

Research Article

Fault Identification of Rolling Bearing Using Variational Mode Decomposition Multiscale Permutation Entropy and Adaptive GG Clustering

Tianjing He , Rongzhen Zhao , Yaochun Wu , and Chao Yang 

School of Mechanical and Electrical Engineering, Lanzhou University of Technology, Lanzhou 730050, China

Correspondence should be addressed to Rongzhen Zhao; zhaorongzhen@lut.cn

Received 14 April 2021; Revised 25 July 2021; Accepted 23 August 2021; Published 8 September 2021

Academic Editor: Xavier Chiementin

Copyright © 2021 Tianjing He et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The nonlinear and nonstationary characteristics of vibration signal in mechanical equipment make fault identification difficult. To tackle this problem, this paper proposes a novel fault identification method based on improved variational mode decomposition (IVMD), multiscale permutation entropy (MPE), and adaptive GG clustering. Firstly, the original vibration signal is decomposed into a set of mode components adaptively by IVMD, and the mode components that are highly correlated with the original signal are selected to reconstruct the original signal. After that, the MPE values of the reconstructed signal are calculated as feature vectors which can differentiate machinery conditions. Finally, low-dimensional sensitive features obtained by principal component analysis (PCA) are fed into the adaptive GG clustering algorithm to perform fault identification. In this method, the residual energy ratio is used to find the optimal parameter K of the VMD and the PBMF function is incorporated into the GG to determine the number of clusters adaptively. Two bearing datasets are used to validate the performance of the proposed method. The results show that the proposed method can effectively identify different fault types.

1. Introduction

Rolling bearings are one of the most prevalent components in rotor system of rotating machinery. Faults may occur in bearings due to wear and tear. These faults may cause malfunctions and failure and may even lead to catastrophic failure of the rotating machinery. Thus, this has prompted a great deal of research on vibration-based fault identification of rolling bearing over the last few decades [1]. Rolling bearing fault identification mainly includes two steps: one is fault feature extraction using signal processing techniques, and the other is fault identification based on pattern recognition methods.

Vibration signals with abundant fault information are often complex, nonlinear, and nonstationary. These characteristics make it difficult for the traditional Fourier transform-based methods to achieve good analysis results [2]. Empirical mode decomposition (EMD) can adaptively extract the intrinsic mode components of the signal

according to its local structural characteristics on the time scale [3]. It is not limited by the uncertainty principle and is particularly suitable to analysis nonlinear and nonstationary signals. With in-depth research, however, it is found that EMD has some obvious shortcomings, such as sensitivity to noise and sampling, mode mixing, endpoint effect, lack of mathematical theory, and being time consuming [4, 5], which significantly limit its further applications. To overcome these deficiencies, variational mode decomposition (VMD) method, which is essentially different from EMD recursive mode decomposition, was proposed by Dragomiretskiy and Zosso [5]. The VMD uses nonrecursive mode decomposition to suppress the mode mixing caused by the accumulation of envelope estimation error. It solves the problem that the components with similar frequency cannot be decomposed effectively and dramatically avoids the endpoint effect. Some successful applications of VMD have been reported recently. Xing et al. [6] proposed a fault diagnosis method based on VMD, Tsallis entropy, and FCM

clustering for rolling bearing. Firstly, they applied VMD to decompose the vibration signals to obtain a series of intrinsic mode functions (IMFs). Then, they calculated the Tsallis entropy of each IMF and used it as fault features. Finally, the features were put into FCM classifier to recognize different fault types. In order to extract more abundant features to distinguish bearing faults, Zhu et al. [7] utilized VMD to decompose the vibration signals firstly. Then, they applied the continuous wavelet transform to draw the time-frequency images of the signals. Finally, those images were used as the input features of the Mobile-Net network to perform fault identification. Liu and Yan [8] decomposed the vibration signal of rolling bearing into a set of mode components using VMD and took the AR model parameters, model variance, and singular value of modal matrix of each mode component as the feature vectors. Finally, the Euclidean distance discriminant function was established to identify the fault types of the rolling bearing. Cui et al. [9] combined generalized morphological filtering and VMD to extract the features of bearing early fault signals collected from bearing and used envelope spectrum analysis to recognize the bearing fault types. Unfortunately, the number of mode components K needs to be set in advance when using VMD algorithm for signal processing. If the preset value of K is less than the number of useful components in the original signal, the decomposition will be insufficient and some intrinsic mode components cannot be accurately separated. On the contrary, if the preset value of K is greater than the number of useful components in the original signal, it will cause overdecomposition and produce some false components. It is obvious that the selection of K directly affects the result of signal processing and is the most important part of VMD algorithm.

Complexity is a nonlinear characteristic reflecting the nature of the system, and the complexity of signal is different under various operating conditions. Permutation entropy (PE) [10] is a nonlinear statistical measure for the complexity and regularity of signals, which has the properties of simple calculation and strong antinoise ability. These properties make it an ideal tool for analyzing vibration signals in fault diagnosis of rotating machinery. Yan et al. [11] applied PE to the feature extraction of vibration signals, the results of which show that the PE can effectively detect and amplify the dynamic changes of vibration signal and characterize the working conditions of rotating machinery.

However, similar to the traditional single scale nonlinear index, PE can only measure the complexity of signals (time series) on a single scale. To compensate for this deficiency, based on PE, multiscale permutation entropy (MPE) which has better robustness than PE was developed by Aziz and Arif [12] and was successfully applied to estimate the complexity and randomness of time series at different scales. For instance, Li et al. [13] used MPE to quantitatively analyze the vibration signals of rotating machinery at different scales to construct the original feature set. Then, they constructed a multichannel fusion convolutional neural network to extract features from PE of multiple scales for fault identification of rolling bearings. In railway transportation industry, Ye et al. [14] proposed a data-driven

method which combines multiscale permutation entropy and linear local tangent space alignment to diagnose the faults of vehicle suspension systems.

Clustering analysis is one of the most important research fields for pattern recognition. Fuzzy C-means Clustering (FCM) and Gustafson Kessel (GK) clustering are two fashionable approaches which have been used for fault identification [15, 16]. Nevertheless, these two approaches are based on the assumption that the analyzed data has a spherical shape, which is rarely achieved in practice. Hence, Gath Geva (GG) clustering analysis method was developed by Gath and Geva [17]. GG clustering analysis is an improvement method of FCM clustering and GK clustering algorithm, which introduces the fuzzy maximum likelihood estimation distance norm to measure the distance between data, so it can reflect data classes with different shapes and directions. Recently, the applications of GG clustering in the field of fault identification and health management have achieved good results. For instance, Wang et al. [18] proposed a method for rolling bearing degradation condition clustering based on multidimensional degradation feature and GG clustering. In their method, the GG clustering is introduced to divide different conditions during the degradation process. The results show that the GG clustering is able to cluster degradation condition of equipment such as bearings accurately and is able to cluster data samples with arbitrary shapes. In [19], a fault diagnosis method that combines empirical mode decomposition, permutation entropy, linear discriminant analysis, and GG clustering was introduced and successfully applied to the fault identification of rolling bearings.

It is true that the above methods have achieved good cluster recognition results. However, it is worth noting that the number of GG clusters must be given correctly beforehand when using GG clustering, which is very difficult in practical applications.

In summary, although VMD holds the potential to overcome the deficiencies of endpoint effect, mode mixing, and other issues, the number of mode components K needs to be set in advance when using it to process signal. Unfortunately, the actual signal is always complex and changeable, and K is usually difficult to determine. GG clustering can reflect data classes with different shapes and directions, but it needs to determine the number of clusters beforehand, which require extensive expert experience or relevant background knowledge.

To solve the above problems, a novel fault identification method is proposed. First, the energy criterion iteration stop condition is introduced into VMD, and an improved variable mode decomposition (IVMD) is developed in this paper; when decomposing the vibration signal, the number of K is automatically determined. Second, the evaluation index PBMFfunction is integrated into GG clustering, and an adaptive GG clustering algorithm is proposed to automatically determine the number of clusters. Last, a new fault identification method combining IVMD, multiscale permutation entropy, and AGG clustering is proposed and applied to analyze the vibration signals collected from bearings.

2. Theoretical Background

2.1. Improved Variational Mode Decomposition. The VMD algorithm redefines the intrinsic mode function (IMF) as an AM-FM signal and decomposes the original signal into a specified number of IMF components by constructing and solving a constrained variational problem. Assuming that the vibration signal is to be decomposed into K IMFs, the corresponding variational problem is constructed as follows.

- (1) For each IMF component $u_k(t)$, use Hilbert transform to obtain its analytic signal.

$$\left(\delta(t) + \frac{j}{\pi t}\right) * u_k(t). \quad (1)$$

- (2) Estimate a center frequency ω_k for each analytical signal and transform the spectrum of each analytical signal to the baseband using frequency shift.

$$\left[\left(\delta(t) + \frac{j}{\pi t}\right) * u_k(t)\right] e^{-j\omega_k t}. \quad (2)$$

- (3) Utilize the Gaussian smoothness index H^1 of the frequency shift signal to estimate the bandwidth of each IMF, and the final constraint variational problem can be expressed as

$$\begin{cases} \min_{\{u_k\}, \{\omega_k\}} \left\{ \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t}\right) * u_k(t)\right] e^{-j\omega_k t} \right\|_2^2 \right\} \\ \text{s.t.} \quad \sum_k u_k = f, \end{cases} \quad (3)$$

where $\{u_k\} = \{u_1, \dots, u_k\}$ are K IMFs decomposed by VMD and $\{\omega_k\} = \{\omega_1, \dots, \omega_k\}$ are the frequency centers of each component.

To obtain the optimal solution of equation (3), a penalty factor α and a Lagrange operator λ_t are introduced. Then, the extended Lagrange operator can be expressed as

$$\begin{aligned} L(u_k, \omega_k, \lambda) = & \alpha \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t}\right) * u_k(t)\right] e^{-j\omega_k t} \right\|_2^2 \\ & + \left\| f - \sum_k u_k \right\|_2^2 + \langle \lambda, f - \sum_k u_k \rangle. \end{aligned} \quad (4)$$

Applying alternating direction method of multipliers to constantly update each mode component u_k and its center frequency ω_k . Finally, the saddle point of equation (4) is the optimal solution of equation (3). All of u_k can be obtained from the frequency domain based on the following equation:

$$\hat{u}_k^{n+1}(\omega) = \frac{\hat{f}(\omega) - \sum_{i \neq k} \hat{u}_i(\omega) + \hat{\lambda}(\omega)/2}{1 + 2\alpha(\omega - \omega_k)^2}, \quad (5)$$

where \hat{u}_k^{n+1} can be regarded as the result of the current residual quantity $\hat{f}(\omega) - \sum_{i \neq k} \hat{u}_i(\omega)$ through Wiener filtering and ω_k^{n+1} represents the gravity center of k th IMF power spectrum, which can be updated through the following formula:

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |\hat{u}_k(\omega)|^2 d\omega}{\int_0^\infty |\hat{u}_k(\omega)|^2 d\omega}. \quad (6)$$

According to the principle of VMD [5], the number of K is required to be predefined when using VMD to process signal. However, the practical signal is always complex and changeable, and K is usually difficult to determine. Inspired by literature [20], a new method to determine K based on residual energy ratio is developed in this work. The energy of the original signal or mode component is calculated as

$$E = \sum_{i=1}^L x^2(i), \quad (7)$$

where E is the energy value of the signal, $x(i)$ represents the signal series (original signal or mode component), and L is the number of sampling points. When the value of K is different, the energy of the decomposed signal is different. In this study, the residual energy ratio perc is defined as

$$\text{perc} = \frac{|\sum_{k=1}^K E_k - E_x|}{E_x}. \quad (8)$$

According to research results in the literature [20], the residual energy ratio perc value becomes progressively smaller as K value increases when VMD decomposes the vibration signal. The perc value is minimum when the signal is completely decomposed, and if the K value continues to increase, the signal is overdecomposed to producing a fictitious component resulting in anperc increase. Based on this, a VMD parameter K selection method from the energy perspective named IVMD is proposed. The specific decomposition procedure of IVMD is shown in Figure 1.

2.2. Multiscale Permutation Entropy and Parameter Selection. MPE is defined as the permutation entropy of time series after multiscale coarsening. First, the original series is coarse-grained to construct the multiscale time series; then, calculate the permutation entropy at each scale. For a given one-dimensional time series $X = \{x(i), i = 1, 2, \dots, N\}$ with length N , the main calculation procedures of MPE can be summarized as follows:

- (1) Perform coarse-graining on the time series $X = \{x(i), i = 1, 2, \dots, N\}$ to obtain its coarse-grained series:

$$y^s(j) = \frac{1}{s} \sum_{i=(j-1)s+1}^{js} x(i), \quad j = 1, 2, \dots, \left[\frac{N}{s}\right], \quad (9)$$

where s is the scale factor, $s = 1, 2, \dots$, and $[\cdot]$ represents the rounding operation. When $s = 1$, the series after coarse-graining is still the original series.

- (2) Perform time reconstruction to the coarse-grained series:

$$Y_t^s = \{y_t^s, y_{t+\tau}^s, \dots, y_{t+(m-1)\tau}^s\}, \quad (10)$$

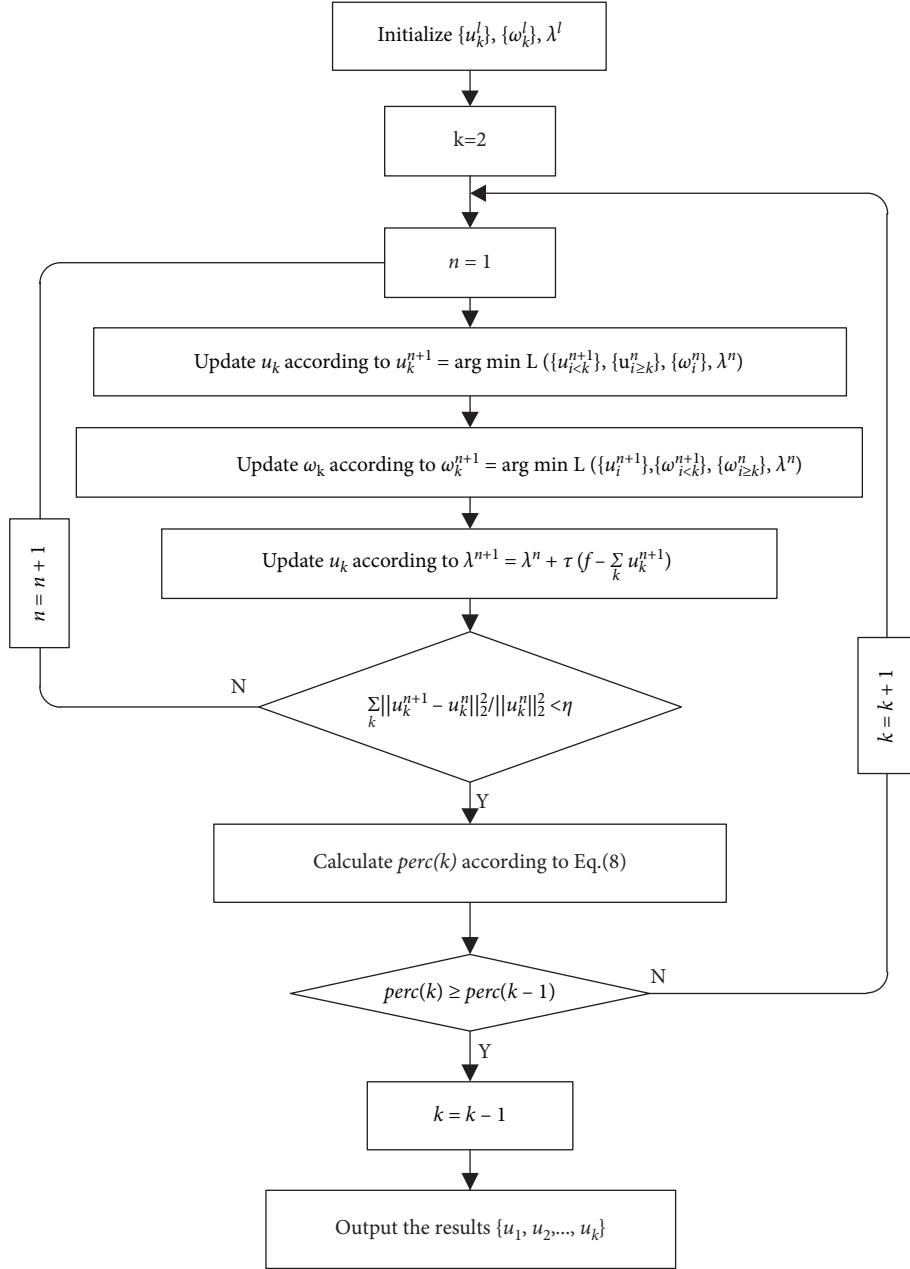


FIGURE 1: Flowchart of IVMD algorithm.

where τ is the time delay and m is the embedded dimension.

- (3) Calculate the permutation entropy of the time series under the scale factor s :

$$H_p^s(m) = - \sum_{j=1}^{m!} P_j^s \ln P_j^s. \quad (11)$$

- (4) Normalize the permutation entropies of different scales:

$$H_p^s = \frac{H_p^s(m)}{\ln(m!)}. \quad (12)$$

2.3. Adaptive GG Clustering. Given a cluster sample set $X = \{x_i | i = 1, 2, \dots, N\}$, each sample has D features, that is, $x_i = (x_{i1}, x_{i2}, \dots, x_{id})$. Suppose we want to cluster the samples in set X into c clusters ($c = 2, 3, \dots, N$). Let the clustering center vector $V = (v_1, v_2, \dots, v_c)$, and the membership matrix $U = [\mu_{ki}]_{c \times N}$, where $\mu_{ki} \in [0, 1]$ represents the membership degree of the i th sample to the k th class $k = 1, 2, \dots, c$; then GG clustering is realized by adjusting (U, V) iteratively to minimize the objective function J_m , which is given as

$$J_m(U, V) = \sum_{k=1}^c \sum_{i=1}^N (\mu_{ki})^m D_{ki}^2, \quad (13)$$

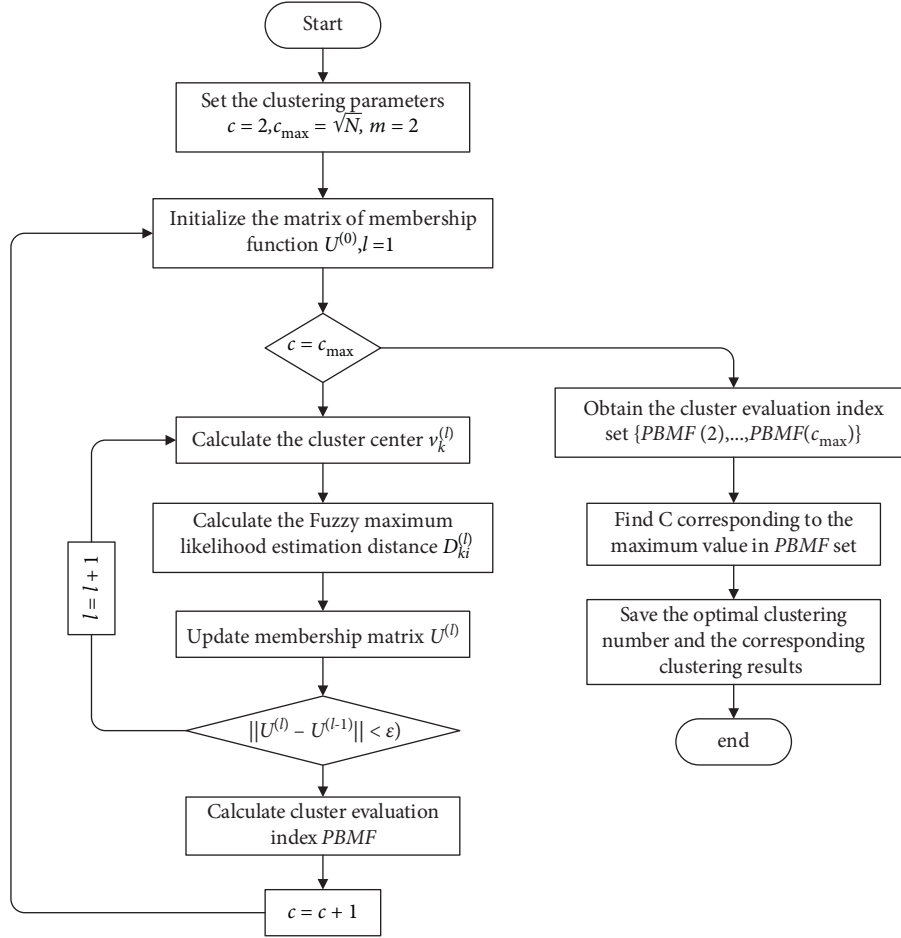


FIGURE 2: Flowchart of the AGG algorithm.

where m is the weighted index; the larger the value is, the more overlap there is between clusters. Generally, $m = 2$. The adjustment process of U , V can be described as follows:

(1) Calculate the cluster center:

$$v_k^{(l)} = \frac{\sum_{i=1}^N (\mu_{ki}^{(l-1)})^m x_i}{\sum_{i=1}^N (\mu_{ki}^{(l-1)})^m}, \quad k = 1, 2, \dots, c. \quad (14)$$

(2) Calculate the fuzzy maximum likelihood estimation distance:

$$D_{ki}(x_i, v_k) = \exp\left(\frac{1}{2}(x_i - v_k^{(l)})^T A_i^{-1}(x_i - v_k^{(l)})\right). \quad (15)$$

(3) Update the classification matrix:

$$\mu_{ki}^{(l)} = \frac{1}{\sum_{j=1}^c (D_{ki}(x_i - v_k) / D_{ji}(x_i - v_j))^{2/m-1}}, \quad i = 1, 2, \dots, N; \quad k = 1, 2, \dots, c. \quad (16)$$

If the condition $\|U^{(l)} - U^{(l-1)}\| < \epsilon$ is satisfied, terminate the iteration; otherwise, let $l \leftarrow l + 1$ and repeat the above steps until the condition is satisfied.

When using GG clustering, the number of cluster centers c is an important parameter that needs to be given in advance. However, it is usually very difficult or even impossible to achieve. In practical application, the number of c is usually determined by expert experience or relevant background knowledge. For the cluster centers c , it is suggested that

$c \in [2, \sqrt{N}]$ [21]; however, this requires several human-computer interactions and it is difficult to determine the optimal number of clusters. This inconvenience is also considered to be a major drawback of GG algorithm. Aiming at this problem, this paper proposes a method of adaptively determining clustering parameters and integrates it into the GG clustering algorithm. Pakhira et al. [22] proposed a clustering evaluation index function PBMF with clustering number c as independent variable, so as to solve the problem

of GG clustering requiring pre-given clustering number, which can be defined as

$$\text{PBMF}(c) = \left(\frac{E_1}{cJ_a} K_c \right)^2, \quad (17)$$

$$K_c = \max_{i,j=1}^c \|v_i - v_j\|, \quad (18)$$

where J_a is the compressibility measure within the class and is also the objective function of the clustering algorithm, K_c is the measure of separation between classes, and E_1 is the value of J_a when $c = 1$.

Analyzing (17) and (18), it is not hard to find that the larger the PBMF value is, the closer the clustering number is to the real cluster number, and consequently, the more accurate the clustering result will be. The flowchart of the adaptive GG clustering algorithm proposed in this paper is shown in Figure 2.

The clustering effect of the AGG clustering algorithm can be tested by clustering accuracy, classification coefficient, and average fuzzy entropy.

The clustering accuracy Acc is defined as

$$\text{Acc} = \frac{1}{N} \sum_{k=1}^c \theta_k, \quad (19)$$

where θ_k represents the number of samples correctly divided into class K .

The classification coefficient PC is defined as

$$\text{PC} = \frac{1}{N} \sum_{k=1}^c \sum_{i=1}^N \mu_{ik}^2. \quad (20)$$

The average fuzzy entropy CE is defined as

$$\text{CE} = -\frac{1}{N} \sum_{k=1}^c \sum_{i=1}^N \mu_{ik} \ln \mu_{ik}. \quad (21)$$

The closer the clustering accuracy and classification coefficient are to 1, and the closer the average fuzzy entropy is to 0, the better the clustering effect will be.

2.4. Principle Component Analysis. Principle component analysis (PCA) is a widely used technique in applications such as dimension reduction, feature selection, lossy data compression, and data visualization. By transforming the data into a lower dimension, PCA simplifies the complexity of high-dimensional data while retaining trends and patterns. It is an unsupervised feature transformation method, which requires no label data, and is completely nonparametric [23].

In general, the objective of the PCA transform is to maximize the variance of the projected data.

$$y_i = W^T x_i, \quad i = 1, 2, \dots, N, \quad (22)$$

where W is the PCA transform matrix which contains of the principal components, x_i is the original data ($d \times 1$), y_i

($k \times 1$) is the low-dimensional data projected onto the k ($k \ll d$) principal components.

3. The Proposed Method

The proposed method based on IVMD and AGG firstly uses IVMD to decompose the original vibration signal for denoising. This process makes full use of the mature theory of VMD and automatically determines the number of the mode components and then selects the most correlated mode components to reconstruct the original signal. Through decomposition and reconstruction, the main information of the signal is preserved and the large amount of noise contained in it is eliminated. Then it adopts MPE as feature vectors to reveal the different conditions of equipment. As mentioned above, MPE can effectively describe the granularity characteristics of the signal, which is more conducive to the feature quantization and extraction of the reconstructed signal. In addition, PCA is applied for dimension reduction to obtain some better feature vectors with low dimensionality, high sensitivity, and small classification error rate. Finally, the obtained feature vectors are fed into the AGG clustering algorithm to perform fault identification. The flowchart explaining the complete methodology is given in Figure 3. The exact procedures of the proposed algorithm can be expressed as follows:

- (1) Apply IVMD to decompose the original vibration signals of the equipment in various states, and obtain K mode components.
- (2) Calculate the cross-correlation between each mode component and the original signal, and select those mode components whose correlation coefficients are greater than the preset threshold to reconstruct the original signal.
- (3) Calculate the MPE of the reconstructed signal, and construct the original feature vectors.
- (4) Use PCA to reduce the dimension of the entropy feature vectors.
- (5) Take the low-dimensional sensitive feature vectors after dimension reduction as the input of adaptive GG clustering. Analyze the clustering effect by evaluation index.

4. Experiment and Analysis

4.1. Case Western Reserve University Bearing Dataset. To verify the effectiveness of the proposed fault diagnosis method (Figure 3), the bearing dataset provided by Bearing Data Center of Case Western Reserve University is analyzed with the proposed method [24]. This dataset contains 4 bearing states: normal condition (NR), inner race fault (IF), outer race fault (OF), and rolling element fault (BF). The 6205-2RS JEM SKF deep groove ball bearing is utilized in this experiment which is at the drive end. The damage of inner race, outer race, and rolling element was manufactured using electrodischarge machining (EDM), with a diameter of 0.1778 mm and a depth of 0.2974 mm, respectively. The

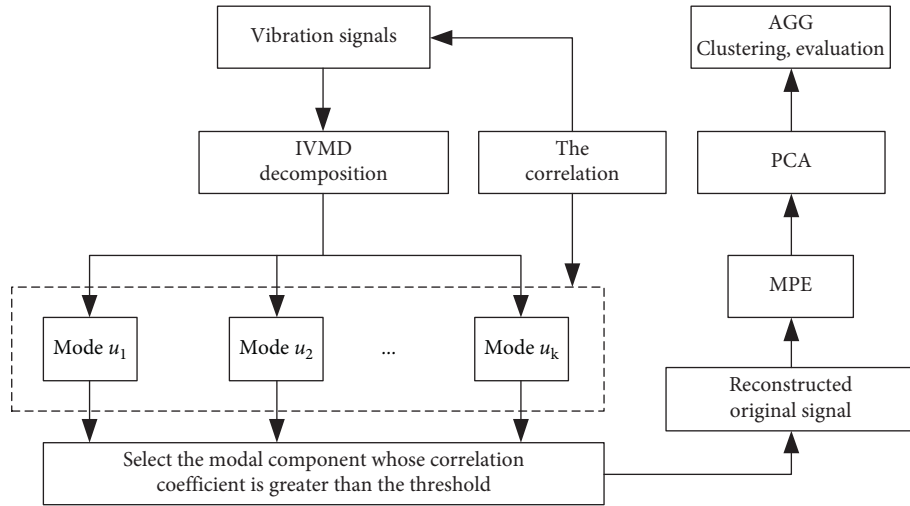


FIGURE 3: Flowchart explaining the complete methodology.

motor speed is set to be 1797 rpm with motor load about 1 hp. Under the sampling frequency of 12 kHz, 50 groups of vibration signals for each of the 4 states were collected with the length $N = 2048$.

Due to the layout, the outer race fault signals are randomly selected for analysis. On the basis of the IVMD (Figure 1) method proposed in Section 2.1, the acquired outer ring fault signal is adaptively decomposed, and the relationship between perc and K is shown in Figure 4. The perc value is minimum when K is 5, so it is appropriate to choose K to be 5. The original signals are adaptively decomposed into 5 IMF components, as shown in Figure 5. The correlation coefficient between each component and the original signal is calculated, which is displayed in Figure 6.

In this paper, correlation coefficient is used to select the intrinsic mode components. The larger the correlation coefficient, the higher the linear correlation degree between two variables. When the correlation coefficient is between 0.5 and 0.8, it is considered to be significantly correlated [25]. In this paper, components with correlation values greater than 0.5 are selected for reconstruction. More specifically, for the normal signals, IMF1 and IMF2 are selected; for the rolling element fault signals, IMF4 and IMF5 are selected; for the inner race fault signals, IMF4 and IMF5 are selected; for the outer race fault signals, IMF3 and IMF4 are selected.

The calculation of MPE is related to the selection of series length N , embedding dimension M , scale factor s , and time delay τ . It is noticeable that the series length N should satisfy the condition $N \geq 5M!$ [26] to obtain reliable statistics. Bandt and Pompe [10] proposed the permutation entropy method and indicated that the method works with embedding dimension $4 \leq M \leq 7$. In addition, Yan et al. [11] have discussed the validity of permutation entropy under different conditions of embedding dimension M . Obviously, when the value of M is too small, there will be too few states in the reconstruction of the original signal. It will bring about the algorithm to lose its physical meaning and effectiveness. On the contrary, if the value of M is too large, it

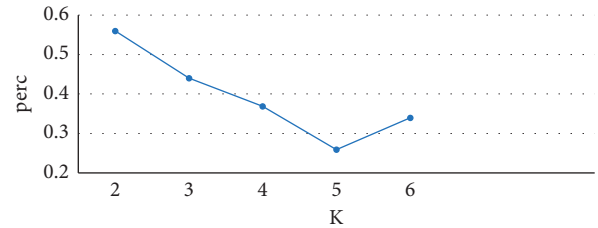


FIGURE 4: Relationship between the residual energy ratio perc and K .

will lead to be time consuming and the greater M will also cause considerable inconvenience for the followed multi-scale research because of the condition $N \geq 5M!$ [26]. For detecting the dynamic variation, embedding dimension M is usually selected by trade-off between information loss and computational complexity; in this paper, M is set as 5. Time delay τ has relatively small influence on the calculation of permutation entropy; we set $\tau = 1$. There is no clear criterion for the selection of the maximum s_m of scale factors. Generally, $s_m \geq 10$, and we set $s = 12$. When the parameters are determined, the permutation entropy of 12 coarse-grained vectors corresponding to the reconstructed signals in four states are calculated. The MPE of the reconstructed signals are shown in Figure 7.

Figure 7 illustrates that the MPE values of rolling bearings in different states are different. This is because the randomness of vibration signals is different when different fault emerges. This is a good indication that MPE can effectively characterize the fault characteristics of vibration signals in different scales. When $s = 1$, although the entropy values of normal state and inner race fault are obviously separated, the entropy values of rolling element and outer race fault are very close, which is difficult to distinguish. In order to get better experimental results, it is necessary to do multiscale analysis on the reconstructed signal. Because the first several entropy values represent the main information of vibration signal, we select the permutation entropy of the first nine scales as the feature vectors in this paper.

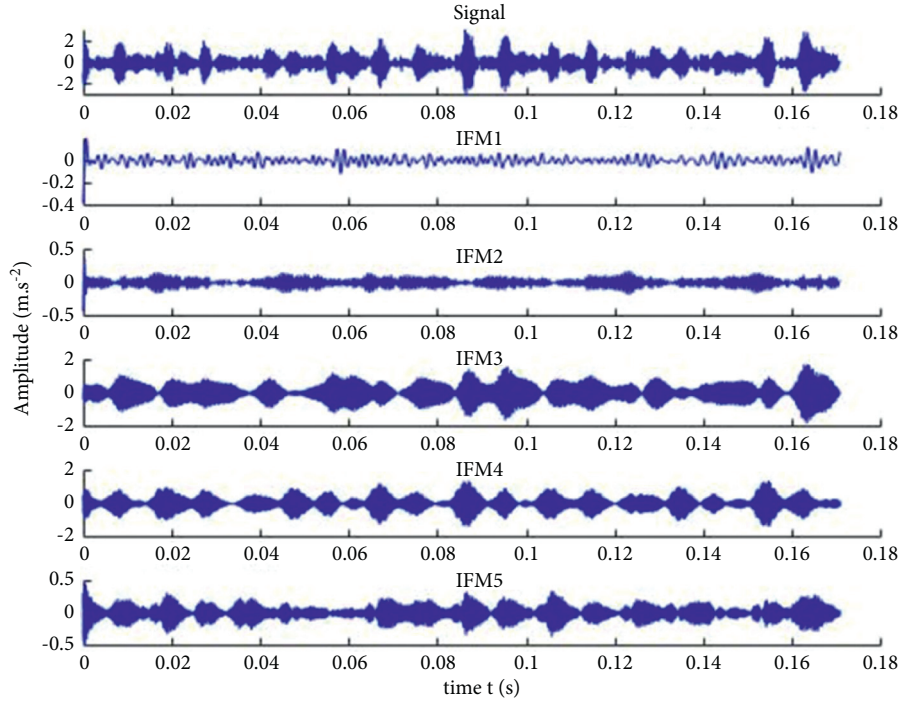


FIGURE 5: IVMD decomposition of outer race fault signals.

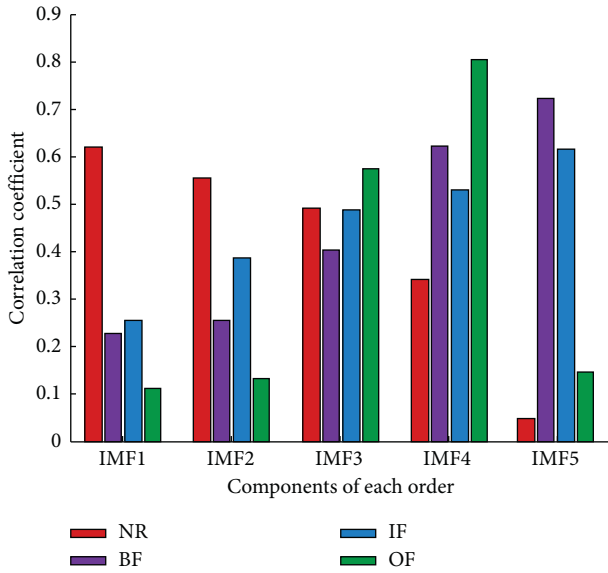


FIGURE 6: Correlation coefficients in different states.

The vibration signals of 4 different states are decomposed adaptively using IVMD, and high correlation IMFs are selected to reconstruct the original signals. The MPE of the reconstructed signal in each state are calculated. The first nine entropy values are selected as feature vectors to construct $4 * 50 * 9$ fault feature set. In order to visualize the data, PCA is used to reduce the dimensions of the feature vector from 9 to 2. Finally, input these low dimension vectors into the AGG algorithm to identify the optimal clustering number c and perform clustering analysis. Setting $c \in [2, 14]$ ($c_{\max} = \sqrt{200}$), then the recognition result of c is obtained as in Figure 8.

As shown in Figure 8, the value of clustering evaluation indicator function PBMFfirst increases and then decreases with the increase of cluster number c . When $c = 4$, the value of PBMF gets the maximum, indicating that the optimal clustering number of fault dataset is 4. This result is consistent with the actual situation, verifying the feasibility of AGG algorithm to automatically determine the optimal clustering number. Set $c = 4$, the fault feature sets are input into the AGG algorithm, and the clustering result is shown in Figure 9.

In Figure 9, v_1, v_2, v_3 , and v_4 are the clustering centers of rolling element fault, normal state, outer race fault, and inner race fault, respectively, and their coordinate values are listed in Table 1. As can be seen, the samples belonging to four different states of bearings are obviously separated and gathered near their clustering centers without aliasing; in addition, the distance of between-classes is large and the distance within-class is small. The average membership degree of each group is listed in Table 2. As can be seen, the average membership degree of the samples in normal state group is 1 to clustering center v_2 , and 0 to other three clustering centers, which means these samples belong to v_2 class completely and are far away from other classes. The average membership degree of the samples in inner race group is 1 to cluster center v_4 , and 0 to other three cluster centers, which means these samples belong to v_4 class completely and are far away from other classes. Similarly, for the samples in rolling element group, the average membership degree is 0.9725 to cluster center v_1 , and 0.0275 to cluster center v_3 , which are significantly higher than that to the other two cluster centers. It shows that v_1 and v_3 are close to each other, and this group of samples belong to v_1 . For the samples in outer race group, the average membership degree

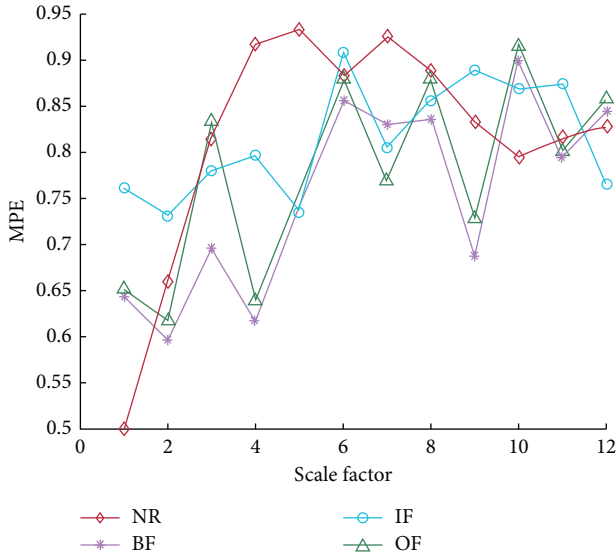


FIGURE 7: Multiscale permutation entropy of reconstructed signals.

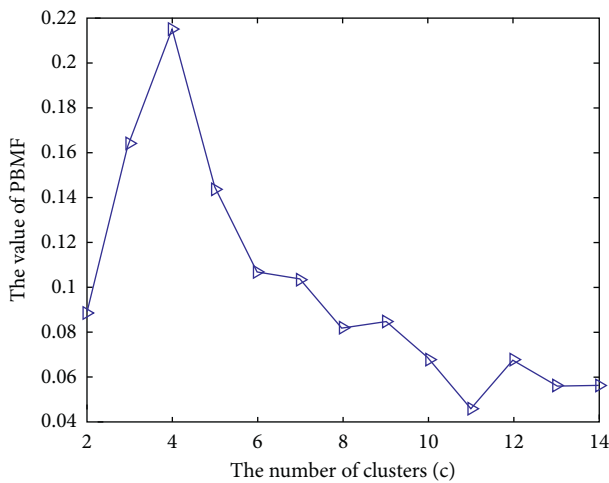


FIGURE 8: Changes of PBMF indexes with cluster number C .

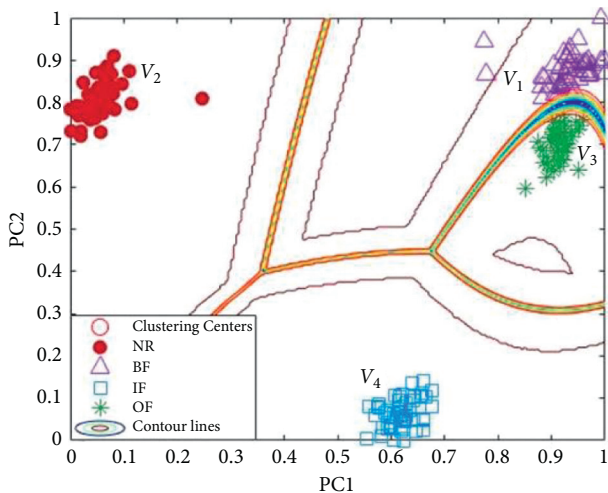


FIGURE 9: The contour map of AGG clustering.

TABLE 1: Clustering centers of 4 different state signals.

The clustering center	Coordinate values	
	x	y
V_1	0.9307	0.8735
V_2	0.0549	0.8052
V_3	0.9135	0.7031
V_4	0.6188	0.0638

TABLE 2: Membership degree of 4 different state signals.

Sample group no.	Membership			
	V_1	V_2	V_3	V_4
NR	0	1	0	0
BF	0.9725	0	0.0275	0
IF	0	0	0	1
OF	0.0028	0	0.9972	0

is 0.9972 to cluster center v_3 , and 0.0028 to cluster center v_1 . Obviously, this group of samples belong to v_3 . It is evident that AGG algorithm has excellent identification effect for fault features set.

In order to illustrate the superiority of the proposed method in this paper, three other fault identification methods including IVMD single scale permutation entropy and AGG clustering, IVMD multiscale permutation entropy and FCM clustering, IVMD multiscale permutation entropy and GK clustering are used for comparative study under the same datasets and identification task. The experimental results are shown in Figure 9. Through the comparative analysis of Figures 9 and 10, the following can be seen:

- (1) Compared with Figure 10(a), the overlap degree between different class samples in Figure 9 is lower, and this shows that the adaptive decomposition of the signal by IVMD eliminates most of the noise contained in the signal and preserves the main information of the signal.
- (2) The compactness of the within-class samples, by comparison, is better in Figure 9 than that in Figure 10(b). This shows that the multiscale permutation entropy can better represent the fault information of the signal, and taking it as the feature vector is more conducive to clustering identification.
- (3) Compared with Figures 10(c) and 10(d), the clustering effect in Figure 9 is better. Specifically speaking, clustering profiles of FCM and GK are more spherical and approximately elliptical, while AGG is arbitrary.

Clustering accuracy Acc, classification coefficient PC, and fuzzy entropy CE are used to quantitatively evaluate the performance of different identification methods in Figures 8 and 9. The results are listed in Table 3.

As can be seen from Table 3, under the precondition that the clustering methods are the same, the values of Acc and PC are 96.5% and 0.97 when MPE is used to directly extract fault features from vibration signals, and the values of Acc and PC are 98% and 0.9812 when IVMD + PE is

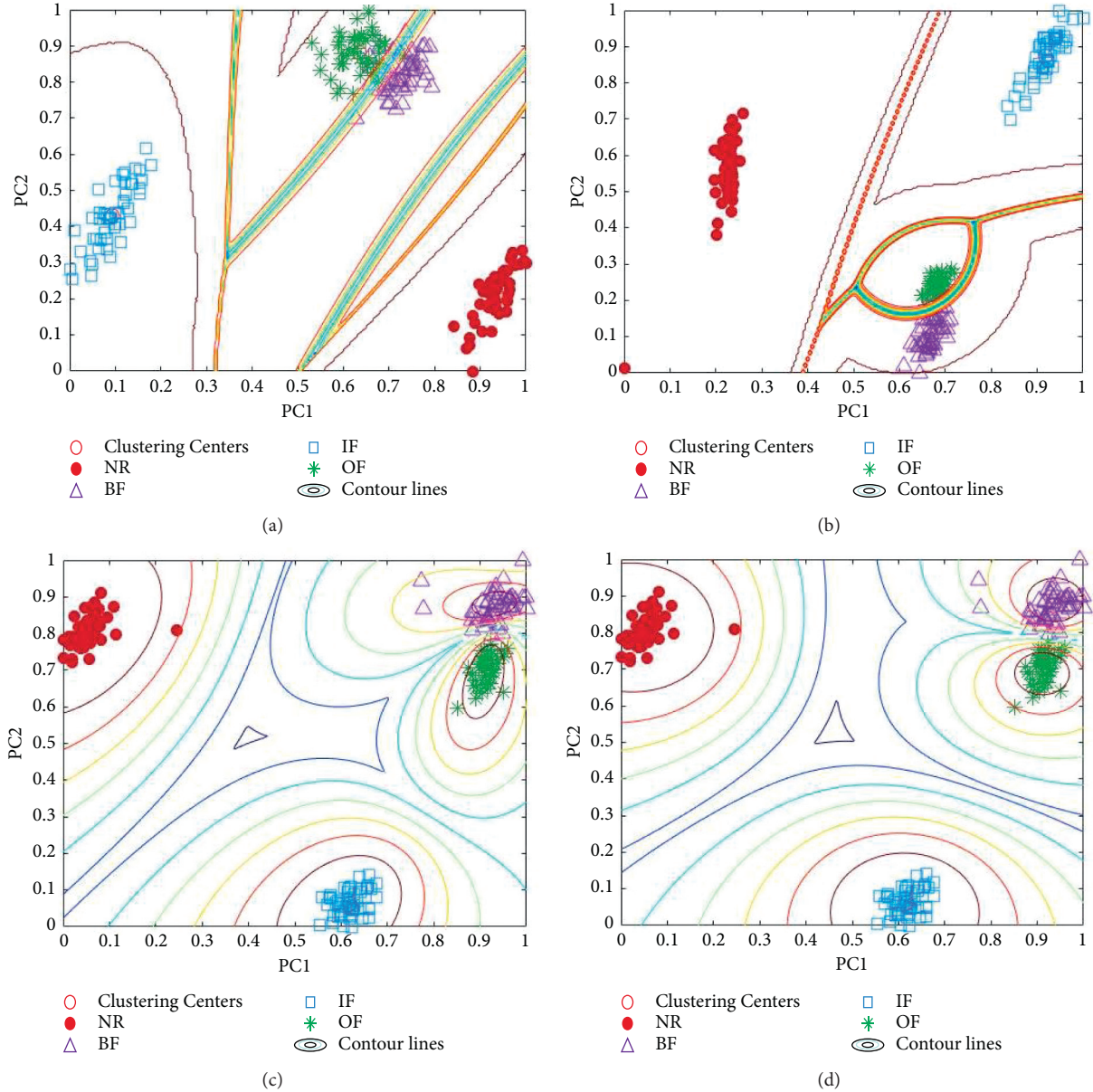


FIGURE 10: Clustering results of different combination modes. (a) Multiscale permutation entropy and AGG clustering. (b) IVMD permutation entropy and AGG clustering. (c) IVMD multiscale permutation entropy and FCM clustering. (d) IVMD multiscale permutation entropy and GK clustering.

TABLE 3: Fault diagnosis results.

Methodology	The evaluation index		
	Acc (%)	PC	CE
MPE + AGG	96.5	0.9709	NaN
IVMD + PE + AGG	98	0.9812	NaN
IVMD + MPE + FCM	98.5	0.9102	0.1774
IVMD + MPE + GK	100	0.9188	0.1675
Proposed method	100	0.9905	NaN

used for fault features extraction; however, the corresponding values of the proposed method are 100% and 0.9905. Our method has obvious advantages. Under the premise of the same feature extraction method, the values

of PC and CE are 0.9102 and 0.1774 separately when FCM clustering is used, and the values of PC and CE are 0.9188 and 0.1675, respectively, when GK clustering is applied; nevertheless, the values of PC and CE using the proposed method are 0.9905 and infinitesimal. It also has certain identification advantages. Therefore, the proposed method based on IVMD, MPE, and AGG clustering is effective.

4.2. *Laboratory Measured Datasets.* To further verify the performance of the proposed method, experiments are carried out on the measured dataset of Houde instrument laboratory.

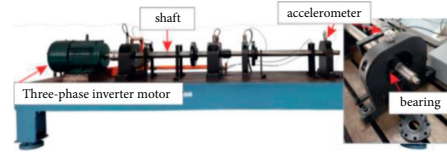


FIGURE 11: Rolling bearing test bench.

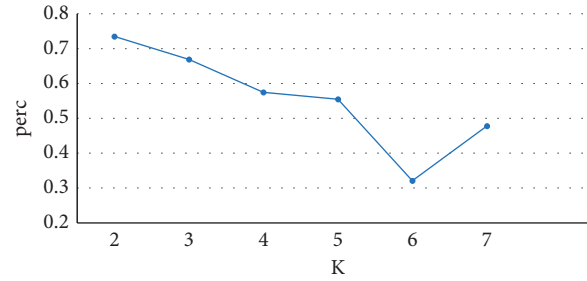
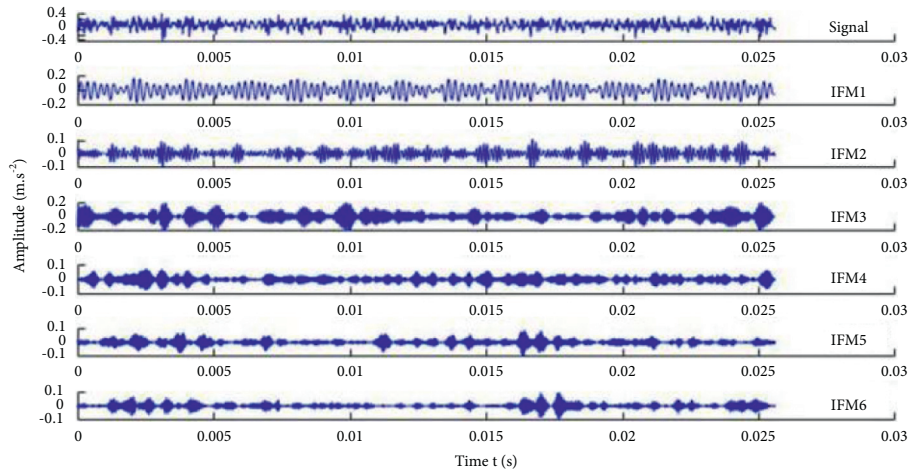
FIGURE 12: Relationship between the residual energy ratio perc and K .

FIGURE 13: IVMD decomposition of rolling element fault signal.

The bearing test bench used in the experiment includes three-phase inverter motor, rolling bearing mounting frame, rolling bearing, and loading device, as shown in Figure 11. The diagnosed object is NSK6308 deep groove ball bearing under 5 conditions, which are normal (NR), outer race peeling (OF), inner race peeling (IF), rolling element peeling (BF), and cage failure (FF). The test bearing speed is 3200 rpm, the type of the acceleration sensor is HD-YD232, and the style of the data acquisition card is HD9200. Under the sampling frequency of 8 kHz, 50 samples for each state were collected, with the sample length $N = 2048$ (Figure 11).

The rolling element fault vibration signals are randomly selected and adaptively decomposed by the proposed IVMD (Figure 1). The relationship between perc and K is shown in Figure 12. The perc is minimum when K is 6, so it is appropriate to choose K to be 6. When K is set to 6, the rolling element fault vibration signal is decomposed as in Figure 13. The correlation coefficients between each component and the original signal are calculated (Figure 14).

Similarly, when the correlation coefficient is between 0.5 and 0.8, it is considered to be significantly correlated [25]. Following this principle, those mode components with correlation coefficients greater than 0.5 are selected to reconstruct the original signal. More specifically, for the normal condition (NR), IMF1 and IMF2 are selected; for the rolling element fault condition (BF), IMF4 and IMF5 are selected; for the cage fault condition (FF), IMF1 is selected; for the inner race fault condition (IF), IMF5 and IMF6 are selected; for the outer race fault condition (OF), IMF5 is selected.

The same parameters are set as those in Section 4.1. The MPE values of reconstructed signal in 5 states are calculated, and the results are shown in Figure 15. Similarly, the first nine entropy values are selected as feature vectors to construct $5 \times 50 \times 9$ fault feature set.

Dimensions of entropy vectors are reduced from 9 to 2 by PCA, and these low dimension vectors are input into the AGG algorithm to identify the optimal clustering number c

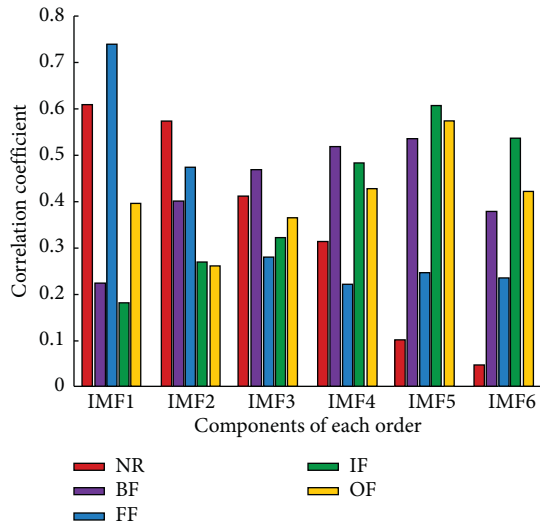


FIGURE 14: Correlation coefficients in different states.

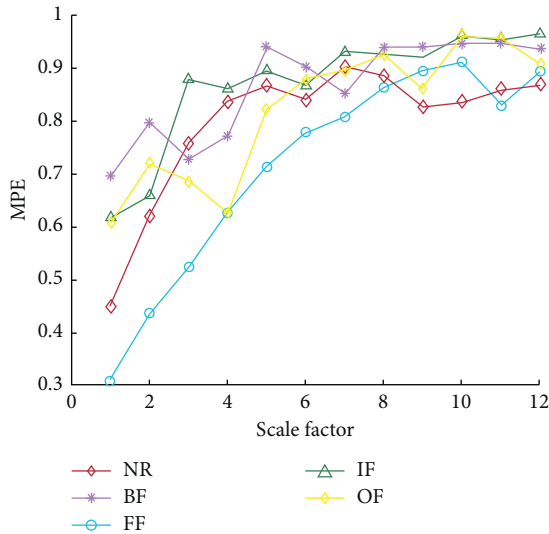


FIGURE 15: Multiscale permutation entropy of reconstructed signals.

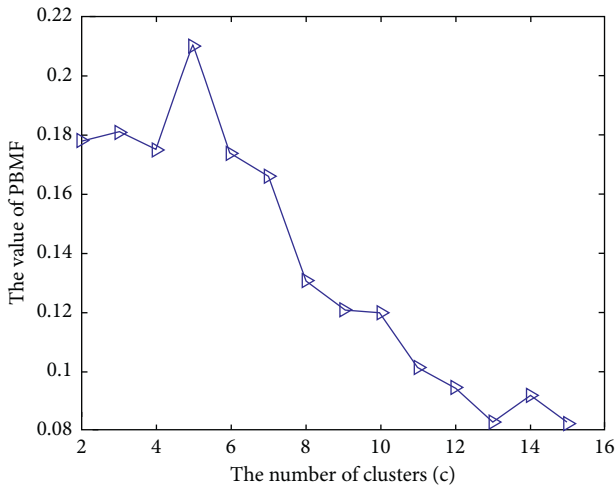


FIGURE 16: Changes of PBMF index with cluster number C.

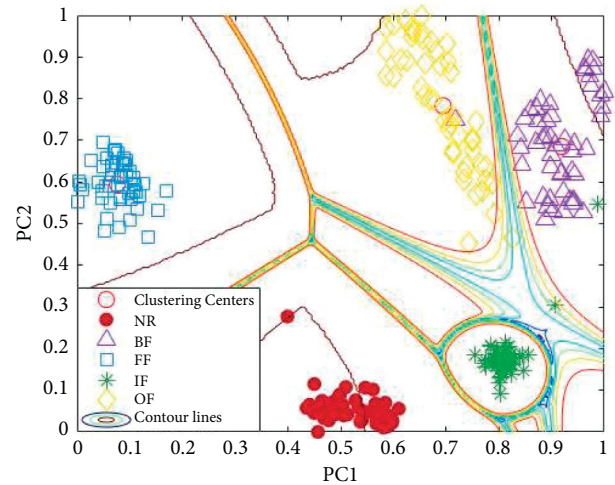


FIGURE 17: The contour map of AGG clustering.

and perform clustering analysis. Set $c \in [2, 15]$ ($c_{max} = \sqrt{250}$); the recognition result of c is shown in Figure 16. It can be seen that when the evaluation index function PBMF reaches the maximum value, the clustering number c is 5, which is consistent with the actual situation. This again proves the feasibility of the AGG algorithm to determine the optimal clustering number c adaptively. When $c = 5$, the result of cluster analysis is shown in Figure 17.

As can be seen from Figure 17, the samples of the same fault type are distributed around the same cluster center, and different types of faults are clearly separated. In addition, the between-classes distance is large and the within-class distance is small. Among them, one rolling element fault sample and two inner race fault samples are incorrectly clustered as outer race fault and rolling element fault separately. Therefore, the fault clustering identification rate is 98.8%. Furthermore, PC is 0.9945, and CE is infinitesimal. The experiment results on laboratory measured dataset once again demonstrated the effectiveness of the proposed method.

5. Conclusion

A new method based on improved variational mode decomposition, multiscale permutation entropy, and adaptive GG clustering algorithm is proposed for rolling bearing fault diagnosis. By introducing the energy criterion iteration stop condition, the improved variational mode decomposition effectively solves the problem that the number of mode components needs to be set in advance. By integrating the PBMF function, the AGG clustering algorithm can adaptively determine the optimal number of clusters. Firstly, IVMD is applied to adaptively decompose the vibration signal into several mode components, and the mode components with high correlation are selected to reconstruct the original signal. Then, the multiscale permutation entropy which can fully describe the irregularity and complexity of the signal is extracted from the reconstructed signals as the feature vectors. Finally, the extracted feature vectors are fed

AGG clustering algorithm for fault identification. The experimental results show that the proposed method can successfully distinguish the different operating states of the equipment, and there is no overlap between classes. Consequently, it is an effective method for fault information mining and identification.

Data Availability

All data included in this study are available upon request by contact with the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (no. 51965034) and Lanzhou City Innovation the Entrepreneurship Project (no. 2018-RC-25).

References

- [1] D. Chen, Y. Zhang, C. Yao, F. Sun, and N. Zhou, "Fault diagnosis based on FVMD multi-scale permutation entropy and GK fuzzy clustering," *Journal of Mechanical Engineering*, vol. 54, no. 1, pp. 16–27, 2018.
- [2] J. Zheng, H. Pan, and J. Chen, "Mean-optimized empirical mode decomposition and its application in rotor fault diagnosis," *Journal of Mechanical Engineering*, vol. 54, pp. 107–115, 2018.
- [3] 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. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998.
- [4] Z. Li and M. Zhu, "Research on mechanical fault diagnosis method based on variational mode decomposition," *Acta Armamentarii*, vol. 38, pp. 180–186, 2017.
- [5] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Transactions on Signal Processing*, vol. 62, no. 3, pp. 531–544, 2014.
- [6] T. Xing, Y. Zen, Z. Meng, and X. Guo, "A fault diagnosis method of rolling bearing based on VMD Tsallis entropy and FCM clustering," *Multimedia Tools and Applications*, vol. 79, no. 39, Article ID 30069, 2020.
- [7] J. Zhu, A. Deang, S. Xue, X. Ding, and S. Zhang, "Rolling bearing fault diagnosis and prediction based on VMD-CWT and MobileNet," *Computer Science & Information Technology*, vol. 11, no. 3, pp. 45–57, 2021.
- [8] C. Liu and X. Yan, "fault diagnosis of a wind turbine rolling bearing based on variational mode decomposition and condition identification," *Journal of Chinese Society of Power Engineering*, vol. 37, pp. 22–27, 2017.
- [9] X. Cui, H. Wang, J. Xing, and Y. Zuo, "Rolling bearing fault diagnosis of generalized morphological filtering and VMD decomposition," *Journal of Electronic Measurement and Instrument*, vol. 32, pp. 51–56, 2018.
- [10] C. Bandt and B. Pompe, "Permutation entropy: a natural complexity measure for time series," *Physical Review Letters*, vol. 88, no. 17, Article ID 174102, 2002.
- [11] 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.
- [12] W. Aziz and M. Arif, "Multiscale permutation entropy of physiological time series," in *Proceedings of the IEEE INMIC International Multi-topic Conference*, pp. 1–6, Karachi, Pakistan, December 2005.
- [13] H. Li, J. Huang, X. Yang, J. Luo, L. Zhang, and Y. Pang, "fault diagnosis for rotating machinery using multiscale permutation entropy and convolutional neural networks," *Entropy*, vol. 32, no. 8, 2020.
- [14] Y. Ye, Y. Zhang, Q. Wang, Z. Wang, Z. Teng, and H. Zhang, "Fault diagnosis of high-speed train suspension systems using multiscale permutation entropy and linear local tangent space alignment," *Mechanical Systems and Signal Processing*, vol. 138, Article ID 106565, 2020.
- [15] Y. Hu, S. Zhang, A. Jiang, L. Zhang, W. Jiang, and J. Li, "A new method of wind turbine bearing fault diagnosis based on multi-masking empirical mode decomposition and fuzzy C-means clustering," *Chinese Journal of Mechanical Engineering*, vol. 32, no. 46, 2019.
- [16] X. Wu, J. Zhu, B. Wu, J. Sun, and C. Dai, "Discrimination of tea varieties using FTIR spectroscopy and allied Gustafson-Kessel clustering," *Computers and Electronics in Agriculture*, vol. 147, pp. 64–69, 2018.
- [17] I. Gath and A. B. Geva, "Unsupervised optimal fuzzy clustering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 773–780, 1989.
- [18] B. Wang, X. Hu, T. Mei, S. D. Jian, and W. Wei, "Rolling bearing degradation condition clustering using multidimensional degradation feature and Gath–Geva fuzzy clustering algorithm," *Journal of Vibration and Control*, vol. 27, no. 17–18, 2020.
- [19] J. Hou, Y. Wu, H. Gong, A. S. Ahmad, and L. Liu, "A novel intelligent method for bearing fault diagnosis based on EEMD permutation entropy and GG clustering," *Applied Sciences*, vol. 10, no. 1, p. 386, 2020.
- [20] G. Tang and X. Wang, "Early fault diagnosis of rolling bearing based on IVMD and singular value difference spectrum," *Journal of Vibration, Measurement & Diagnosis*, vol. 36, pp. 700–707, 2016.
- [21] N. R. Pal and J. C. Bezdek, "On cluster validity for the fuzzy c-means model," *IEEE Transactions on Fuzzy Systems*, vol. 3, no. 3, pp. 370–379, 1995.
- [22] M. K. Pakhira, S. Bandyopadhyay, and U. Maulik, "Validity index for crisp and fuzzy clusters," *Pattern Recognition*, vol. 37, no. 3, pp. 487–501, 2004.
- [23] R. Machlev, D. Tolkachov, Y. Levron, and Y. Beck, "Dimension reduction for NILM classification based on principle component analysis," *Electric Power Systems Research*, vol. 187, Article ID 106459, 2020.
- [24] Bearing Data Center Website, The Case Western Reserve University Bearing Data Center Website Bearing Data Center Test Seeded Fault Test Data [EB/OL], 2013, <http://csegroups.case.edu/bearingdatacenter/pages/download-data-file>.
- [25] J. Lu, W. Ma, D. Ye, and C. Jiang, "Algorithm of sound signal enhancement based on VMD," *Journal of Mechanical Engineering*, vol. 54, no. 10, pp. 10–15, 2018.
- [26] Y. Li, M. Xu, Y. Wei, and W. Huang, "A new rolling bearing fault diagnosis method based on multiscale permutation entropy and improved support vector machine based binary tree," *Measurement*, vol. 77, pp. 80–94, 2016.