An adjustable envelope based EMD method for rolling bearing fault diagnosis

Y F Hu1 and Q Li2

1 School of Electrical Engineering and Automation, Harbin Institute of Technology, Harbin, 150001, China
2 Beijing Institute of Tracking and Telecommunications Technology, Beijing, 100094, China

E-mail: yifanhu_123@163.com

Abstract. Empirical mode decomposition (EMD) can decompose complex non-stationary signals into the sum of several intrinsic mode functions (IMF), and it is widely used in fault diagnosis of mechanical devices. In order to solve the mode mixing problem and enhance the fault feature extraction ability, an adjustable envelope based EMD (AE-EMD) method was proposed. Firstly, the envelope fitting method is replaced by rational Hermite interpolation, which can effectively avoid the outstanding over and undershoot problem in the conventional fitting curve. Secondly, the optimal parameter in the AE-EMD is searched by grey wolf optimizer (GWO) using the maximum Kurtosis index as the objective function. Then, the AE-EMD method combined with the Hilbert envelope spectrum is employed to extract the fault characteristic frequency. Case study demonstrates that AE-EMD can restrain the mode mixing effectively, and it also has better fault feature extraction ability compared with the conventional EMD. This proposed method has potential significance to the fault diagnosis and condition-based maintenance (CBM) of rolling bearings in the large and complex equipment.

1. Introduction

Recently, “Made in China 2025” strategy calls for more our attention to the reliability, fault diagnosis, and prognostics of the large and complex equipment. Rolling bearing is one of the most important components of rotating machinery in aerospace and weapon equipment. Because rolling bearings are in harsh working condition of high temperature and high load for a long time, they often have a high failure rate and easily damaged. Hence, its working state directly affects and determines the reliability and safety of the whole system. If the incipient fault features of rolling bearings can be extracted exactly and the corresponding maintenance strategy can be made in time, it will be of great significance to maintain the reliability and safety of mechanical systems [1].

When the rolling bearing has local damage or defect, it will produce periodic vibration impact, which makes the equipment noise increase and some abnormal vibration will occur. Through a large number of literature research and case studies, it is concluded that the condition monitoring method based on vibration signal is the most effective mean to diagnose rolling bearing faults [2,3]. Many experts and scholars have done a lot of research work in fault diagnosis of rolling bearings. In 1998, N. E. Huang et al. [4] proposed the empirical mode decomposition (EMD) method which is an adaptive signal processing method for the first time. Once the method was proposed, it has attracted
extensive attention of scholars and it has been successfully applied to the fault diagnosis of rolling bearings. Gai and Hu [5] proposed a fault diagnosis method based on EMD-SVD and fuzzy neural network (FNN). Zhang [6] designed a fault diagnosis scheme for rolling bearings based on EMD and spectral kurtosis analysis. Although EMD has obvious advantages in rolling bearing fault diagnosis, it also has inherent disadvantages such as modal mixing and endpoint effect. In order to improve this phenomenon, CEEMD was proposed by adding specific white noise at each stage of EMD, and the phenomenon of mode mixing was suppressed greatly [7]. Chen [8] proposed a comprehensive fault diagnosis method based on CEEMD, sample entropy and correlation analysis. Huang [9] decomposed the vibration signals using CEEMD, and the key IMF components obtained were analyzed by 1.5-dimensional spectrum. Many researchers believed that the fundamental reason for the mode mixing and endpoint effect in conventional EMD is that there exists the outstanding over and undershoot in the upper and lower envelope curve when using cubic spline interpolation. And this fitting process is a crucial step in the whole procedures of EMD, which will affect the decomposition result seriously. So some envelope interpolation methods have been used to improve the conventional EMD [10,11]. But the envelope curve in [10,11] is stationary, which can’t fit the various conditions. Considering the rational Hermite interpolation method has good shape-preserving property, so it is suitable to fit the extremum points of various signals with strong non-stationary characteristics. Consequently, an adjustable envelope based EMD (AE-EMD) was proposed, the optimal parameter in the AE-EMD is searched by grey wolf optimizer (GWO) using the maximum Kurtosis as the objective function. Case study demonstrates that AE-EMD can restrain the mode mixing effectively, and it also has better fault feature extraction ability compared with the conventional EMD.

2. Algorithm of adjustable envelope based EMD method

In this section, the rational Hermite interpolation and the parameter selection criterion are introduced prior to the AE-EMD method.

2.1. Rational Hermite interpolation

For a given sequence \((x\_i, y\_i)\), \(i=1,2,3,...,N\), \(h\_i=x\_i+\_1-x\_i\), and \(t=\frac{x-x\_i}{h\_i}\). \(d\_i=P'(x\_i)\) denotes the slope of interpolation basis function at \(x\_i\), and the basis functions can be described as follows [12]:

\[
F_1(t) = 1 - (\lambda - 3)t^2 - (2\lambda - 2)t^3 + \lambda t^4 \\
F_{i+1}(t) = -(\lambda - 3)t^2 + (2\lambda - 2)t^3 - \lambda t^4 \\
G_1(t) = t + (\lambda - 2)t^2 - (2\lambda - 1)t^3 + \lambda t^4 \\
G_{i+1}(t) = -(\lambda + 1)t^2 + (2\lambda + 1)t^3 - \lambda t^4
\]

(1)

where \(\lambda\) represents the shape parameter. And the basis functions also satisfy the following relationships:

\[
F_1(0) = F_{i+1}(1) = 1, F_1'(0) = F_{i+1}'(0) = 0, \\
F_1(0) = F_1'(0) = F_{i+1}'(0) = F_{i+1}'(0) = 0, \\
G_1(0) = G_{i+1}(1) = G_{i+1}(0) = 0, \\
G_1'(0) = G_{i+1}'(1) = 1, G_1'(0) = G_{i+1}'(0) = 0, \\
F_1(t) = F_{i+1}(t) = 1, G_1(t) = -G_{i+1}(1-t)
\]

(2)

Therefore, it can be expressed as:

\[
P(t) = y\_i + h\_i d\_i t + \left[-(\lambda - 3)t^2 h\_i + (\lambda - 2)h\_i d\_i - (\lambda + 1)h\_i d_{i+1}\right]t^2 \\
+ \left[(2\lambda - 2)t^2 h\_i - (2\lambda - 1)h\_i d\_i + (2\lambda + 1)h\_i d_{i+1}\right]t^3 \\
+ \left[-\lambda t^2 h\_i + \lambda h\_i d\_i - \lambda h\_i d_{i+1}\right]t^4
\]

(3)
where \( \tau_i = \frac{y_{i+1} - y_i}{h_i} \) is the differential coefficient.

Fig. 1 displays the rational Hermite-based envelope curves with different shape parameters, it can be seen that the new envelope algorithm can restrain the over and undershoot problem effectively, meanwhile, the optimal parameter selection is also necessary.

Figure 1. The rational Hermite-based envelope curves with different shape parameters.

2.2. Parameter selection criterion: maximum Kurtosis index

Because Kurtosis can reflect the sparse characteristics of vibration signal, so it is commonly used as a time-domain index to measure the damage degree of rotating machinery [13]. Consequently, if the obtained IMF component contains abundant fault feature information and regular impulse in the waveform, it will show strong sparse characteristic, which has maximum Kurtosis value, vice versa. The calculation of Kurtosis is expressed as follow:

\[
K = \frac{1}{N} \frac{\sum_{i=1}^{N} (X(i) - \overline{X})^4}{\left(\frac{1}{N} \sum_{i=1}^{N} (X(i) - \overline{X})^2\right)^2}
\]

where \( K \) represents the Kurtosis index of vibration signal, \( N \) is the length of signal. Based on the above discussion, the maximum Kurtosis index is constructed and employed as the objective function.

2.3. Proposed method

Based on the above discussion, the proposed AE-EMD method is designed, in which the maximum Kurtosis index is defined as the objective function as the form of (5). And the GWO [14] is employed to search for the optimal parameter.

\[
\begin{align*}
\text{fitness} &= \max_{\gamma = \lambda} \left\{ K_i \right\} \\
\text{s.t.} \; \lambda &\in [-2.5, 2.5]
\end{align*}
\]

where \( \text{fitness} \) denotes the objective function, \( K_i (i = 1, 2, ..., N) \) is the Kurtosis index of each mode, and \( \gamma = \lambda \) is the shape parameter to be optimized. And in this paper, \( \lambda \) takes the real number in the interval of \([-2.5, 2.5]\). The AE-EMD algorithm can be summarized as follows:

1) Input the vibration signal \( X(t) \), then set the range of the shape parameter \( \lambda \), and initialize GWO algorithm.
2) Construct the envelope curves using the optimized rational Hermite interpolation and then obtain the optimal modes by optimization and decomposition.

3) Select the sensitive IMF modes using correlation coefficient for further fault feature extraction by Hilbert envelope spectrum.

The flowchart of the proposed method is presented in Fig. 2.

Figure 2. The flowchart of the AE-EMD method

3. Case validation
In this section, a case study of bearing inner race fault diagnosis is conducted to verify the effectiveness of the AE-EMD method, and the comparative analysis highlight its superiority in alleviating mode mixing and extracting more explicit fault feature.

3.1. Data acquisition
The experimental data are collected from the CWRU. And the test device diagram is shown in Fig. 3 [15]. As this test device and data set are widely known and applied, for more detail information, please refer to [15]. In this experiment, the inner race fault was taken as an example to extract its fault features. After calculation, the characteristic frequency is 155.3Hz.
3.2. Bearing inner race fault diagnosis and comparative analysis

Vibration signals of bearing in inner race fault state were collected as shown in Fig. 4. Because the vibration signal is non-stationary and disturbed by the environment, it is difficult to get more explicit fault information from the diagram directly.

Then we decomposed the original signal using the AE-EMD method. After the parameter optimization, \( \lambda = 0.78 \) for generating the first component is adopted. The decomposition results of the first five improved IMF (I-IMF) were shown in Fig. 5.
Figure 5. The first five I-IMF components generated by AE-EMD method

Then the correlation analysis was used to further select the key IMFs to extract the frequency characteristics of inner race fault. The correlation coefficient calculation results were shown in Table 1.

Table 1 The correlation coefficient calculation results

| I-IMF components | 1     | 2     | 3     | 4     | 5     |
|------------------|-------|-------|-------|-------|-------|
| Correlation coefficient | 0.8644 | 0.4487 | 0.2829 | 0.1409 | 0.0906 |

From Table 1, we could clearly see that the correlation coefficient of the first three IMFs were bigger than the others, so they should be considered as the key I-IMFs. Then, we make the Hilbert envelope spectrum analysis based the selected I-IMFs. And the analysis results were shown in Fig. 6.

Figure 6. The Hilbert envelope spectrum of the key I-IMF components

We have known that the characteristic frequency of the inner race fault is 155.3Hz ($f_r$) after calculation. From the Hilbert envelope spectrums, the fault characteristic frequency $f_r$ as well as its harmonics ($2f_r$, $3f_r$, $4f_r$, $5f_r$) can be clearly observed in the spectrum of I-IMF1. Meanwhile, $f_r$ and $2f_r$ are also visible in IMF2, besides, we can still extract $f_r$ in IMF3. So this case fully demonstrates that the proposed AE-EMD method can identified more explicit and abundant fault feature information.
In addition, the conventional EMD is also used to decompose the same signal for comparison. The decomposition results and corresponding Hilbert analysis of key modes (also selected by correlation analysis) was shown in Fig. 7 and 8, respectively.

![Figure 7](image1.png)

Figure 7. The first five IMF components generated by conventional EMD

![Figure 8](image2.png)

Figure 8. The Hilbert envelope spectrum of the key IMF components

The mode mixing appears in the fifth IMF component, which is denoted by the red rectangles, while it has been restrained apparently by AE-EMD method as shown in Fig. 5. From Fig. 8, we could also observe the fault characteristic frequency $f_i$ as well as its harmonic frequency $2f_i$ in the first IMF, but the fault characteristic frequency was submerged in the second and third IMF, while it is still clearly visible in the corresponding spectrum of AE-EMD method.

Furthermore, the construction of the envelope curves for generating the first component is also presented for comparison to explain the definite reason of mode mixing mentioned in Introduction part. In Fig. 9, it can be clearly seen that the rational Hermite interpolation with optimal parameter effectively overcomes the over and undershoot problems, and construct more tense and explicit envelope curves.
In summary, this case study is sufficient to prove the rationality and validity of the AE-EMD method proposed in this paper, meanwhile, the comparative analysis highlights its advantages and superiorities.

4. Conclusions
In this paper, an adjustable envelope based EMD (AE-EMD) method was proposed, the maximum Kurtosis index is employed as the parameter selection criterion, and the optimal parameter is searched by GWO algorithm. This method can restrain the mode mixing phenomenon evidently, and it also outperforms in fault feature extraction and decomposition accuracy compared with the conventional EMD. Case study combined with comparative analysis highlights the effectiveness and superiority of the proposed method. So this method can be considered as a potential improvement in fault diagnosis and condition-based maintenance (CBM) of rotating machinery.

Acknowledgments
This work was supported by the National Key R&D Program of China [grant number 2017YFB1300800] and the National Natural Science Foundation of China [grant number 61671172].

References
[1] Hong S, Zhou Z, Zio E and et al 2014 Condition assessment for the performance degradation of bearing based on a combinatorial feature extraction method J. Digital Signal Processing 27 159-66
[2] El-Thalji I and Jantunen E 2015 A summary of fault modelling and predictive health monitoring of rolling element bearings J. Mechanical Systems and Signal Processing 60-61 252-72
[3] Mcfadden P D and Smith J D 1984 Vibration monitoring of rolling element bearings by the high-frequency resonance technique — a review J. Tribology International 17(1) 3-10
[4] Huang N E, Shen Z, Long S R and et al 1998 The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis J. Proceedings A 454(1971) 903-95
[5] Gai J and Hu Y Research on Fault Diagnosis Based on Singular Value Decomposition and Fuzzy Neural Network J. Shock and Vibration (2018-4-8) 2018(1) 1-7
[6] Zhang Z, Shi X, Shi and et al 2013 Fault feature extraction of rolling bearings based on improved EMD and spectral kurtosis method J. Vibration, testing and diagnosis 33(3):478-82

[7] Yeh J R, Shieh J S and Huang N E 2010 Complementary Ensemble Empirical Mode Decomposition: a Novel Noise Enhanced Data Analysis Method J. Advances in Adaptive Data Analysis 2(2) 135-56

[8] Chen J, Zhou D, Lyu C and et al 2017 An integrated method based on CEEMD-SampEn and the correlation analysis algorithm for the fault diagnosis of a gearbox under different working conditions J. Mechanical Systems and Signal Processing 113

[9] Huang H, Sun B, Ren X and et al 2019 Early fault diagnosis method for rolling bearings based on CEMDAN and 1.5-dimensional spectrum J. China Test 45(2) 151-6

[10] S. D. Hawley, L. E. Atlas and H. J. Chizeck Jan. 2010 Some Properties of an Empirical Mode Type Signal Decomposition Algorithm IEEE Signal Processing Letters 17(1) 24-7

[11] Q. H. Chen, N. Huang, S. Riemenschneider and Y. S. Xu Jul. 2008 A B-spline approach for empirical mode decompositions Advances in Computational Mathematics 29(1) 93

[12] Y. B. Li, M. Q. Xu, Y. Wei and W. H. Huang Mar. 2015 An improvement EMD method based on the optimized rational Hermite interpolation approach and its application to gear fault diagnosis Measurement vol. 63 330-45

[13] X. Zhang, Q. Miao, H. Zhang and L. Wang Aug. 2018 A parameter-adaptive VMD method based on grasshopper optimization algorithm to analyze vibration signals from rotating machinery Mechanical Systems and Signal Processing 108 58-72

[14] Zhang X, Liu Z, Miao Q and et al 2018 An optimized time varying filtering based empirical mode decomposition method with grey wolf optimizer for machinery fault diagnosis J. Journal of Sound and Vibration 418 55-78

[15] Bearing Data Center, Case Western Reserve University. <http://csegroups.case.edu/bearingdatacenter/pages/download-data-file>.