Fault Diagnosis of Gearbox Based on Complementary Ensemble Empirical Mode Decomposition and Kernel Fuzzy Clustering Algorithm

Hao Guo*

Inner Mongolia Power (Group) Co., Ltd, Hohhot Power Supply Branch, Hohhot, Inner Mongolia, China

**Corresponding Author.*

Abstract: A bearing fault diagnosis method based on complementary ensemble empirical mode decomposition (EEMD) and kernel fuzzy c-means (KFCM) algorithm is proposed to address the difficulties in feature extraction and fault diagnosis of wind turbine gearbox vibration signals. Based on empirical mode decomposition method, complementary ensemble empirical mode decomposition is proposed for the decomposition of gearbox vibration signals, obtaining multiple intrinsic mode functions. By calculating the sample entropy of the intrinsic mode function components as feature vectors, the kernel fuzzy c-means algorithm is used to achieve gearbox fault diagnosis. The experimental results show that the proposed method can effectively identify gearbox faults. In order to verify the progressiveness of the proposed method, the proposed method is compared with other methods. The experimental results show that the proposed method has higher fault diagnosis accuracy, which verifies the progressiveness of the proposed method.

Keywords: Gearbox; Fault Diagnosis; CEEMD; KFCM; Intrinsic Mode Function

1. Introduction

The gearbox in wind turbines is an important mechanical component, and the gearbox mechanism is the most widely used transmission device in modern machinery, with advantages such as compact structure, high reliability, and high transmission efficiency. As an important mechanical component of wind turbines, the gearbox is of great significance for fault diagnosis[1-3].

The vibration signal of the gearbox is a typical nonlinear signal and has strong non stationarity. At present, traditional signal analysis methods include wavelet analysis, empirical mode decomposition, ensemble empirical mode decomposition, etc. The effectiveness of wavelet analysis methods in signal processing is limited by the selected wavelet basis function, so the effect is not stable[4]. Although the empirical mode decomposition method does not require the selection of wavelet basis functions, it has strong mode aliasing and endpoint effects[5]. Ensemble Empirical Mode Decomposition (EEMD) has a certain suppression effect on mode aliasing by adding white noise to the original signal, but it does not completely eliminate mode aliasing and endpoint effects[6-9]. Complementary ensemble empirical mode decomposition is an improved algorithm based on EMD, which is not only suitable for processing non-stationary and nonlinear signals, but also effectively suppresses mode confusion caused by EMD methods. Compared with ensemble empirical mode decomposition, its computation time is shorter. Based on this, this article applies the Complementary Ensemble Empirical Mode Decomposition (CEEMD) method to feature extraction of gearbox vibration signals.

Due to the fuzziness of gearbox fault characteristics, this paper introduces the fuzzy c -means clustering algorithm into the fault diagnosis of gearboxes. Using the kernel fuzzy c -means (KFCM) clustering algorithm with better clustering performance combined with sample entropy, clustering diagnosis of gearbox faults is carried out.

2. CEEMD

Add noise signals to the original signal in two different ways to obtain two new signals.

$$\begin{cases} c_m^{+}(t) = x(t) + n_m(t) \\ c_m^{-}(t) = x(t) + n_m(t) \end{cases} \tag{1}$$

Where, $n_m(t)$ is the added white noise, and $c_m^+(t)$ and $c_m^-(t)$ are the signals after adding white noise. The Intrinsic Mode Function (IMF) components obtained by decomposing the $c_m^+(t)$ and $c_m^-(t)$ signals using EMD method are:

$$\begin{cases} c_1^+(t) = \sum_{j=1}^{h} c_{j,1}^+(t) \\ c_1^-(t) = \sum_{j=1}^{h} c_{j,1}^-(t) \end{cases} \quad (2)$$

In the equation, $c_1^+(t)$ and $c_1^-(t)$ are the signals of the original signal after the first addition of white noise. After EMD decomposition, h IMF components of these two signals are obtained. The IMF obtained by decomposing the signal containing added and subtracted noise into a set of IMF components is:

$$c_m(t) = \sum_{j=1}^{h} \frac{(c_{j,m}^+(t) + c_{j,m}^-(t))}{2} \quad (3)$$

The vibration signal $x(t)$ extracted from the rolling bearing can be expressed as the sum of each IMF component and the residual $r(t)$.

$$x(t) = \sum_{j=1}^{m} c_j(t) + r(t) \quad (4)$$

3. Sample Entropy

Sample entropy is an improvement on the approximate entropy algorithm and is currently a widely used method for calculating entropy eigenvalues[10-12]. Sample entropy is also a measure of time series complexity. The smaller the sample entropy value, the lower the time series complexity and the higher the self similarity. Usually used in the field of fault diagnosis. The calculation steps for sample entropy for the original vibration signal with $x = \{x_1, x_2, x_3, \cdots x_j\}$ and $j = 1,2,\cdots,L$ are as follows:

1. Form a set of ordinal numbers into a phase space with dimension m, resulting in state vectors $x_i = (x_i, x_{i+1}, \cdots x_{L-m+1})$ and $i = 1,2,\cdots,L-m+1$.
2. Calculate the distance (the one with the largest difference between the two corresponding elements).

$$d_{ij} = \max(|x_{i+1} - x_{j+1}|) \quad (5)$$

Where, $i, j = 1,2,\cdots L-m$ and $i \neq j$.

3. Assuming a threshold ($r>0$) is set, the unit step function $H(\gamma - d_{ij})$ can be used to calculate the proportion of state vectors that are similar to the given state vector x_i.

$$B_l^m(r) = \frac{\sum jH(\gamma - d_{ij})}{L-m-1} \quad (6)$$

4. Solve for all $B_l^m(r)$ and then take the average of them.
5. Increase the dimension to $m+1$ and repeat steps 1 to 3 to obtain $B_i^{m+1}(r)$.
6. When L is a finite value, the sample entropy of the signal sequence is obtained as:

$$SampEn(m,r) = \ln Bm(r) - \ln Bm + 1(r) \quad (7)$$

4. KFCM Algorithm

The KFCM clustering algorithm changes the distance function of fuzzy c-means clustering through the kernel function, and its objective function is defined as[13,14]:

$$J_{kfcm}(U,V) = \sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^m \left\| \varphi(x_j) - \varphi(v_i) \right\| = \sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}\left[2 - 2K(v_i, x_j)\right] \quad (8)$$

Where, v_i is the clustering center of the clustering data, c is the number of categories in the clustering, n is the number of samples participating in the clustering, u_{ij} is the membership degree of the i-th category corresponding to the j-th sample, and $0<u_{ij}<1$, m is the fuzzy index. In this paper, we take 2 and $K(v_i, x_j)$ as Gaussian radial basis functions.

$$K(v_i, x_j) = e^{-\|xj - vi\|/2\delta^2} \quad (9)$$

Regarding membership degree, the following constraints must be met:

$$\sum_{i=1}^{c} u_{ik} = 1 \quad (k = 1,2,\cdots,n) \quad (10)$$

The iterative formula for membership degree and cluster center derived from the necessary conditions of Lagrange extremum is:

$$u_{ij} = \frac{[1 - K(x_j, v_i)]^{-1/(m-1)}}{\sum_{k=1}^{c}[1 - K(x_j, v_i)]^{-1/(m-1)}} \quad (11)$$

$$v_i = \frac{\sum_{j=1}^{n} u_{ij} K(x_j, v_i) x_j}{\sum_{j=1}^{n} u_{ij} K(x_j, v_i)} \quad (12)$$

The membership matrix is:

$$U(i, j) = u_{ij} \quad (13)$$

The specific steps of the algorithm are as follows:

1. Initialize U , Gaussian radial basis parameter σ ; The number of clusters c (in this article, $c = 4$, i.e. normal, broken teeth, worn, and pitting). Fuzzy index $m = 2$, convergence accuracy $\varepsilon = 0.0001$, iteration count $k = 0$; Initial cluster center matrix V .
2. Calculate U^K using equations (11) and (13).
3. Calculate v_i using equation (12), where $K = K + 1$.
4. Repeat steps 2 and 3 until termination condition $\left\| U^K - U^{K-1} \right\| < \varepsilon$ is met.

5. Fault Diagnosis Method for Gearbox Based on CEEME and KFCM

1. Firstly, collect the vibration signal of the gearbox.
2. Perform CEEMD decomposition on the collected vibration signal to obtain several IMF components, and select the first n order components with more fault features, where $n = 6$ in this paper.
3. Calculate the sample entropy of the first n IMF components and form a feature vector as the clustering sample.
4. Randomly select one set of samples from each of the four working states of the gearbox (normal, broken teeth, abrasion, and pitting), and use known fault samples as labels to determine the fault diagnosis results.
5. Cluster the samples using the KFCM method to obtain the final fuzzy membership matrix U . Cluster them using the maximum membership method.
6. Determine unknown fault types. Samples that are classified into the same category as the labels of known sample faults, and whose fault types are consistent with those of known samples.

6. Experimental Analysis

The data in this article comes from the fault diagnosis platform of a certain electrical group. The experimental platform consists of a fan gearbox, a variable speed drive motor, bearings, sensors, a speed regulator, and a spindle. Through installation and debugging, various faults of the gearbox can be simulated. Set the sampling frequency to 2000 × 2.56Hz, and collect 25 sets of normal, broken, worn, and pitting vibration signals each at a speed of 880r/min and a loading current of 0.1A through an accelerometer, for a total of 100 sets. Taking the vibration signal of a gearbox with broken teeth as an example, the acceleration waveform of the vibration signal is shown in Figure 1.

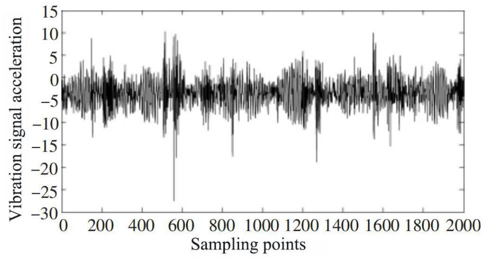

Figure 1. Acceleration Waveform of Vibration Signal

To verify the effectiveness of the method proposed in this article, both CEEMD and EMD methods were used to decompose the vibration signal of the gearbox, obtaining a series of IMF components. The effectiveness of modal decomposition was measured by three indicators: completeness, orthogonality, and the number of IMF components. The completeness index measures the reconstruction error of empirical mode decomposition, and a small index indicates good completeness of the reconstruction. Orthogonality is used to measure the degree of component mode mixing, and the smaller the value, the lower the degree of mode mixing and the better the decomposition effect. The specific values of the completeness, orthogonality, and number of IMF components of the EMD and CEEMD decomposition of gear box tooth breakage signals are shown in Table 1.

Table 1. Comparison of EMD and CEEMD Methods

Index	Completeness	Orthogonality	Number of IMF components
EMD	5.986	0.048	9
CEEMD	4.583	0.007	6

As can be seen from Table 1, the completeness of CEEMD decomposition of gearbox is slightly better than that of EMD, and the orthogonality index value of CEEMD decomposition is far less than that of EMD decomposition. The number of components of CEEMD is 6 and that of EMD is 9, which reduces the number of components. It can be shown that CEEMD algorithm is more complete than EMD algorithm, and it can also improve modal aliasing. The vibration signals of gearbox decomposed by EMD and CEEMD are obtained respectively, and the sample entropy of the first six components is calculated as the feature vector of feature recognition. The sample entropy of EMD and CEEMD under the same group of samples under four conditions of gearbox is shown in Table 2 and Table 3.

Table 2. IMF Component Sample Entropy of EMD Decomposition

Category	Normal	Broken teeth	Abrasion	Pitting
IMF1	1.638	1.684	1.719	1.673
IMF2	1.723	1.734	1.842	1.783
IMF3	1.142	1.231	1.431	1.012
IMF4	0.612	0.641	0.914	0.463
IMF5	0.351	0.315	0.412	0.385
IMF6	0.154	0.185	0.174	0.124

Table 3. Sample Entropy of IMF Components Decomposed by CEEMD

Category	Normal	Broken teeth	Abrasion	Pitting
IMF1	0.387	0.336	0.824	0.364
IMF2	0.156	0.136	0.378	0.188
IMF3	0.082	0.094	1.431	0.081
IMF4	0.032	0.051	0.091	0.053
IMF5	0.005	0.009	0.012	0.008
IMF6	0	0.001	0.004	0.002

The form of sample entropy from IMF1 to IMF6 is $[SampEn1, SampEn2, \cdots SampEn6]$. Taking CEEMD sample entropy as an example, the eigenvectors under normal conditions are [0.387, 0.156, 0.082, 0.032, 0.005, 0]. All feature vectors are calculated to form a feature vector set, that is, clustering samples are obtained.

This experiment is conducted using four methods for comparison, namely EMD-FCM, EMD-KFCM, CEEMD-FCM, CEEMD-KFCM. To verify the effectiveness of the CEEMD-KFCM algorithm, two different vibration signals were collected for experiments. Experiment 1 shows the vibration signal of the gearbox at a speed of 880r/min and a loading current of 1A, while Experiment 2 shows the vibration signal of the gearbox at a speed of 880r/min and a loading current of 0.05A. The experimental results are shown in Table 4.

Table 4. Comparison of Experimental Results

Category	EMD-FCM(%)	EMD-KFCM(%)	CEEMD-FCM(%)	CEEMD-KFCM(%)
Experiment 1	92	94	96	99
Experiment 2	87	93	99	100

According to Table 4, compared to EMD sample entropy, using CEEMD decomposed sample entropy as a feature vector can better reflect the fault characteristics of the gearbox. Compared to FCM algorithm, KFCM algorithm has better fault clustering effect. Among the four methods, CEEMD-KFCM algorithm has the best effect on gearbox fault diagnosis.

7. Conclusion

This article focuses on the nonlinear and non-stationary characteristics of the vibration signal of the wind turbine gearbox. The CEEMD algorithm is used to decompose the signal into IMF. By comparing with the traditional EMD algorithm, the superiority of this algorithm in gearbox fault extraction is verified. According to the principle of sample entropy, the sample entropy of IMF components is used as the feature vector for fault recognition, and its effectiveness in fault recognition is demonstrated through experiments. Aiming at the fuzziness of gearbox fault characteristics, the KFCM algorithm, which is more suitable for handling nonlinear problems, is introduced. Combined with the CEEMD algorithm, a gearbox fault recognition algorithm CEEMD-KFCM is proposed. After analyzing the experimental results, the effectiveness of the algorithm is verified, and a new idea is proposed for vibration signal processing and fault diagnosis of wind turbine gearboxes.

References

[1] Li J, Dong L, Zhang W W Z. Research on a novel fault diagnosis method for gearbox based on matrix distance feature. Measurement and Control: Journal of the Institute of Measurement and Control, 2024, 57(4):454-468.

[2] Chuanhui L, Xiaojing C, Xiaoxiao H, et al. Fault Diagnosis Analysis of Variable Working Condition Gearbox based on DWAE and GRUNN Combination Model. Journal of Mechanical Transmission, 2022,

46(2):155-159.
[3] Chen Y, Liang R, Ran W, et al. Gearbox fault diagnosis method based on deep learning multi-task framework. International Journal of Structural Integrity, 2023, 14(3):401-415.
[4] Jiaoyi, Wu, Xinyan, et al. A wavelet-analysis-based differential method for engine wear monitoring via on-line visual ferrograph. Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology, 2013, 227(12):1356-1356.
[5] Sheng, Fu, Lei, et al. Fault identification and severity assessment of rolling element bearings based on EMD and fast kurtogram. Journal of Vibroengineering, 2016, 18(6):3668-3683.
[6] Jiang F, Zhu Z, Li W, et al. Robust condition monitoring and fault diagnosis of rolling element bearings using improved EEMD and statistical features. Measurement Science & Technology, 2014, 25(2):1-14.
[7] Wang Y, Na X, Kang S, et al. State Recognition Method of a Rolling Bearing Based on EEMD-Hilbert Envelope Spectrum and DBN Under Variable Load. Zhongguo Dianji Gongcheng Xuebao/Proceedings of the Chinese Society of Electrical Engineering, 2017, 37(23):6943-6950.
[8] Amarouayache I I E, Saadi M N, Guersi N, et al. Bearing fault diagnostics using EEMD processing and convolutional neural network methods. International Journal of Advanced Manufacturing Technology, 2020, 107(4):1-19.
[9] Liu Y M, Zhao C C, Xiong M Y, et al. Assessment of bearing performance degradation via extension and EEMD combined approach. Journal of Central South University, 2017, 24(5):1155-1163.
[10]Wensheng S, Fengtao W, Hong Z, et al. Feature Extraction of Rolling Element Bearing Fault Using Wavelet Packet Sample Entropy. Journal of Vibration, Measurement & Diagnosis, 2011, 31(2):162-380.
[11]Chen Y, Pham T D. Sample entropy and regularity dimension in complexity analysis of cortical surface structure in early Alzheimer's disease and aging. Journal of Neuroscience Methods, 2013, 215(2):210-217.
[12]Song Y, Zhang J. Discriminating preictal and interictal brain states in intracranial EEG by sample entropy and extreme learning machine. Journal of Neuroence Methods, 2016, 257:45-54.
[13]Hu D, Sarosh A, Dong Y F. A novel KFCM based fault diagnosis method for unknown faults in satellite reaction wheels. ISA Transactions, 2012, 51(2):309-316.
[14]Jin L, Hao L, Yuxiu W. Performance Evaluation of Cotton Yarns based on KFCM and KFDA. International Journal of Advancements in Computing Technology, 2012, 4(11):414-422.

