RESEARCH ON FAULT FEATURE EXTRACTION OF ROLLING BEARING BASED ON IMPROVED CEEMDAN

Maohua Xiao¹, Cunyi Zhang¹, Kai Wen¹, Yue Zhu¹, Yilidaer Yiliyasi²
¹College of Engineering, Nanjing Agricultural University, Nanjing 210031, China
²College of Mechanical and Electrical Engineering, Xinjiang Agricultural University, Urumqi 830052, China
E-mail: xiaomaohua@njau.edu.cn

Abstract - As a core component of machinery, rolling bearings play an important role in modern industry. In order to avoid unpredictable accidents and undesired downtime costs, fault diagnosis techniques for rolling bearings are widely used in the industrial field. Although there are many troubleshooting methods, such as empirical mode decomposition (EMD), collective empirical mode decomposition (EEMD) and wavelet transform, their disadvantages are very obvious. For example, the eigenmode function (IMF) generated by EEMD always contains residual noise. Furthermore, adding different types of Gaussian noise to the signal will result in a different number of IMFs, which makes the average difficult. In order to eliminate these shortcomings, this paper proposes an improved combination of wavelet transform and full set empirical mode decomposition and adaptive noise (CEEMDAN). Simulation experiments were used to illustrate the effectiveness of CEEMDAN. To further illustrate the proposed method, the improved CEEMDAN is applied to fault diagnosis motor rolling bearings. In addition, two indicators, signal-to-noise ratio (SNR) and root mean square error (RMSE), are established to evaluate the fault diagnosis capability of the proposed method. The analysis results show that the method can improve the fault diagnosis ability of rolling bearings.

Keywords: Rolling Element Bearing, Fault Diagnosis, Signal-noise Ratio, Root Mean Square Error, Improved CEEMDAN.

1. Introduction

Rolling element bearings have been widely used in most industrial fields to support shafts, which are often referred to as key components of rotating machines. However, they often survive in harsh working environments, including corrosion, poor lubrication, plastic deformation, and incorrect design or installation, which often causes bearings to be damaged [1]. In order to avoid unpredictable accidents and unnecessary downtime costs, fault detection and diagnosis of bearings in rotating machines is of great significance [2]. Extracting fault features from the collected signals during the work process plays an important role in the presentation of rolling bearing support conditions [3]. Therefore, the vibration signal processing method has been widely used in various bearing condition monitoring technologies, and has long been the leading analytical tool for bearing fault diagnosis [4].

When the working surface of the bearing component fails, the impact of the bearing fault will produce a rapid transient response with damping. While the analysis of vibration signals is complicated owing to the stochastic movement of rolling elements, much researches have built sound theoretical basis and methods to diagnose bearing failure [5]. Many studies concentrate on obtaining fault information through time and frequency domain signal processing techniques [6-8]. However, rolling element bearing vibration signals are characterized by non-linear, non-stationary, especially under variable speed conditions. The conventional signal analysis methods based on linear system fail to extract nonlinear characteristics in vibration signals. For accurately analyzing the non-linear, non-stationary signals of rolling element bearings, a series of advanced techniques have been employed to extract fault characteristics. Common nonlinear, non-stationary signal processing methods include wavelet transform, morphological filtering, Hilbert-Huang transform, empirical mode decomposition (EMD), and set empirical mode decomposition (EEMD). The combination of Hilbert and wavelet transform The study of bearing failures has shown that these methods greatly improve the extraction ability of bearing characteristics [1,9-11]. Huang improved the Hilbert transform and proposed empirical mode decomposition (EMD), which has been cited by many scholars [12-16]. As a valuable aid to EMD, Collective Empirical Mode Decomposition (EEMD) eliminates the problem of EMD pattern blending to some extent and is also cited by many papers [17-21]. Mathematical
morphological analysis has been proposed to classify the bearing failure, identify the location of failure and defect severity of bearing failure effectively [22-25].

The above methods have obtained good effects on extraction of bearing fault characteristics, but all have their own limitations. Wavelet transform can effectively suppress white noise, but the suppression capability of pulse interference is weak. Mathematical morphology has a strong ability to suppress impulsive interference, and the algorithm is simple and feasible, but there are statistical deviation problems and the choice of the best structural elements. EMD is a powerful tool for analyzing non-stationary signals and non-linear signals, but there is a lack of rigorous mathematical basis, low algorithm efficiency, mode mixing problems. As a valuable aid to EMD, EEMD would solve the problem of mode mixing to some extent by adding the Gaussian white noise and averaging the IMFs. However, with different Gaussian white noises added to the signal, the different numbers of IMFs produced by signal decomposition could make difficult the averaging. In order to make improvement in the final averaging problem produced by EEMD, Torres et al. proposed a method for performing collective empirical mode decomposition using adaptive noise (CEEMDAN) [26]. This approach solves the problem of the final average by adding specific noise instead of Gaussian white noise at each decomposition stage. Great improvement notwithstanding, there are still some “false” IMFs in the earlier decomposition. To enhance the capability of mode decomposition of CEEMDAN, Colominas et al. use the IMF of the Gaussian white noise decomposed by EMD as a unique noise to be added at each decomposition stage [27]. As a result, the noises remained in IMFs are much eliminated.

With the motivation to improve the capability of extraction of weak fault characteristics from noise signals, a new fault diagnosis method of rolling element bearings based on wavelet transform and CEEMDAN is proposed in this paper. In this paper, heuristic threshold wavelet transform is used to preprocess the noise signal, followed by modal decomposition of the processed signal with CEEMDAN. A series of effective IMFs were selected by correlation coefficient method for envelope analysis. The hierarchical envelope and the reconstructed signal envelope analysis methods are employed in this paper. The results of the analysis show that the proposed method can improve the fault diagnosis capability significantly of rolling element bearings. The rest of this paper is organized as follows. In section 2, both wavelet transform and CEEMDAN are briefly introduced. A flow chart of the fault diagnosis of the proposed method will be presented. A simulation illustration is dedicated to make comparison between EEMD and CEEMDAN in section 3. And section 4 gives the real vibration data of the motor bearing to evaluate the proposed method. Two indicators, the signal-noise ratio (SNR) and the root mean square error (RMSE) are employed to further evaluate the fault diagnosis capability of the proposed method as well. Conclusions are drawn in the last section.

2. Improved CEEMDAN

*CEEMDAN*

EMD was originally proposed by Huang, but it has the disadvantage of mode blending. By adding Gaussian white noise to the original signal to change the distribution of extreme points, EEMD can alleviate the shortcomings of mode mixing to some extent. However, the IMF of the EEMD decomposition results still contains more residual noise, and EEMD is not a complete decomposition process. Therefore, in order to overcome the limitations of EEMD, Torres et al. An algorithm called CEEMDAN is proposed. In addition, Colominas et al. Continue to improve CEEMDAN. Add specific noise at each decomposition stage. Each IMF is obtained by calculating the difference between the current remaining portion and its local average. Therefore, the problem of mode mixing in the IMF is largely eliminated, as is the incompleteness of EEMD.

Let \( M(.) \) be the operator which produces the local mean of the signal, and \( E_0(.) \) be the operator which produces the \( k \)th mode decomposed by EMD. It can be viewed that:

\[
E_i(x) = x - M(x) 
\]  

(1)

Let \( w^{(i)} \) be the realization of Gaussian white noise.

\[
x^{(i)} = x + w^{(i)} 
\]  

(2)

And \{ \} the operator of averaging the realizations. Therefore, the first mode we have:

\[
\tilde{d}_1 = \left[ E_i\left(x^{(i)}\right) \right] \left[ \left(x^{(i)} - M\left(x^{(i)}\right) \right) \right] 
\]  

(3)

Which can be simplified as:

\[
\tilde{d}_1 = x - \frac{1}{l} \sum_{i=1}^{l} M\left(x^{(i)}\right) 
\]  

(4)

And there exists:

\[
\frac{1}{l} \sum_{i=1}^{l} M\left(x^{(i)}\right) = r_1 
\]  

(5)

The decomposition steps using the improved CEEMDAN can be shown as follows:

Step 1. Use EMD to calculate the local means of \( l \) realizations:

\[
x^{(i)} = x + \beta_i E_i\left(w^{(i)}\right) 
\]  

(6)

And average them to obtain the first residue as:

\[ E_i(x) = x - M(x) \]  

(1)
Research on Fault Feature Extraction of Rolling Bearing based on Improved CEEMDAN

\[ r_i = \left[ M\left( x^{(i)} \right) \right] = \frac{1}{I} \sum_{i=1}^{I} M\left( x^{(i)} \right) \]  

(7)

Step 2. Calculate the first IMF:
\[ \tilde{d}_1 = x - r_1 \]  

(8)

Step 3. Estimate the second residue as:
\[ r_2 = r_1 + \beta_1 E_1\left( w^{(1)} \right) \]  

And the second IMF is:
\[ \tilde{d}_2 = r_1 - r_2 = r_1 - \frac{1}{I} \sum_{i=1}^{I} M\left( r_i + \beta_1 E_1\left( w^{(1)} \right) \right) \]  

(10)

Step 4. Also the kth IMF is:
\[ \tilde{d}_k = r_{k-1} - r_k \]  

(11)

Where,
\[ r_k = \frac{1}{I} \sum_{i=1}^{I} M\left( r_{k-1} + \beta_{k-1} E_k\left( w^{(1)} \right) \right), k = 2, \ldots, N \]  

(12)

Constants \( \beta_k = \varepsilon_k \text{std} (r_k) \) are dedicated to gain an expected signal-noise ratio (SNR) between the added noise and its residue.

- The method based on wavelet transform and CEEMDAN

CEEMDAN is a non-linear, non-stationary signal processing method, which has a stronger anti-noise, higher decomposition efficiency, as well as the better suppression in mode mixing than EEMD.

CEEMDAN decomposes the signal into a series of intrinsic mode functions (IMFs) which frequencies from high to low and selects the effective component to reconstruct.

The modal components obtained by the decomposition of the signal often contain pseudo-modal components which are independent of the original signal and cannot reflect the characteristics of the original signal. Therefore, the signal decomposition needs to identify and remove the pseudo-modal components. In this paper, we use the correlation coefficient method to select the effective mode to draw the envelope spectrum.

The correlation coefficient method can well reflect the degree of linear correlation between each modal variable and the original signal, and can quantitatively represent the degree of correlation between two random variables.

The formula for the correlation coefficient is as follows:
\[ r_{xy} = \frac{\sum_{i=1}^{N} \left( x_i - \bar{x} \right) \left( y_i - \bar{y} \right)}{\sqrt{\sum_{i=1}^{N} \left( x_i - \bar{x} \right)^2} \sqrt{\sum_{i=1}^{N} \left( y_i - \bar{y} \right)^2}} \]  

(13)

Rolling element bearing fault characteristics extraction steps and its flow chart are as follows.

Figure 1: Flow chart of the fault diagnosis principle of the proposed method

Step 1. The original signal is processed by wavelet transform to obtain the preprocessed signal.
Step 2. Decompose the preprocessed signal by improved CEEMDAN and obtain a series of IMFs.
Step 3. Find the correlation between each IMF and the original signal, and preserve the components which correlation values are larger than the given threshold.
Step 4.1. Draw the envelope spectra of each IMF component.
Step 4.2. Reconstruct the chosen IMFs in Step 3.
Then perform envelope spectrum analysis of reconstructed signals.
Step 5. Judge whether there is a fault and where the fault location is based on the fault characteristics in the envelope spectrum.

3. Simulation Illustration

To verify the effectiveness of the method based on CEEMDAN, a simulation signal \( y(t) \) is implemented here.
\[
y(t) = y_1 + y_2 + y_3 + y_4
\]  

(14)

\[
\begin{align*}
y_1(t) &= 12 \sin \left( 2\pi 35t - 2\pi/5 \right) \\
y_2(t) &= 7 \sin \left( 2\pi 13t - 3\pi/5 \right) \\
y_3(t) &= 4 \sin \left( 2\pi 5t - 4\pi/5 \right)
\end{align*}
\]  

(15)
where, $y(t)$ is a random noise with an amplitude of 0.2. Simulation signal, noise and their mixture are plotted in Figure 2.

![Figure 2: Time domain waveform of simulation signals and its components](image)

In order to compare the decomposition performance between EEMD and the improved CEEMDAN, we use absolutely the same parameters, i.e., the noise amplitude is 0.1 and the overall size is 100. Set the sampling frequency to 1000 Hz and the sampling time to 2 seconds. The decomposition results are shown in Figure 3 and Figure 4, respectively.

Figure 2 and Figure 3 reflect the time domain waveform of a series of IMFs decomposed by EEMD and CEEMDAN, respectively. IMF1 denotes a random signal that is somewhat contaminated by the added noise, and IMF2-IMF5 represent the different components of the simulation signal. We can view that the IMFs decomposed by EEMD have some mode mixing, and the added Gaussian white noise is not completely neutralized. However, it can be seen from Figure 4, CEEMDAN basically eliminates the phenomenon of mode mixing, the added Gaussian white noise is completely neutralized. In addition, in the case of the same parameters, CEEMDAN consumes a shorter time, greatly improving the efficiency of the operation.

![Figure 3: Time domain waveform of the decomposition of EEMD](image)

Through the above simulation signal, it can be proved that compared with EEMD, CEEMDAN has the better denoising effect on the noise-contaminated signal, can better eliminate the mode mixing, better neutralize the Gaussian white noise and improve the efficiency of the operation.

4. Experimental Demonstrations

To further validate the proposed method, we used real motor bearing data from Case Western Reserve University.

The vibration data was picked up by an accelerometer at the 6 o’clock position of the motor housing drive end at a sampling frequency of 12 KHz. A single point of failure was handled in the bearing using EDM, with a fault diameter of 0.021 inches and a fault depth of 0.011 inches.

In addition, the motor speed is 1797 rpm.

The characteristic frequency of the bearing is shown in Table 1, which is calculated by the formula (16) and the formula (17).

**Formula of outer race frequency:**

$$f_o = \frac{1}{2} \left( 1 - \frac{d}{D} \cos \alpha \right) f_z$$

(16)

**Formula of inner race frequency:**

$$f_i = \frac{1}{2} \left( 1 + \frac{d}{D} \cos \alpha \right) f_z$$

(17)

Where $d$ and $D$ represent the diameter of the sphere and the diameter of the pitch circle, respectively. $Z$ is the number of rolling elements, and $\alpha$ is the contact angle of the bearing.

**Table 1. Fault characteristic frequency of bearing**

<table>
<thead>
<tr>
<th>Shaft frequency $f_o$</th>
<th>Outer race frequency $f_o$</th>
<th>Inner race frequency $f_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>29.32Hz</td>
<td>107.32Hz</td>
<td>162.29Hz</td>
</tr>
</tbody>
</table>

The raw outer race vibration signals and its spectrums are shown in Figure 5. As can be seen
from Figure 5, due to the signal mixed with serious noise, making the spectrum of high-frequency components increased, it is difficult to identify the fault characteristic frequency.

Figure 5: The raw vibration signals of the motor bearing with an outer race fault

After decomposing the signals with EEMD and CEEMDAN, the correlation coefficient between each IMF component and the original signal is calculated by using the correlation coefficient method, and then the first five layers of IMFs are selected to draw the envelope spectrum.

To fairly compare the decomposition efficiency of EEMD and CEEMDAN, the ensemble size and the noise amplitude are absolutely the same, respectively, 100 and 1.5.

The wavelet transform uses the db10 wavelet basis for 16 layers decomposition.

Figure 6: The time-domain diagram of first five layers of IMFs decomposed by EEMD

Figure 7: The time-domain diagram of first five layers of IMFs decomposed by CEEMDAN

Figure 8: The envelope of the first five layers of IMFs decomposed by EE

As can be seen from the comparisons, the extraction effect of Figure 9 is significantly better than that of Figure 8.

Thus the effectiveness of the proposed method is verified.
Research on Fault Feature Extraction of Rolling Bearing based on Improved CEEMDAN

To quantify the superiority of the proposed method, the signal-noise ratio (SNR) and the root mean square error (RMSE) are used to evaluate the denoising effect.

The signal-noise ratio reflects the purity of the signal, the larger the SNR, the stronger the signal purity.

RMSE reflects the degree of deviation of the measured data from the true value; the smaller the RMSE, the higher the measurement accuracy.

The relative formulas are as follows.

\[ S_{NR} = 10 \times \log \left( \frac{\sum_{i=1}^{N} S_i^2}{\sum_{i=1}^{N} (S_i - S_i')^2} \right) \]  \hspace{0.5cm} (18)

\[ R_{mse} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (S_i - S_i')^2} \]  \hspace{0.5cm} (19)

In the outer race fault diagnosis, we compare the signal-noise ratio and root mean square error of the first five layers of IMFs generated by the two methods to evaluate the denoising effect.

As is shown in Figure 10 and Figure 11, the proposed method is superior to EEMD-based methods in terms of denoising.

Further analysis of the bearing inner race fault signal is carried out. The ensemble size and the noise amplitude are absolutely the same, respectively, 200 and 1.5. The wavelet transform uses the db10 wavelet basis for 16 layers decomposition. The raw inner race vibration signals and its spectrums are shown in Figure 12. Similar to the inner race, it is difficult to identify the fault characteristic frequency. Therefore, reconstruct the selected first five IMFs based on EEMD and CEEMDAN, respectively. Then perform envelope spectrum analysis of reconstructed signals and identify bearing characteristic frequency and its harmonics. The analysis results obtained by the method base on EEMD and CEEMDAN are shown in Figure 13 and Figure 14. Both methods can extract the fault characteristic frequency at fn, and its harmonics can be identified visually as well. To verify the advantages of the proposed method, we use the SNR and RMSE to quantify the denoising effect in Figure 15.
5. Conclusions

Even though the combination of wavelet transform and EEMD has the capability in detecting the faults of bearing, it also shows poor effect in identification of fault-related characteristics.

Aiming to improve the drawbacks of the method based on EEMD, a new improved CEEMDAN is used in this article.

The imitate illustration verified the method based on CEEMDAN can decrease the residual noise and improve the accuracy of decomposition of IMFs.

Two experiments were carried out on the outer and inner rings of the rolling bearing of the motor, and the method was comprehensively verified.

The advantages of this method are further proved by the two indicators of signal-to-noise ratio (SNR) and root mean square error (RMSE).

As a result, the outer ring fault characteristics of the motor bearing are extracted by finding the fault frequency in the envelope spectrum of the first five layers of IMF.

In addition, the inner ring fault signature can be found in the envelope spectrum of the reconstructed signal.

In addition, a comparison is made between the proposed method and the EEMD-based method.

On the basis of comparison, it is concluded that this method can significantly improve the fault diagnosis ability of rolling bearings.

Glossary

\( \omega \) : The wavelet coefficients
\( \lambda \) : The given threshold
\( \eta(w) \) : The wavelet coefficient after applying the threshold
\( \text{sign} \) : The function of obtaining a value
\( \lambda_1 \) : The given threshold
\( N \) : The length of wavelet coefficients
\( \sigma_n \) : Standard deviation of noise
\( \lambda_2 \) : The given threshold
\( \omega_b \) : Risk function
\( \sigma_n \) : Standard deviation of noise
\( M(\cdot) \) : The operator which produces the local mean of the signal
\( E_k(\cdot) \) : The operator which produces the \( k \)th mode decomposed by EMD
\( y(t) \) : A simulation signal
\( d \) : Ball diameter
\( D \) : Pitch diameter
\( Z \) : The number of rolling elements
\( \alpha \) : The bearing contact angle
**Acknowledgements**

The research is funded partially by the Agricultural Science and Technology Independent Innovation Fund of Jiangsu Province (CX(19)3081) and the Key Research and Development Program of Jiangsu Province (BE2018127).

**References**


