57 groups have been classified correctly, while 3 groups have failed. The overall classification accuracy is up to 95%.

For purpose of comparison, the values of approximate entropy (ApEn) from the chosen IMFs are also calculated and applied in the SVM classifier. Table 4 shows the classification results. It can be summarized that the method is actually efficient for differentiating the gearbox faults. Furthermore, the effectiveness of the approach is compared with that of the EEMD-PE approach. It can be seen that the rates of classification in these two approaches are very similar. However, computational load of the developed approach is lower than that of the EEMD-PE approach.

To further study the effectiveness of the developed approach, a 10×10-fold cross validation procedure is employed

(a) Normal condition
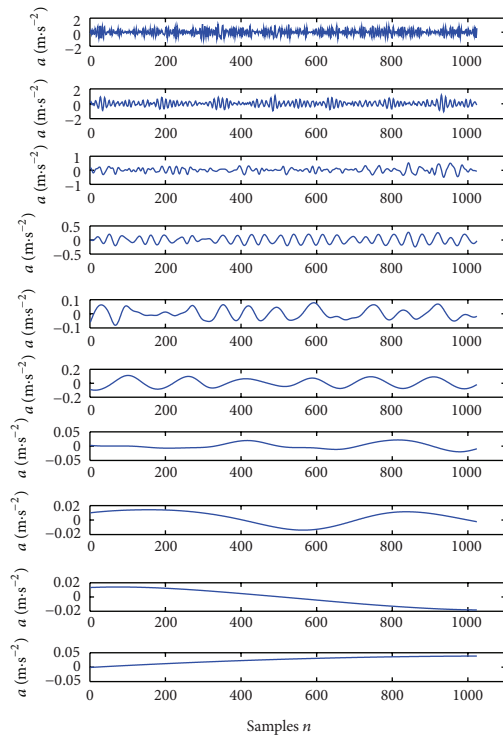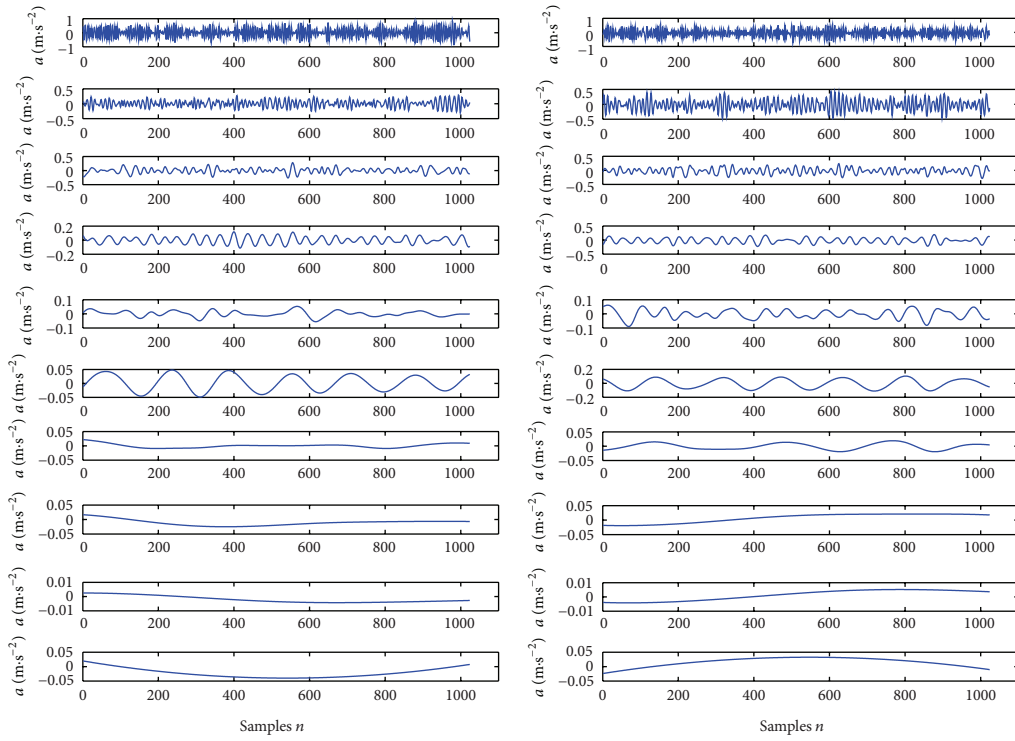
(b) Light fault condition

(c) Severe fault condition

FIGURE 8: The decomposition result by CEEMD under different conditions.

with the selected 120 samples. The average classification rate of the 10×10-fold cross validation is 94.82%. The result is close to the classification result in Table 3.

4. Conclusions

This study develops an integrated approach by combining PE algorithm with CEEMD to diagnose gearbox faults. With the CEEMD, gearbox vibration signals can be decomposed into sets of IMFs with low computational load. Then PE method can efficiently extract fault characteristic from the selected IMFs. Without mathematical model and the study of the fault mechanism of the system, this developed approach can directly recognize gearbox fault severity. Furthermore, the CEEMD, as a preprocessing step, can be utilized to purify the signal for PE calculation, leading to increased classification rate (e.g., 95% for experimental data). It is envisioned that the approach developed in this study could be used in a wide range of applications in the field of fault diagnosis.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

References

[1] W. Fan, G. Cai, Z. K. Zhu, C. Shen, W. Huang, and L. Shang, "Sparse representation of transients in wavelet basis and its application in gearbox fault feature extraction," *Mechanical Systems and Signal Processing*, vol. 56–57, pp. 230–245, 2015.

[2] D. Wang, Q. Miao, Q. Zhou, and G. Zhou, "An intelligent prognostic system for gear performance degradation assessment and remaining useful life estimation," *Journal of Vibration and Acoustics*, vol. 137, no. 2, Article ID 021004, 2015.

[3] R. Yan and R. X. Gao, "Wavelet domain principal feature analysis for spindle health diagnosis," *Structural Health Monitoring*, vol. 10, no. 6, pp. 631–642, 2011.

[4] D. Wang, Q. Miao, and R. Kang, "Robust health evaluation of gearbox subject to tooth failure with wavelet decomposition," *Journal of Sound and Vibration*, vol. 324, no. 3–5, pp. 1141–1157, 2009.

[5] R. Yan, Y. Liu, and R. X. Gao, "Permutation entropy: a nonlinear statistical measure for status characterization of rotary machines," *Mechanical Systems and Signal Processing*, vol. 29, pp. 474–484, 2012.

[6] S.-D. Wu, P.-H. Wu, C.-W. Wu, J.-J. Ding, and C.-C. Wang, "Bearing fault diagnosis based on multiscale permutation entropy and support vector machine," *Entropy*, vol. 14, no. 8, pp. 1343–1356, 2012.

[7] J. Zheng, J. Cheng, and Y. Yang, "Multiscale permutation entropy based rolling bearing fault diagnosis," *Shock and Vibration*, vol. 2014, Article ID 154291, 8 pages, 2014.

[8] X. Zhang, Y. Liang, J. Zhou, and Y. Zang, "A novel bearing fault diagnosis model integrated permutation entropy, ensemble empirical mode decomposition and optimized SVM," *Measurement*, vol. 69, pp. 164–179, 2015.

[9] K. Keller and M. Sinn, "Ordinal analysis of time series," *Physica A: Statistical Mechanics and its Applications*, vol. 356, no. 1, pp. 114–120, 2005.

[10] Z. H. Wu and N. E. Huang, "Ensemble empirical mode decomposition a noise assisted data analysis method," *Advances in Adaptive Data Analysis*, vol. 1, no. 1, pp. 1–41, 2009.

[11] J. Zhang, R. Yan, R. X. Gao, and Z. Feng, "Performance enhancement of ensemble empirical mode decomposition," *Mechanical Systems and Signal Processing*, vol. 24, no. 7, pp. 2104–2123, 2010.

[12] Y. Lei, J. Lin, Z. He, and M. J. Zuo, "A review on empirical mode decomposition in fault diagnosis of rotating machinery," *Mechanical Systems and Signal Processing*, vol. 35, no. 1-2, pp. 108–126, 2013.

[13] M. E. Torres, M. A. Colominas, G. Schlotthauer, and P. Flandrin, "A complete ensemble empirical mode decomposition with adaptive noise," in *Proceedings of the 36th IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '11)*, pp. 4144–4147, Prague, Czechia, May 2011.

[14] 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*, vol. 454, no. 1971, pp. 903–995, 1998.

[15] R. Tiwari, V. K. Gupta, and P. K. Kankar, "Bearing fault diagnosis based on multi-scale permutation entropy and adaptive neuro fuzzy classifier," *Journal of Vibration and Control*, vol. 21, no. 3, pp. 461–467, 2015.

[16] G. Xiong, W. Huang, and L. Zhang, "Fault severity identification of rolling bearing based on multiscale entropy," *Journal of Applied Sciences*, vol. 13, no. 13, pp. 2404–2408, 2013.

[17] Z. H. Wu and N. E. Huang, "A study of the characteristics of white noise using the empirical mode decomposition method," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 460, no. 2046, pp. 1597–1611, 2004.

[18] R.-X. Chen, B.-P. Tang, and J.-H. Ma, "Adaptive de-noising method based on ensemble empirical mode decomposition for vibration signal," *Journal of Vibration and Shock*, vol. 31, no. 15, pp. 82–86, 2012.

[19] D. Fernández-Francos, D. Marténez-Rego, O. Fontenla-Romero, and A. Alonso-Betanzos, "Automatic bearing fault diagnosis based on one-class v-SVM," *Computers & Industrial Engineering*, vol. 64, no. 1, pp. 357–365, 2013.

[20] L. M. R. Baccarini, V. V. Rocha E Silva, B. R. de Menezes, and W. M. Caminhas, "SVM practical industrial application for mechanical faults diagnostic," *Expert Systems with Applications*, vol. 38, no. 6, pp. 6980–6984, 2011.