An Improved Signal Processing Approach Based on Analysis Mode Decomposition and Empirical Mode Decomposition

Zhongzhe Chen 1,*, Baqiao Liu 2, Xiaogang Yan 1 and Hongquan Yang 1

1 The School of Mechanical and Electrical Engineering, University of Electronic and Science Technology of China, Chengdu 611731, China
2 Department of Computer Science, University of North Carolina, Chapel Hill, NC 27516, USA
* Correspondence: zhzhchen@163.com; Tel.: +86-139-8179-1418

Received: 25 June 2019; Accepted: 5 August 2019; Published: 9 August 2019

Abstract: Empirical mode decomposition (EMD) is a widely used adaptive signal processing method, which has shown some shortcomings in engineering practice, such as sifting stop criteria of intrinsic mode function (IMF), mode mixing and end effect. In this paper, an improved sifting stop criterion based on the valid data segment is proposed, and is compared with the traditional one. Results show that the new sifting stop criterion avoids the influence of end effects and improves the correctness of the EMD. In addition, a novel AEMD method combining the analysis mode decomposition (AMD) and EMD is developed to solve the mode-mixing problem, in which EMD is firstly applied to dispose the original signal, and then AMD is used to decompose these mixed modes. Then, these decomposed modes are reconstituted according to a certain principle. These reconstituted components showed mode mixing phenomena alleviated. Model comparison was conducted between the proposed method with the ensemble empirical mode decomposition (EEMD), which is the mainstream method improved based on EMD. Results indicated that the AEMD and EEMD can effectively restrain the mode mixing, but the AEMD has a shorter execution time than that of EEMD.

Keywords: empirical mode decomposition; analysis mode decomposition; analysis-empirical mode decomposition; mode mixing; sifting stop criterion

1. Introduction

The analysis of time-frequency of vibration signals is one of the most effective and important methods for fault diagnosis of rotating machinery, since the vibration signal includes massive information that reflects the running state of rotating machinery [1]. The empirical mode decomposition (EMD) is one of most commonly used methods for signal processing in the time-frequency domain. EMD can decompose a complex signal into the finite intrinsic mode functions (IMF) based on the local characteristic time scale of signals, and each IMF represents one intrinsic vibration mode of the original signal. Then the characteristic information of the signal can be extracted by analyzing such stationary stable IMFs [2]. EMD has attracted increasing attention since it appeared [3–5], and it has been widely used in economics, biomedicine and engineering science fields, especially for fault diagnosis. Ali et al. [6] applied the EMD method and artificial neural network in fault diagnosis of rolling bearings automatically. Xue et al. [7] presented an adaptive, fast EMD method and applied it to rolling bearings fault diagnosis. Yu et al. [8] introduced various applications of EMD in fault diagnosis. Cheng et al. [9] combined EMD with a Hilbert transform to conduct the recognition for mode parameters. Then, they introduced how to apply the EMD to the fault diagnosis for local rub-impact of rotors [10]. Rilling [11] investigates how the EMD behaves under the case of a composite two-tone signal.
Though EMD has been commonly utilized in reality, some shortcomings are exposed, including mode mixing, end effects, stop criteria of IMF, over-envelope and under-envelope. These deficiencies normally restrict the further promotion and application of EMD [12]. Since the criterion to judge IMF is that the average value of its upper and lower envelope spectrums is zero in the EMD method, but in the real decomposition process, that average value is impossible to be zero due to the disturbance of cubic spline interpolation, the influence of end effect and adopted frequency. Thus, it is necessary to define a valid stop criterion for the decomposition process in engineering, namely the stop criterion problem of IMF for the EMD method, also named as the sifting stop criterion. Some researchers have attempted to improve the efficiency of EMD. Among them, Pustelnik [13] mainly developed an alternative to the sifting process for EMD, based on non-smooth convex optimization allowing integration flexibility in the criteria, proposed algorithm and its convergence guarantees.

Mode mixing means that one IMF contains vastly different characteristic time scales, or similar characteristic time scales that are distributed in different IMFs, which results in the waveform mixing of the two adjacent IMFs. These mixing modes influence each other so that it is difficult to identify them. Mode mixing is the fatal flaw of EMD. It makes the physical significance of IMF components uncertain finally, which influences the correctness of signal decomposition and seriously restricts its application in engineering [14].

Until now, various methods have been developed to restrain the mode-mixing problem in the EMD. Zhao [15] directly filtered abnormal information related to the IMF and fitted filtered data segment by spline interpolation, but it only proved to dispose the mode mixing problems caused by a known transient abnormity. The masking signal method [16] and the high frequency harmonic method [17] are simple and effective, but they are susceptible to distortion and need to be reprocessed for practical engineering signals.

The ensemble empirical mode decomposition (EEMD) presented by Huang is generally considered an effective one [18]. EEMD takes advantage of the statistical characteristics of white Gaussian noise while frequency is uniformly distributed, so the signal after adding white Gaussian noise shows continuity in different scales, which solved the mode mixing problem to some extent. However, it also raised some other issues. For example, the number and the amplitude of white noises which are added in the signal are greatly subjective, and the EEMD sacrifices some adaptivity. In addition, although the number and the amplitude of added white noises are chosen reasonably, the mode mixing in low frequency may be aroused artificially while high-frequency mode mixing is restricted [19]. Moreover, the algorithm of the EEMD method is complex and it takes a long time to run the program, which will restrict its application on the signal process that demands to be processed in real-time. Accordingly, some researchers presented improved methods to overcome the deficiencies of EEMD. Among them, Lei [20] proposed adaptive EEMD to improve its adaptivity. Zheng [21] developed partial EEMD and Tan [22] presented multi-resolution EMD to solve mode mixing problem. Mohammad [23] uses approximate entropy and mutual information to improve EEMD to generate statistical features in order to increase the performance of early fault appearance detection, as well as the fault type and severity estimation.

In this paper, a new sifting stop criterion was proposed based on valid data segments to solve the problem of sifting stop criteria in the EMD, and an improved method, namely AEMD, combined the analysis mode decomposition (AMD) and EMD. It was developed to solve the mode-mixing problem. The sifting stop criteria and AEMD proposed were applied to a simulation signal and an engineering case to illustrate the validity and superiority of the proposed method. The structure of the paper is as follows: Section 2 introduces the basic principles of EMD; Section 3 narrates the proposed sifting stop criteria based on valid data segment and compares it with the original one; in Section 4, the principle and steps of AEMD are expounded firstly, then applied to decompose simulation signals and a rotor vibration signal, and it is compared with the EMD and EEMD methods; finally, Section 5 draws a brief conclusion of current work.
2. Basic Principles of EMD

EMD refers to a “sifting” process, in which, the component with the smallest extreme time feature scale is sifted out firstly, then those with larger extreme time feature scales, and the component with the largest feature scales are finally sifted out. That is to say, the average frequency of IMF components obtained from the EMD is reduced gradually, and the main steps of the EMD are introduced as follows [24]:

(1) Firstly, identify all the local maximum points and minimum points of original signal $x(t)$, then match the upper and lower envelope spectrums of extreme points with cubic spline line respectively. Furthermore, ensure that the signal $x(t)$ is between the upper and lower envelope spectrums; (2) calculate the local mean of upper and lower envelope spectrums, denoted as $m_1$; (3) calculate the first component $h_1(t)$ according to Equation (1):

$$ h_1(t) = x(t) - m_1 $$ (1)

(4) Judge that whether $h_1(t)$ can satisfy the conditions to be an IMF. If not, $h_1(t)$ should be treated as an original signal to repeat steps (1), (2) and (3) until $h_1(t)$ can meet the conditions, designated as $C_1(t) = h_1(t)$. $C_1(t)$ is the first IMF component after decomposition. (5) Separate $C_1(t)$ from the signal $x(t)$, i.e., $r_1(t) = x(t) - C_1(t)$. (6) Treat $r_1(t)$ as an original signal to repeat steps (1)–(5), and after $n$ cycles, $n$ IMF components and 1 residual value $r_n(t)$ can be derived; i.e.,

$$ \begin{align*}
  r_2 &= r_1 - c_2 \\
  r_3 &= r_2 - c_3 \\
  \vdots \\
  r_n &= r_{n-1} - c_n
\end{align*} $$ (2)

Then, the original signal $x(t)$ can be expressed as:

$$ x(t) = \sum_{i=1}^{n} c_i + r_n $$ (3)

where $c_i$ refers to the $i$th IMF component and $r_n$ is the residual function.

This cycle comes to the end when $r_n$ becomes a monotonic function, but in the real decomposition process, when $r_n$ meets the conditions of monotonic function, the cycle number is usually large, and the number of IMF components will be too large. Furthermore, a mass of EMD tests have revealed that most of these final components obtained from EMD are false IMF components, without substantive physical significance. Thus, a good sifting stop criterion can not only improve the decomposition efficiency, but also increase the decomposition accuracy.

The stop criteria of decomposition process proposed by Huang was realized by limiting the standard deviation of the two adjacent IMFs and the sifting times are eventually controlled by the iteration threshold $S_d$ [25]. In particular, $S_d$ is defined as:

$$ S_d = \sum_{i=0}^{T} \frac{|h_k(t) - h_{k-1}(t)|^2}{h_{k-1}^2(t)} $$ (4)

where $T$ refers to the time span of a signal, $h_k(t)$ and $h_{k-1}(t)$ denote the two adjacent processing sequences in the process of EMD and the value of $S_d$ is usually between 0.2 and 0.3.

It is of great significance to determine reasonable iteration threshold $S_d$. If the threshold is too small, the computational cost will increase greatly and the IMF components finally obtained will be of no significance; while if it is too large, it will be difficult to satisfy the condition of IMF.
3. The Sifting Stop Criteria Based on the Valid Data Segment

The sifting stop criteria of IMF proposed by Huang is valid for most cases, but there are some problems; it can be noted from Equation (3) that when $h_k(t) = 0$, the iteration threshold will become an uncertain value. It is evidently inappropriate if the iteration threshold is used to control sifting number at the time. In addition, the sifting stop criteria ignored the influence of end effect. Consequently, the decomposition result may produce errors if the distorted endpoint data are adopted. In most cases, the end effect of EMD is obvious, though some effective measures will been taken to restrain it, the end effect still exists.

In this paper, an improved sifting stop criterion of IMF based on the valid data segment is proposed, where the valid data segment means the residual data segment after kicking out the distorted endpoint data. The sifting stop criteria fully consider the influence of end effect to sifting stop criteria, and only use the valid data segment to calculate the iteration threshold. The sifting stop criterion is expressed as follows:

If $h_k(t)$ is an IMF, it satisfies the following inequality:

$$-\delta < \bar{m}_{k+1} < \delta$$

where $\bar{m}_{k+1}$ denotes the valid data segment of the mean curve of upper envelope and lower envelope; $\delta$ represents the error threshold, which ranges from 0.01 to 0.1.

The EMD is applied to decompose a simulation signal respectively based on the proposed sifting stop criterion and the traditional one below. The simulation signal is

$$x(t) = x_1(t) + x_2(t), t \in [0, 1]$$

$$x_1(t) = (1 + 0.5 \sin(2\pi \cdot 4t)) \cdot 2 \cos(2\pi \cdot 80t),$$

$$x_2(t) = \sin(2\pi \cdot 24t)$$

The simulation signal consists of $x_1(t)$ and $x_2(t)$ (as shown in Figures 1–3) are the decomposition results by using EMD based on the proposed sifting stop criterion and the traditional one, respectively. It is worth noting that the EMD can efficiently decompose the original signal by using the two sifting stop criteria, but it is clear from the two edges of these two figures that the decomposition result $c_1$ and $c_2$ in Figure 2 are more accurate than $imf_1$ and $imf_2$ in Figure 3, which show an evident end effect.

Figure 1. The simulation signal and its two components.
Figure 2. The result of empirical mode decomposition (EMD) method using the sifting criterion based on the valid data segment.

Figure 3. The result of EMD method using the sifting criterion proposed by Huang.

Figure 4 plots the vibration signal of a real gear with broken teeth. The number of gear teeth $z = 75$; module $m = 2$; the rotating frequency $f = n_2/60 \approx 13.6$. EMD based on the proposed sifting stop criteria is used to decompose the vibration signal of a gear, and obtain five IMFs. The first IMF $imf_1$ contains massive fault information of the gear. Note from its time domain graph (see Figure 5) that the waveform of $imf_1$ presents an obvious modulation feature, and its cycle of modulation wave ($T$, approximately 0.074 s) accordingly had a frequency of about 13.6 Hz, which is exactly the rotating frequency of the faulty gear. Thus, it can be concluded that the fault information has been exacted from the vibration signal of the practical gears by using the proposed sifting stop criterion.
4. AEMD Method

4.1. Analytical Mode Decomposition (AMD) Method

AMD can separate the harmonic components of each frequency band from a multiband signal, its principle is as follows:

A signal \( x_0(t) \) is separated into the two components by boundary frequency \( f_{br} \), one is the fast-changing signal \( x_f(t) \) and another is slowly-changing signals \( x_s(t) \):

\[
x_0(t) = x_f(t) + x_s(t)
\]

Their corresponding Fourier transforms \( X_f(\omega) \), \( X_s(\omega) \) have no overlap in the frequency band. The Hilbert transforms of \( \cos(2\pi f_b t) \cdot x_0(t) \) and \( \sin(2\pi f_b t) \cdot x_0(t) \) are as follows:

\[
H[\cos(2\pi f_b t) \cdot x_0(t)] = x_s(t) \cdot H[\cos(2\pi f_b t)] + \cos(2\pi f_b t) \cdot H[x_f(t)]
\]

(9)

\[
H[\sin(2\pi f_b t) \cdot x_0(t)] = x_s(t) \cdot H[\sin(2\pi f_b t)] + \sin(2\pi f_b t) \cdot H[x_f(t)]
\]

(10)


As aforementioned, the signal \( x_0(t) \) is decomposed into \( x_s(t) \) and \( x_f(t) \) by using the analytical mode decomposition.

4.2. Steps of the AEMD Method

When the signal to be decomposed contains high-frequency intermittent signals or some similar compositions in time feature size, it is inevitable that the mode mixing appears in decomposition result by using EMD directly [11]. In this section, the AEMD that combines AMD with EMD is proposed to restrain the mode mixing in EMD. In particular, EMD is first applied to dispose the original signal, and then AMD is used to decompose these mixed modes. Then, these decomposed modes are reconstituted according to a certain principle. These reconstituted components show mode mixing phenomena alleviated. The specific steps are given as follows:

1. EMD is applied to the original signal and obtain some IMFs and a residual component \( r \);

2. A fast Fourier transform (FFT) is conducted on IMFs, and the spectral diagrams of these IMFs are obtained. The first \( IMF_1(t) \) with mode mixing is performed by AMD. The boundary frequency \( f_{b1} \) can be determined according to the spectrum. The average frequency of the two confused modes is used as the boundary frequency in this analysis. Then \( IMF_1(t) \) is separated into the two signals \( c_1(t) \) and \( \hat{c}_1(t) \) by using AMD. \( c_1(t) \) is the correction component of \( IMF_1(t) \) and \( \hat{c}_1(t) \) is the residual component of \( c_1(t) \). Thus, \( IMF_1(t) \) can be expressed as follows:

\[
IMF_1(t) = c_1(t) + \hat{c}_1(t)
\]

Figure 5. The first IMF of EMD for the vibration signal with broken teeth.
(3) Add $\hat{c}_1(t)$ to the $IMF_2(t)$ and get the renewed $IMF_2^*(t)$, denoted as $IMF_2^*(t)$. The boundary frequency $f_2$ can be obtained according to the amplitude spectrum of $IMF_2^*(t)$. Then $IMF_2^*(t)$ is decomposed into the two signals $c_2(t)$ and $\hat{c}_2(t)$ by using AMD. $c_2(t)$ is the correction component of $IMF_2^*(t)$ and $\hat{c}_2(t)$ is the residual component of $c_2(t)$. Similarly, $IMF_2^*(t)$ is expressed as follows:

$$IMF_2^*(t) = c_2(t) + \hat{c}_2(t)$$ (14)

(4) By that analogy, the similar approach is utilized to dispose all the subsequent IMF components with mode mixing. The final residual component $c_k(t)$ is added to the residual error $r$ and obtain the final residual error, denote as $\tau_k(t)$.

4.3. The Comparison of Simulation Signal Analysis by Different Methods

A simulation signal was decomposed by using EMD, EEMD and the proposed AEMD respectively, and their disposal results were compared, to demonstrate AEMD efficiency and superiority. The simulation signal $x(t)$ is:

$$x(t) = x_1(t) + x_2(t) + x_3(t); t \in [0, 1]$$ (15)

$$x_1(t) = \begin{cases} 
0.8 \cos(600\pi t); & 0.12 \leq t \leq 0.18 \\
0.9 \cos(600\pi t); & 0.42 \leq t \leq 0.48 \\
0.8 \cos(600\pi t); & 0.75 \leq t \leq 0.81 \\
0; & \text{else} 
\end{cases}$$ (16)

$$x_2(t) = 2 \cos(60\pi t); \quad t \in [0, 1]$$

$$x_3(t) = 2.5 \cos(28\pi t); \quad t \in [0, 1]$$ (17)

The simulation signal is shown in Figure 6. The decomposition result of EMD, in Figure 7, shows more serious mode mixing. For the first mode, the high frequency intermittency signal $c_1(t)$ is confused with the sinusoidal signal at 30 Hz, for the second mode $c_2(t)$, the sinusoidal signal at 30 Hz with the sinusoidal signal at 14 Hz. The third and fourth modes are false IMF components. The decomposition result of EEMD is $imf_1$, $imf_2$, $imf_3$, $imf_4$, as shown in Figure 8, with ensemble average number $N = 30$, and the amplitude of noise is set as 0.01 standard deviation of the original signal. The decomposition result of AEMD is shown in Figure 9. A successful decomposition result should distinctly obtain the several components of the original signal. In order to compare the results of decomposition results by these different methods, the green lines in Figures 7–9 represent the several components of the original signal; The blue lines represent the composition results by these different methods. Table 1 lists the run time by these three methods using the same computer.

Note from Figures 7–9 that the goodness of fit between the blue line and green line by EEMD and AEMD methods is higher than that of EMD. The comparison of these three methods reveals that EEMD and AEMD can restrain effectively the mode confusion phenomena, but the running time of EEMD is much longer than those of AEMD and EMD, because the EEMD method adds white noise to original signals to alleviate mode mixing problem, which makes the decomposition process more complex and time-consuming. The executed time of EMD is approximately equal with AEMD’s, but it has inevitable mode mixing problem. Thus, AEMD has shown certain comprehensive advantages in alleviating mode mixing and making decomposition efficient.
Figure 6. The time domain chart of simulation signal $x(t)$.

Figure 7. The decomposition result of the simulation signal by EMD.

Figure 8. The decomposition result of the simulation signal by ensemble empirical mode decomposition (EEMD).

Table 1. The running time of the three methods (unit: seconds).

<table>
<thead>
<tr>
<th>Method</th>
<th>First</th>
<th>Second</th>
<th>Third</th>
<th>Fourth</th>
<th>Fifth</th>
<th>Sixth</th>
</tr>
</thead>
<tbody>
<tr>
<td>EMD</td>
<td>1.3988</td>
<td>1.3860</td>
<td>1.3930</td>
<td>1.3846</td>
<td>1.3820</td>
<td>1.3940</td>
</tr>
<tr>
<td>EEMD</td>
<td>41.4875</td>
<td>44.8383</td>
<td>44.1800</td>
<td>42.8480</td>
<td>43.2321</td>
<td>44.024</td>
</tr>
<tr>
<td>AEMD</td>
<td>1.4027</td>
<td>1.3832</td>
<td>1.3843</td>
<td>1.3934</td>
<td>1.3880</td>
<td>1.3848</td>
</tr>
</tbody>
</table>
The blue lines represent the composition results
The green lines represent the components of the original signal.

Figure 9. The decomposition result of the simulation signal by AEMD (combined AMD and EMD).

4.4. Case Study

Figure 10 presents the experimental device to simulate the local rubbing fault of rotors, during which the rotate-speed was at 300 RPM, controlled by the inputted motor. The displacement sensor that was horizontally installed was employed to sample the vibration signals, with sampling frequency 4k Hz. Figure 11 plots the vibration signal of a practical rotor system with rub-fault. The decomposition result for the vibration signal by using the EMD and the AEMD are shown as Figures 12 and 13, respectively. Figure 14a,b is the spectrum diagram by fast Fourier transform (FFT) for the composition results of EMD and AEMD, respectively.

Note from Figures 12 and 14a that the mode mixing problem still distinctly exists in the decomposition results by EMD. In the imf2 diagram, the characteristic frequency of the rotor rubbing fault is confused in multiple frequencies, caused by the slight rubbing fault of the rotor, which affects the identification for the fault. In the imf3 diagram, the triple frequency of the characteristic frequency of the rotor rubbing fault—150 Hz is nearly invisible, which is mixed in the fundamental frequency of the rotor rubbing fault—50 Hz.

Figure 10. The experimental device for simulating local rubbing fault of rotors.

Figure 11. The vibration signal of rub-fault for rotors.
This spectrum character with multiple higher harmonics is in accordance with the slight rubbing fault. Energies frequency—500 Hz of the characteristic frequency of the rotor rubbing fault are distinctly decomposed. However, the proposed method has certain limitations; the boundary of the AEMD needs less execution time than that of EEMD. The proposed method avoids the influence of end effects and improves the correctness of the EMD. In addition, note from Figures 13 and 14b that the tripling frequency 150 Hz and ten times frequency of the characteristic frequency of the rotor rubbing fault—150 Hz is nearly invisible, which is mixed in the fundamental frequency of the rotor rubbing fault—50 Hz.

In addition, note from Figures 13 and 14b that the tripling frequency 150 Hz and ten times frequency—500 Hz of the characteristic frequency of the rotor rubbing fault are distinctly decomposed. This spectrum character with multiple higher harmonics is in accordance with the slight rubbing fault.
of the rotors. Thus, the characteristic frequency of the rotor rubbing fault is successfully obtained by AEMD, though there is still slighter mode mixing.

5. Conclusions

In this paper, an improved sifting stop criterion based on valid data segment was proposed. Compared with the traditional one, results indicated that the newly proposed sifting stop criterion avoids the influence of end effects and improves the correctness of the EMD. In addition, a novel method, namely AEMD, which combines the AMD and EMD, was developed to solve the mode mixing problem of EMD. Model comparison was conducted between the proposed method and the EEMD. It is worth mentioning that both the AEMD and EEMD can effectively restrain the mode mixing, but the AEMD needs less execution time than that of EEMD. The proposed method overcomes some shortcomings of EMD to some extent and makes it better for use in the fault diagnosis of rotating machinery. However, the proposed method has certain limitations; the boundary frequency of AMD method requires human intervention, which means worse adaptivity than EMD. It can achieve better results for more simple signals, but its advantage is not obvious for complex engineering signals, since they take more time than EMD.

Author Contributions: Z.C. conceived the methodology, conducted the research and wrote the original draft. B.L. performed the programming to fulfill the method and edited the manuscript. X.Y. collected data and analyzed the data. H.Y. investigated the experiment and discussed the results.

Funding: This research was funded by the National Natural Science Foundation of China grant number 11672070.

Conflicts of Interest: We declare no conflict of interest.

References


