IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received April 14, 2021, accepted April 30, 2021, date of publication May 10, 2021, date of current version May 20, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3078823

A Fault Feature Extraction Method for Rolling Bearings Based on Refined Composite Multi-Scale Amplitude-Aware Permutation Entropy

YOUSHUO SONG AND WEIYU WANG

Department of Mechanical Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China

Corresponding author: Youshuo Song (wssys2006@163.com)

ABSTRACT Aiming at the problems of unclear early fault characteristics and difficult extraction of rolling bearings, a new nonlinear dynamic analysis method called refined composite multi-scale amplitude-aware permutation entropy (RCMAAPE) is introduced in this paper. Firstly, RCMAAPE is used to extract features from the bearing life data, and Chebyshev's inequality is used to establish a health threshold to evaluate the performance degradation state. Secondly, RCMAAPE is also used for bearing fault diagnosis. Both experimental results prove that RCMAAPE could extract fault characteristics effectively. RCMAAPE can accurately reflect the degradation trend of bearing in the whole life process, and is especially sensitive to the early failure of the bearing. RCMAAPE is also able to effectively identify the states of bearing faults. Especially after selecting the features, RCMAAPE only needs a small number of features to effectively identify the different states of the bearing and the recognition accuracy is up to 100%. Compared with the existing methods, the proposed method can extract fault features more effectively, has higher computational efficiency and obvious advantages.

INDEX TERMS Rolling bearing, refined composite multi-scale amplitude-aware permutation entropy, performance degradation assessment, fault diagnosis.

I. INTRODUCTION

Rolling bearing is widely used in rotating machinery and a common fault component, which has a great impact on the normal operation of machinery. Therefore, the performance degradation assessment and fault diagnosis of rolling bearings play a vital role in reducing catastrophic accidents and major economic losses, and can improve the reliability of rotating equipment. There are many ways to monitor fault signals of rolling bearings, such as vibration signals and acoustic emission (AE) signals [1], [2]. The vibration signals carry the dynamic information of the mechanical part itself, and the characteristic information of the bearing state can be effectively extracted from a certain method. Therefore, it is widely used in the extraction of rolling bearing fault features [3].

The most important step in the performance degradation assessment and fault diagnosis of rolling bearings is fault feature extraction. However, the vibration signal of the bearing is often non-linear due to the impact and load during the operation of the bearing. Therefore, researchers have conducted in-depth research on nonlinear dynamic analysis methods, and entropy method is one of them. Approximate entropy (ApEn), sample entropy (SampEn), fuzzy entropy (FuzzyEn) and permutation entropy (PE) have been widely used.

Pincus [4], [5] proposed approximate entropy (ApEn) to quantify the irregularity of data. Yan and Gao [6] first applied ApEn to fault diagnosis and operation state monitoring of rolling bearing. However, ApEn is highly dependent on the length of data and may cause inconsistency in the calculation results. To solve these shortcomings, Richman *et al.* [7] proposed sample entropy (SampEn), which reduced the dependence on the data length and was applied for the processing of medical and mechanical signals. Liang *et al.* [8] and Cheng *et al.* [9] combined ensemble empirical mode decomposition (EEMD) and SampEn for fault diagnosis of bearings and gears, respectively. Zhang *et al.* [10] combined the wavelet packet transform and SampEn for bearing fault diagnosis. In addition, Costa *et al.* [11] proposed multi-scale entropy (MSE), and MSE has also been widely used in fault

The associate editor coordinating the review of this manuscript and approving it for publication was Joanna Kołodziej.

diagnosis [12]–[17]. However, as the scale factor increases, the coarse-grained time series will rapidly become shorter, and the entropy value will undergo abrupt changes, and the error will increase. Moreover, MSE only analyzes the low-frequency components of the original signal and ignores the high-frequency components. In view of the shortcomings of MSE, different improvement methods have been proposed, and some of them were also introduced into the field of fault diagnosis [18]–[22].

In order to improve the shortcomings of SampEn, Chen *et al.* [23] used fuzzy functions instead of heaviside functions to measure the complexity of the data, thereby solving the problem of inaccurate estimates, but the computational efficiency dropped greatly. Fuzzy entropy (FuzzyEn) is further improved on the basis of SampEn, which is also widely used. Zheng *et al.* combined it with Intrinsic scale component (LCD) and partially ensemble empirical mode decomposition (PEEMD) for fault diagnosis, respectively [24], [25]. Similar to SampEn, researchers have also made many improvements to FuzzyEn, such as multi-scale fuzzy entropy (MFE) [26] and improved multi-scale fuzzy entropy (IMFE) [27]. These methods have also been applied to the bearing diagnosis field [28], [29]. Li *et al.* [30] proposed hierarchical fuzzy entropy (HFE) based on hierarchical entropy and applied it to rolling fault diagnosis.

Christoph and Bernd [31] proposed permutation entropy(PE), which is very sensitive to the mutations of random signals. PE and its improved methods have high computational efficiency and have been widely used in rolling fault diagnosis [32]–[38]. Furthermore, deep learning and the combination of machine learning model method and intelligent algorithm agent model is very promising and widely used [39]–[43]. Adaptive decomposition methods are also widely used in bearing fault diagnosis [44], [45].

Generally, fault diagnosis is to study how to identify different types of faults. In practical application, we not only need to identify different types of faults, but also monitor the state of mechanical operation and accurately judge the degradation of bearings. In this regard, many researchers have also carried out studies. For example, Yu [46] used the Hidden Markov Model to describe the bearing degradation process and established health indicators. Zhu *et al.* [47] introduced rough support vector data description (RSVDD) into rolling bearing performance degradation assessment. Pan *et al.* [48] combined wavelet packet decomposition with fuzzy clustering to evaluate bearing performance degradation. Moreover, it is a promising way to combine surrogate models involving machine learning model methods and intelligent algorithms. Song *et al.* [49] proposed a proxy model based on distributed cooperative wavelet neural network regression, and expounded the probabilistic framework of turbine disk LCF life evaluation.

As mentioned above, PE is very sensitive to random signal mutations, and its calculation is very fast. PE is widely used in the field of fault diagnosis. However, PE only considers the order of data and does not consider the amplitude relation between elements, so some effective information will be lost. In order to solve this problem, Azami and Escudero [50] put forward an improved method called amplitude-aware permutation entropy (AAPE). Chen *et al.* [51] combined intrinsic time-scale decomposition (ITD) with improved multi-scale AAPE (IMAAPE) for bearing fault diagnosis. This article will further to improve, propose a new method called RCMAAPE, and apply it to the analysis of simulation signals and bearing vibration signals. This paper proposes a new method for bearing performance degradation assessment and fault diagnosis based on RCMAAPE. Firstly, RCMAAPE is used to extract features to judge the correlation between vibration signal health data and test data, and Chebyshev's inequality is used to establish a health threshold to evaluate the performance degradation state of the bearing. Secondly, RCMAAPE and PSO-LIBSVM are combined for bearing fault diagnosis. Both experiments prove that the method proposed could extract fault characteristics effectively. The experimental results show that, compared with the existing methods, the proposed method can effectively extract the fault features, and has good effects in bearing performance degradation assessment and fault diagnosis.

The rest of this paper is organized as follows. Section II mainly presents the fundamentals of RCMAAPE method and its simulation signal test results. The experimental data are used to verify the effectiveness of the proposed method in rolling bearing performance degradation assessment in section III. In section IV, the experimental data are used to verify the effectiveness of the proposed method in rolling bearing fault diagnosis. Section V draws the conclusion of this paper.

II. METHODOLOGIES

A. PERMUTATION ENTROPY

Let the one-dimensional time series V with length N be shown in equation (1) [31].

$$V = \{v_1, v_2, \ldots, v_N\} \tag{1}$$

According to the embedded dimension m, the time series is reconstructed in space and the reconstructed vector is obtained as shown in equation (2).

$$V_i = (v_i, v_{i+t}, \ldots, v_{i+(m-1)t}) \tag{2}$$

where: m and t are the reconstruction dimension and delay time respectively.

The elements in the i-th reconstructed matrix are arranged according to the principle of from small to large, and it can be obtained as follows:

$$V_{i+(j_1-1)t} \leq V_{i+(j_2-1)t} \leq V_{i+(j_m-1)t} \tag{3}$$

where: $j_1, j_2, \cdots, j_m$ is the ordinal number of each element in the sequence before the permutation. If there are equal values in the sequence j, the sorting continues according to their own order, for any V_i , the sequence T can be obtained.

$$T_l = \{j_1, j_2, \ldots j_m\} \tag{4}$$

where: T_l is the l-th arrangement order, $l = 1, 2, \cdots, k$, and there are $k = m!$ different permutations.

The occurrence times of each situation are counted to obtain its occurrence probability. The relation is shown in equation (5).

$$\sum_{g=1}^{k} p_g = 1 \tag{5}$$

where: p_g is the probability of the ranking order of g.

The obtained permutation entropy after normalization is shown in equation (6).

$$H = \frac{-\sum\limits_{g=1}^{k} p_g In(p_g)}{In(m!)} \tag{6}$$

where: H is the normalized permutation entropy, which represents the randomness of the time series. The value is larger, the time series is more random, otherwise, the time series is more regular.

B. AMPLITUDE-AWARE PERMUTATION ENTROPY

Traditional permutation entropy has two main problems in describing complex time series. Firstly, PE only considers the sequence of time series amplitude from small to large, but it does not take into account the amplitude information of corresponding time series elements. Secondly, it does not clearly explain the influence of elements with equal amplitude in time series on the final PE value [50].

Aiming at the above problems of PE, Azami *et al.* proposed an improved method called amplitude-aware permutation entropy (AAPE), to make up for the shortcomings of PE. AAPE considers the mean value of signal amplitude and the deviation between the amplitudes, and introduces the relatively normalized probability to replace the counting rule of PE.

Assuming that the initial p_g value is zero, for time series V, i increases from 1 to N-m+1, and for different g, p_g is also updated simultaneously.

$$p_g^{update} = p_g + \Big(\frac{A}{m}\sum_{n=1}^{m} |v_{i+(n-1)t}| + \frac{1-A}{m-1}\sum_{n=2}^{m} |v_{i+(n-1)t} - v_{i+(n-2)t}|\Big) \tag{7}$$

where: A is the adjustment coefficient, which is used to adjust the weight of the signal amplitude mean value and the deviation between the amplitudes. Generally, $A = 0.5$ is taken; p_g^{update} is the updated p_g value. Then, in the whole time series, the probability of occurrence of g is shown in equation (8), as shown at the bottom of the page.

AAPE can be obtained as shown in equation (9):

$$AAPE(N, m, t) = -\sum_{g=1}^{k} p_g In(p_g) \tag{9}$$

C. REFINED COMPOSITE MULTI-SCALE AMPLITUDE-AWARE PERMUTATION ENTROPY

Multi-scale analysis is to cut the original sequence into short sequences at equal distance, and then calculate the mean value of these short sequences to form a new coarse-grained sequence. Although this method is simple and fast, the relationship between the elements of the new coarse-grained sequence is not considered when cutting the time series with equal distance. If the starting point is different, the final entropy will also fluctuate to a certain extent.

Therefore, the multi-scale analysis method is further refined. For the one-dimensional time series V, when the scale factor is s, the original sequence is divided into s series according to the starting point from 1 to s. The multi-scale analysis of each sequence with the scale factor of s is used to obtain s coarse-grained sequences. Then p_g^b of coarse-grained sequence is calculated, respectively. Then $\overline{p_g}$ is calculated according to equation (11), and RCMAAPE is calculated with equation (12). The process of refined composite multi-scale analysis is shown in Figure 1.

For one-dimensional time series V, its i-th coarse-grained sequence $w_l^s = \{w_{l,1}^{(s)}, w_{l,2}^{(s)}, \cdots w_{l,l}^{(s)}\}$.

$$w_{k,j}^{(s)} = \frac{1}{s}\sum_{i=k+s(j-1)}^{k+js-1} v_i 1 \leq \mathrm{j} \leq \frac{N}{s}, 1 \leq \mathrm{k} \leq s \tag{10}$$

$$\overline{p_g} = \frac{1}{s}\sum_{b=1}^{s} p_g^b \tag{11}$$

$$RCMAAPE(N, m, t) = -\sum_{g=1}^{k} \overline{p_g} In(\overline{p_g}) \tag{12}$$

The obtained amplitude-aware permutation entropy after normalization is shown in equation (13):

$$RCMAAPE = \frac{-\sum\limits_{g=1}^{k} \overline{p_g} In(\overline{p_g})}{In(m!)} \tag{13}$$

D. PARAMETERS SELECTION

The main parameters considered in calculating RCMAAPE are embedding dimension (m), time delay (t), sample length

$$p_g = \frac{p_g^{update}}{\sum\limits_{i=1}^{N-m+1} \Big(\frac{A}{m}\sum\limits_{n=1}^{m} |v_{i+(n-1)t}| + \frac{1-A}{m-1}\sum\limits_{n=2}^{m} |v_{i+(n-1)t} - v_{i+(n-2)t}|\Big)} \tag{8}$$

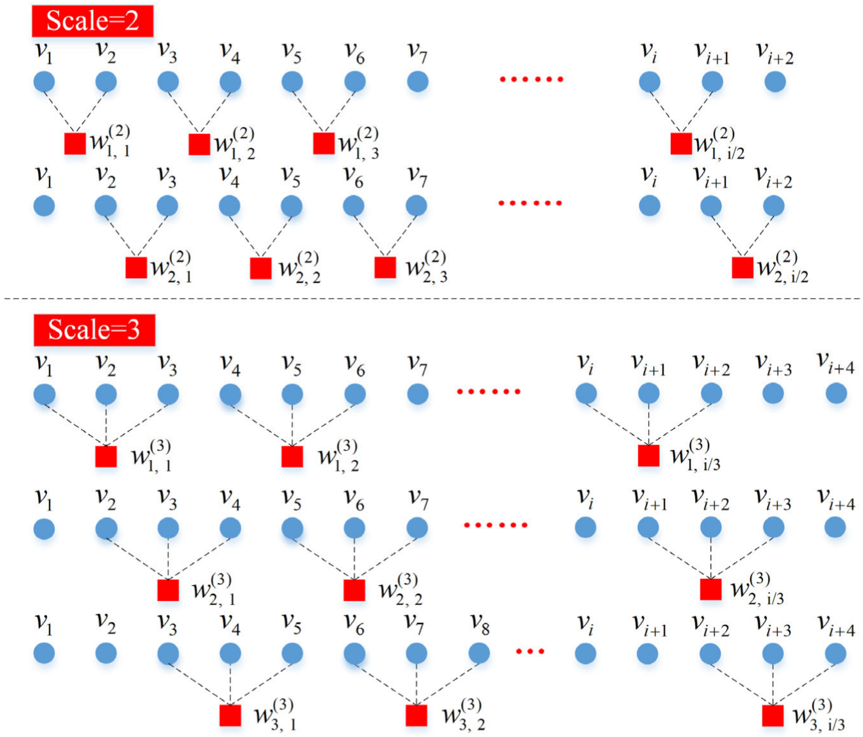

FIGURE 1. Coarse-grained analysis process of refined composite multi-scale amplitude-aware permutation entropy.

(N), and adjustment factor (A), respectively. According to reference [48], t is usually the minimum positive integer 1, A is usually 0.5. The embedded dimension m is the length of the sequence to be compared, if m is too small, the number of permutations is very limited, which will lead to information loss. The larger the m is, the more detailed the reconstruction of the dynamic process will be, but a too large m value will lead to a large amount of calculation of entropy value. According to reference [49], m is fixed as 4. RCMAAPE is less dependent on sample length, so N in this paper is chosen as 2048.

E. SIMULATION SIGNAL ANALYSIS

To illustrate the effectiveness of the proposed method, white gaussian noise (WGN) and 1/ f noise containing 2048 data points are adopted to perform the comparison of RCMAAPE and other three entropies (MAAPE, RCMSE and RCMFE). Both RCMAAPE and MAAPE use the parameters selected in previous section. For RCMSE and RCMFE, we set m $= 2$, $r = 0.15$ and scale $= 20$. WGN is higher than 1/ f noise in terms of the degree of irregularities, but 1/ f noise is more complex than WGN in terms of structure. Entropy value of WGN is theoretically greater than those of 1/f noise. The four entropies of 50 sets of WGN and 1/f noise are respectively calculated and the mean values of 50 sets are presented in Figure 2.

As shown in Figure 2, the RCMAAPE and MAAPE values of WGN are greater than those of 1/f noise, and the RCMSE and RCMFE values of WGN are greater than those of 1/f noise at the first two scales, which is consistent with theoretical analysis. It can be also seen from Figure 2 that the measurement scales of the four entropies are quite different. It is not appropriate to directly use the standard deviation(SD) to compare the degree of data dispersion. In order to eliminate the influence of measurement scale, the coefficient of variation(CV) is used to compare the degree of dispersion of the four sets of data in this paper(CV is defined as the SD divided by the mean value). To highlight the advantages of RCMAAPE, CV of four entropies for WGN and 1/ f noise are respectively calculated and the results are shown in Figure 3. As can be seen in in Figure 3, whether it is WGN or 1/f noise, the CV of RCMAAPE is the smallest in four entropies. Therefore, RCMAAPE can achieve more accurate results in the calculation of entropy than other three entropies, and is the most suitable for detecting dynamic changes of the complex signal.

III. CASE 1: ROLLING BEARING PERFORMANCE DEGRADATION ASSESSMENT

In this section, the proposed method is applied to the performance degradation assessment of bearing. The essence of rolling bearing performance degradation assessment is to judge the correlation between normal signals and tested signals of the rolling element bearings, and then to evaluate the performance degradation state of the bearing. In this paper, RCMAAPE is used to extract the features and draw

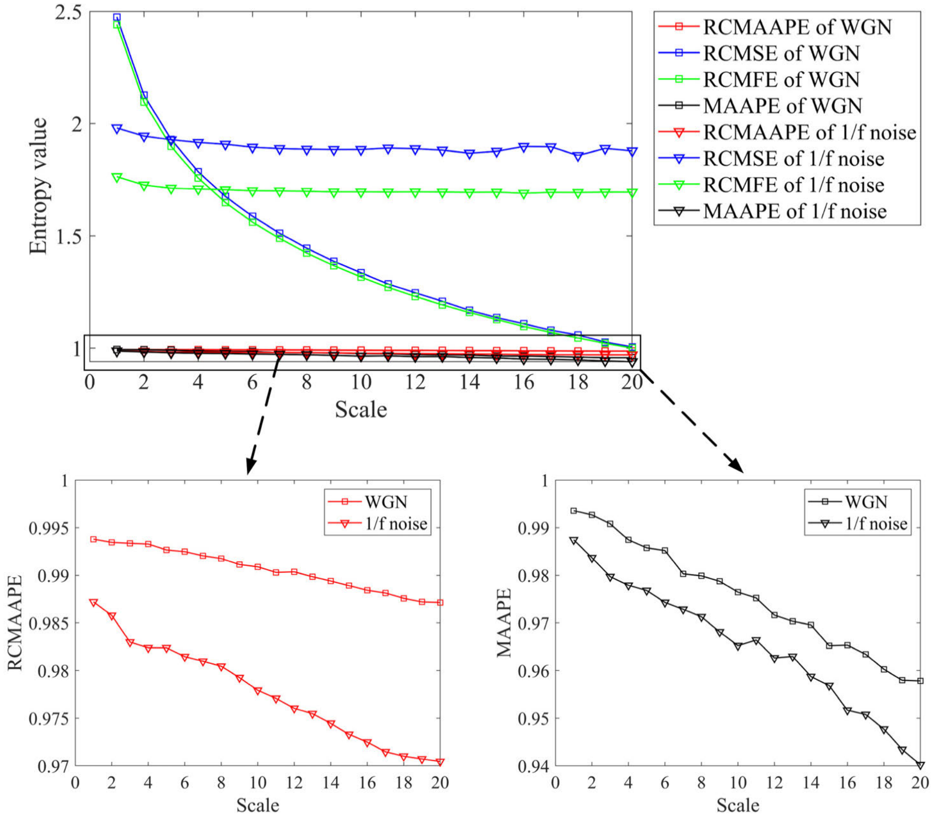

FIGURE 2. Simulation signal analysis of four entropies (MAAPE-multi-scale amplitude-aware permutation entropy, RCMSE- refined composite multi-scale sample entropy, RCMFE- refined composite multi-scale fuzzy entropy).

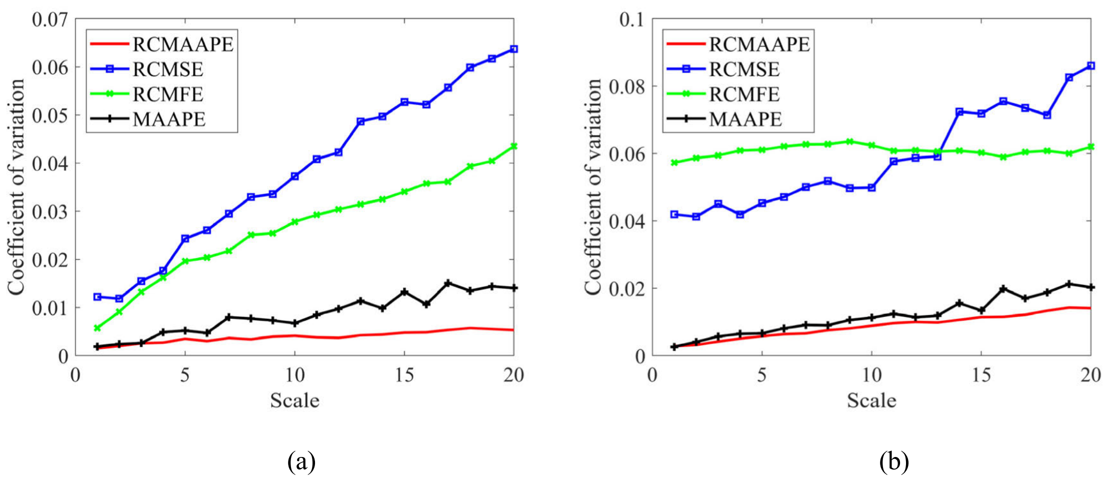

FIGURE 3. (a) CV values of four entropies for WGN, (b) CV values of four entropies for 1/f noise.

the entropy curve of the bearing life to evaluate the bearing performance degradation.

A. DATA SOURCES

The experimental data selected in this paper are bearing data set provided by Center on Intelligent Maintenance Systems (IMS), University of Cincinnati [52]. The physical and structural drawings of the bearing test platform are shown in Figure 4. The AC motor rotated at a constant speed of 2000r/min and was coupled by friction belts to drive the rotating shaft installed with four Rexnord ZA-2115 double row bearings, and the bearings have 16 rollers in each row, a pitch diameter of 2.815 in, roller diameter of 0.331 in, and

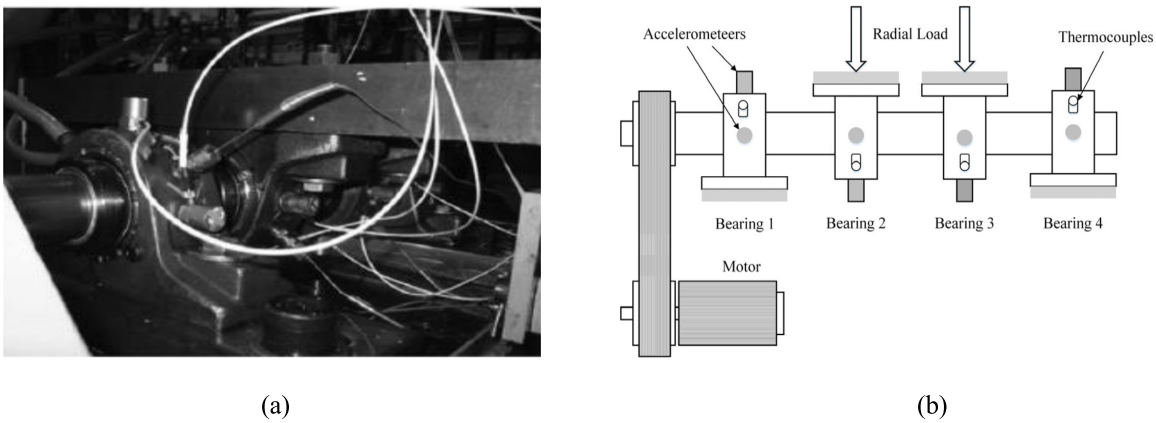

FIGURE 4. Bearing test platform: (a) physical diagram, (b) structural sketch.

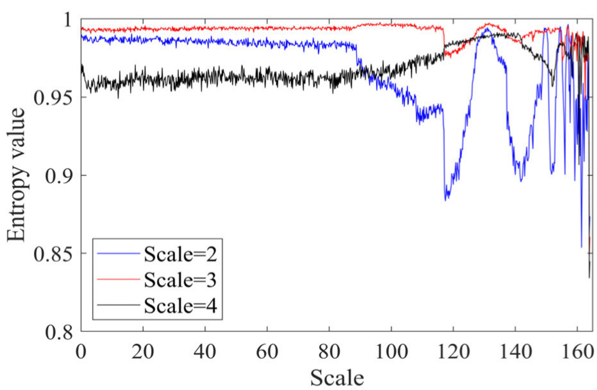

FIGURE 5. Effects of scale factor on RCMAAPE.

a tapered contact angle of 15.17°. The detailed information of the bearings is as follows. A spring mechanism was used to apply 6,000 pounds of radial load to the shaft and bearing. The magnetic plug was equipped in the oil return pipe of the lubrication system to judge the degradation degree of the bearing. The fault measurement standard was that the magnetic plug collected the metal wear particles in the lubricating oil, when the metal particles reached a certain amount, it indicated that the bearing has suffered serious fault, thus ending the experiment. A PCB 352B33 High Sensitivity Quarts ICP accelerometer was installed on each bearing housing to collect the vibration signals.

In this paper, the data set was collected in the second experiment. The sampling frequency is 20 kHz, and each file contains 20480 data points, the interval of file collection is 10 minutes. Data collection started at 10:32:39, February 12, 2004 and ended at 06:22:39, February 19, 2004. A total of 984 files were collected and the test lasted 163.83 hours. After the test, the outer ring of bearing 1 was severely worn.

B. EFFECT OF SCALE FACTOR SELECTION

The parameters of RCMAAPE have been determined in section II. Bearing entropy life curve is single-scale, so only the value of the scale factor needs to be determined. This article compares and analyzes the experimental data used to determine the scale factor. Figure 5 is the effects of scale factor on RCMAAPE. As shown in Figure 5, if scale factor is too large, some information will be lost. Entropy value is insensitive to the early failure of the bearing and affects the evaluation result of bearing performance degradation. The scale factor should be as small as possible, so the scale factor is fixed as 2. At the same time, to make full use of experimental data and avoid random errors. The data in each file is divided into 10 sections, their entropy value is calculated, respectively. The final result is the average of ten entropy values [54].

C. HEALTH THRESHOLD

After calculating the RCMAAPE, the performance degradation process of the rolling bearing can be monitored. In practical application, establishing an appropriate and easy to implement health threshold is the key to identifying the running state of bearings. In this paper, according to literature [53], [54], Chebyshev inequality theory is an effective method to establish health threshold, which is defined in equation (14).

$$P\{|X-\mu_h| \geq \varepsilon_h\} \leq \frac{\sigma_h^2}{\varepsilon_h^2} \text{ or } P\{|X-\mu_h| < \varepsilon_h\} \geq 1-\frac{\sigma_h^2}{\varepsilon_h^2} \tag{14}$$

In the formula, X represents the sequence of RCMAAPE values under the failure-free operation of the bearing, μ_h is the mean value of X, and σ_h is the standard deviation of X.

According to Chebyshev's inequality theory [53], $\varepsilon_h = 5\sigma_h$. This means that the probability that the entropy value of the bearing is greater than $\mu_h \pm 5\sigma_h$ is not more than 4%, and the probability that the bearing fault can be accurately identified is 96%. To minimize the possibility of false positives, the first 60 hours of bearing operation are selected as the health state to establish the health threshold. The data in the whole life of the bearing are processed as test signals.

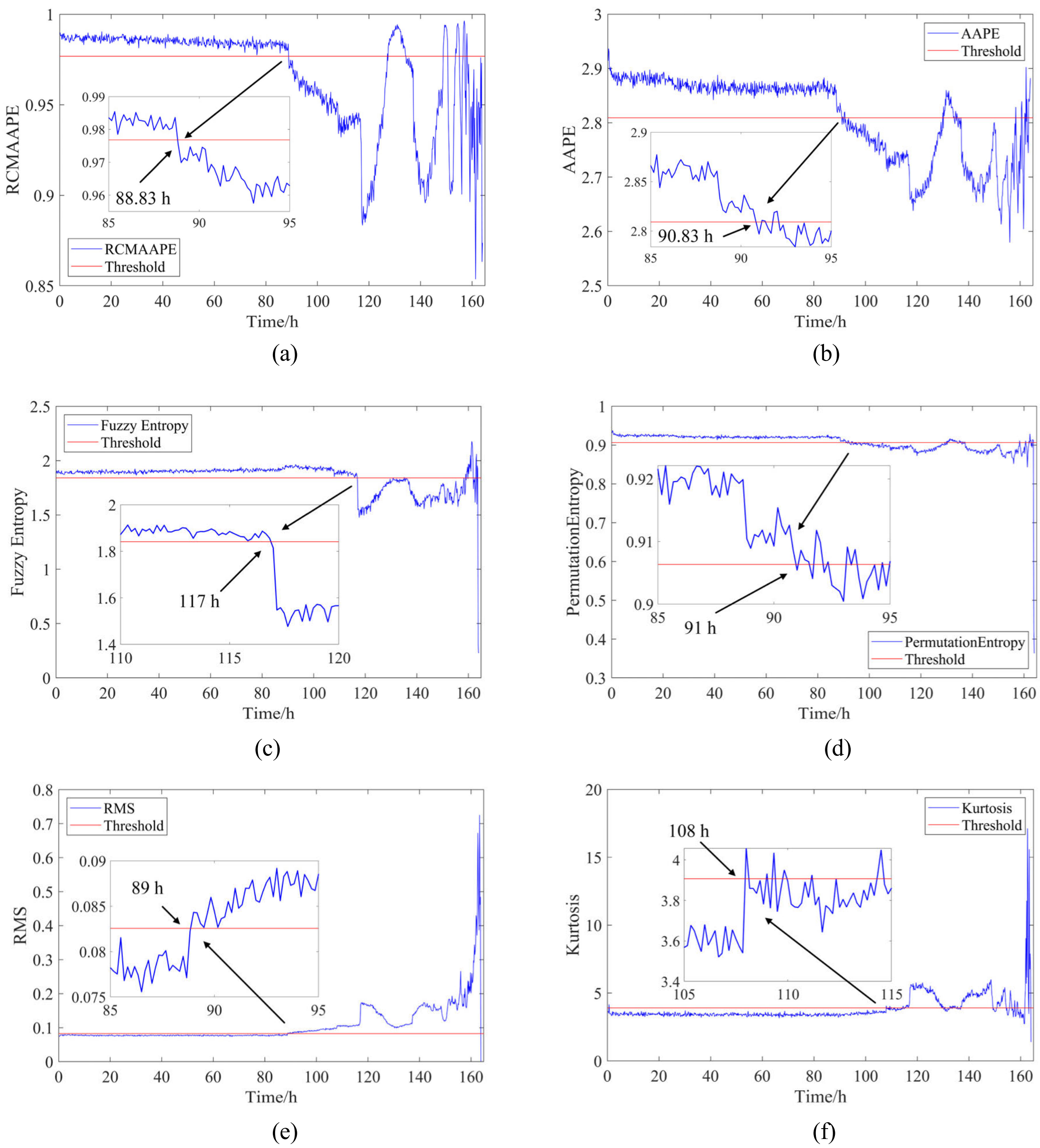

FIGURE 6. **Six indicators with health threshold over whole lifetime of bearing 1.**

D. RESULTS AND ANALSYS

RCMAAPE with health threshold over whole lifetime of bearing 1 is shown in Figure 6(a). RCMAAPE can extract the fault characteristics of bearing to evaluate the performance state of the bearing. In the fault-free stage, the entropy value keeps a stable state. When an early failure occurs, the entropy curve has a mutation, which means that the vibration signals of healthy bearing are more complex than those of faulty bearing. This can be explained by the fact that the vibration signal under normal conditions has low self-similarity due to its irregularity and randomness, and the self-similarity will increase with the occurrence of faults. It can be seen from Figure 6(a) that RCMAAPE values indicate the occurrence of degradation at 88.83 h, the change range of the curve is larger when an early failure occurs, and the health threshold can provide a warning in time. For the reason, RCMAAPE is sensitive to bearing early failure and is very suitable for bearing performance degradation.

In order to further show the advantage of RCMAAPE over the existing method, the existing methods are also adopted for comparing their performance. Three entropies with health threshold over whole lifetime of bearing 1 are shown from Figure 6(b) to (d). As shown in Figure 6(b), AAPE is also very sensitive to early failures, but its entropy values fluctuate

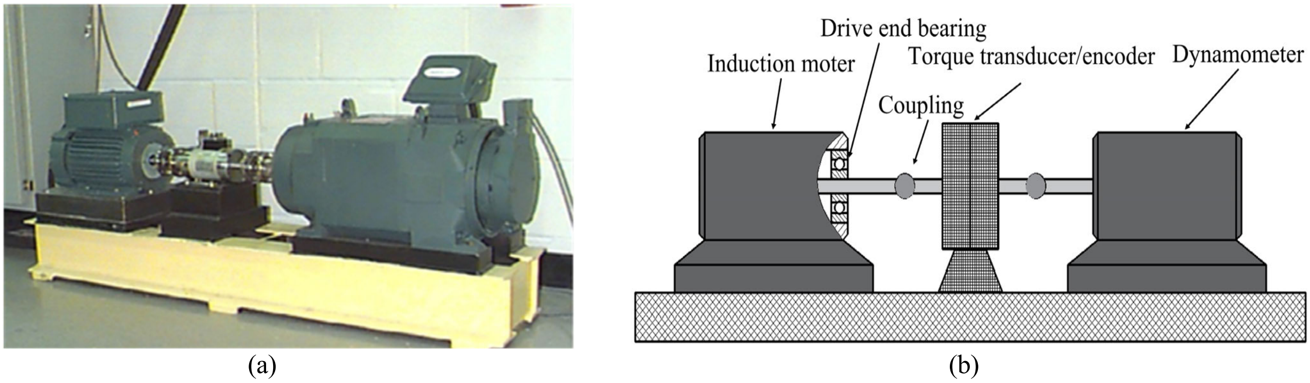

FIGURE 7. Experimental apparatus for bearing: (a) physical diagram, (b) structural sketch.

greatly, so the health threshold is lower. The intersection of the health threshold and the curve is at 90.83 h, which means that RCMAAPE can detect the early degradation 2 h in advance. Refined composition is a process of averaging multiple sets of data, RCMAAPE has better statistical stability. From Figure 6(c), When an early failure occurs, FuzzyEn values do not change significantly, which means that FuzzyEn is not sensitive to early failures. It takes a lot of time to calculate FuzzyEn, especially for a large amount of data. It can be seen from Figure 6(d) that PE values indicate the occurrence of degradation at 91 h, the decrease of PE values is less obvious than that of RCMAAPE. PE values have a certain change when the early fault occurs, but the change range is small.

For comparison purpose, another two degradation assessment indicators, root mean square (RMS) and the kurtosis value, are also presented in Figure 6(e) and Figure 6(f), respectively. As shown in Figure 6(e), the health threshold and curve intersect at 89 hours, 10 minutes later than the RCMAAPE value. Moreover, RMS values have a certain change when the early fault occurs, but the change range is also small. It can be seen from Figure 6(f) that kurtosis values indicate the occurrence of degradation at 108 h, which is more than 19 h later than that indicated RCMAAPE. Kurtosis is very sensitive to shock failures, kurtosis values fluctuate greatly. It is obvious that kurtosis is not sensitive to early failure. Table 1 is the results of performance degradation evaluation of rolling bearing.

According to the above analysis, RCMAAPE can extract the fault characteristics and assess the bearing performance degradation process over their whole life time clearly and effectively. Compared with the existing methods, the proposed method can detect the initial degradation of the bearing in advance, which is of great importance to the prediction in condition monitoring. Therefore, RCMAAPE is the most suitable for performance degradation evaluation of rolling bearings.

IV. CASE 2: ROLLING BEARING FAULT DIAGNOSIS

In the previous section, the proposed method is applied to the performance degradation evaluation of rolling bearing.

TABLE 1. Results of performance degradation evaluation of rolling bearing.

Methods	The warning time /h	Amplitude of curve slope change
RCMAAPE	88.83	Obvious
AAPE	90.83	Obvious
FuzzyEn	117	Very small
PE	91	Small
RMS	89	Small
Kurtosis	108	Very small

The results show that RCMAAPE can effectively extract fault features and find faults in time. In this section, the proposed method is combined with the PSO-LibSVM classifier and further applied to the rolling bearing fault diagnosis to judge the location and severity of the fault.

A. DATA SOURCES

This paper uses the rolling bearing data provided by the bearing data center of Case Western Reserve University (CWRU) to experimentally verify the effectiveness of the proposed fault diagnosis method [55]. The experimental platform and its structure diagram are shown in Figure 7. The experimental apparatus includes a 2 horsepower motor, a torque encoder and a dynamometer. The bearings used in the test were the 6205-2RS JEM SKF deep groove balling bearing. Three types of bearing faults were generated by electric discharge machining: bearing inner ring fault (IR), outer ring fault (OR) and ball element fault (B), and 4 different degrees of damage are done (some category data is not available). The experimental data is collected by the acceleration sensor, and the sampling frequency is 12 kHZ.

The experimental data selected in this article is the motor speed of 1730 r/min. The test data include normal signal (Nor), inner ring fault (IR), ball elements fault (B) and outer ring fault (OR). Each fault type includes three kinds of faults with different severities (different fault sizes), respectively. The specific category and labels are shown in Table 2. In this paper, nine fault signals and one normal signal are selected. The time-domain waveforms of typical samples of each category are shown in Figure 8. The data length of the sample is fixed 2048, and 50 data samples are taken for each type, and

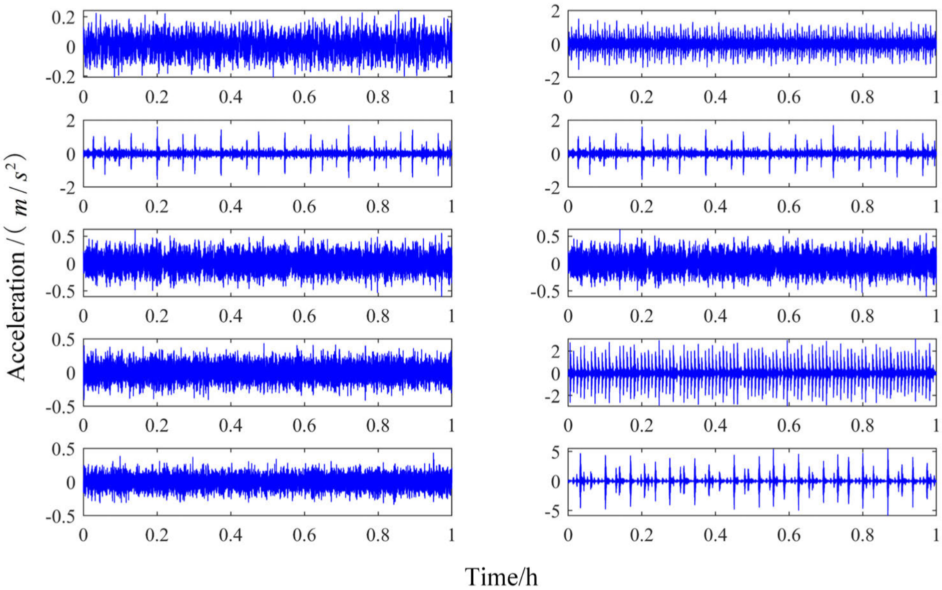

FIGURE 8. Time-domain waveform of selected signals in the experimental data.

a total of 500 samples form the sample set. In order to realize the automatic diagnosis of rolling bearing faults, 10 samples of each type are randomly selected as training samples, and the remaining 40 ones are used as test samples.

B. MULTI-CLUSTER FEATURE SELECTION

In this paper, multi-cluster feature selection(MCFS) is used to select the best feature subset from all fault features. MCFS was proposed by He *et al.* in 2010 [56]. MCFS considers the interrelationships between the selected features to identify the multi-type structure of the original data. It has been applied in many fields and has been proven to be an effective feature selection method [57].

Given a raw data set $V = [v_1, v_2, \cdots, v_N], v_i \in R^M$, and each data point has M features. The purpose of the MCFS method is to find the most informative feature among M features. In other words, the points $V = \left[v_1', v_2', \cdots, v_N'\right]$ represented in the d-dimensional space R^d can well retain the geometric structure of the data represented in the original M-dimensional space. The main calculation steps of the MCFS algorithm are as follows.

Construct a graph with N vertices, each vertex corresponds to a data point. For each data point v_i, its p nearest neighbors are found and an edge is placed between v_i and its neighbors. These edges are defined by weights, and the weight matrix W of this graph can be determined by a variety of weighting methods. The commonly used one is the 0-1 weighting method: $W_{ij} = 1$, if and only if there is an edge between node i and j. This is the simplest weighting method, and it is easy to calculate.

The elements of the diagonal matrix D are the column sums of W, and it can be obtained as follows:

$$D_{ii} = \sum_j W_{ij} \tag{15}$$

The graph Lapalcian can be obtained as follows:

$$L = W - D \tag{16}$$

In order to find the flat embedding of the data points of the unfolded data manifold, the generalized feature problem needs to be solved.

$$Ly = \lambda Dy \tag{17}$$

The flat embedding of each data point is each row $Y = [y_1, y_2, \cdots, y_K]$, where y_k is the eigenvector of the minimum eigenvalue in the generalized eigenvalue problem above, the K is the intrinsic dimensionality of the data, and each y_k is the reflection of the data distribution along this dimension.

Given y_k, a column of Y, a relevant subset of features can be found by minimizing the fitting error as follows:

$$\min_{a_k} y_k - X^T a_k^2 \quad s.t|a_k| < \gamma \tag{18}$$

where a_k is a M-dimensional vector, and $|a_k|$ is the $L1$-norm of a_k, which is defined as

$$|a_k| = \sum_{j=1}^{M} |a_{k,j}| \tag{19}$$

In the process of approximating y_k, a_k essentially contains the combination coefficients of different features. Corresponding to the non-zero coefficients in a_k, a subset containing the most relevant features can be selected for y_k. Equation (18) is essentially a regression problem that can be solved by the Least Angle Regression(LARs) algorithm.

For each feature j, the MCFS score can be calculated in equation(20) [56].

$$MCFS(j) = \max_k |a_{k,j}| \tag{20}$$

TABLE 2. Composition of experimental data for fault diagnosis.

Fault mode	Fault size(mils)	Fault degree	Label	Training Data	Test Data
Normal	*	*	1	10	40
IR7	7	Slight	2	10	40
IR14	14	Middle	3	10	40
IR21	21	Severe	4	10	40
B07	7	Slight	5	10	40
B14	14	Middle	6	10	40
B21	21	Severe	7	10	40
OR7	7	Slight	8	10	40
OR14	14	Middle	9	10	40
OR21	21	Severe	10	10	40

where $a_{k,j}$ is the j-th element of the vector a_k. Finally, all features are sorted in descending order according to the their own MCFS scores, and the d features are selected from all feature candidates.

C. CLASSIFICATION STEPS

On the basis of the above analysis, this paper proposes a method based on the combination of RCMAAPE and PSO-LibSVM for rolling bearing fault categories and severity. The specific steps are described as follows:

Step 1: the experimental data of different categories and severity are selected from the database. Each experimental data is divided into experimental samples according to the selected data length to form a sample set.

Step 2: RCMAAPE values of the first 20 scales of each sample are calculated. Part of the samples are randomly selected as the training set, and the rest as the test set.

Step 3: MCFS is used to select a subset of features from the main 20 features according to its MCFS score, and the appropriate number of features are selected to train the PSO-LibSVM.

Step 4: the test set is input into the PSO-LibSVM classifier, and the fault category and severity of rolling bearing can be obtained according to the output results.

D. RESULTS AND ANALYSIS

The widely used classifiers are artificial neural networks (ANN) and support vector machine(SVM). Compared with ANN, SVM is more suitable for processing learning problems with small-quantity samples. Therefore, LibSVM is selected to complete the bearing fault diagnosis in this paper. LibSVM is used to achieve automated bearing fault diagnosis. LibSVM is a support vector machine (SVM) library proposed by Professor Lin. It has the advantages of small program, flexible, few input parameters, open source, easy to expand, etc [58]. It is currently the most widely used SVM library. The kernel function parameter g and penalty factor c are the most important parameters of LibSVM, which have a great influence on the accuracy of classification. Particle swarm optimization (PSO) is used to optimize it to choose the appropriate parameters. The particle swarm size and number of iterations are 30 and 100, respectively.

To begin with, single-scale entropy is used to diagnose bearing faults. When the scale factor $s = 1$, RCMAAPE is AAPE. Compare it with PE, SampEn and FuzzyEn. In order to realize the automatic diagnosis of rolling bearing faults, 10 samples of each type are randomly selected as training samples, and the remaining 40 ones are used as test samples. The classification accuracy of AAPE, PE, SampEn and FuzzyEn are 68.5%, 73.5%, 82% and 75%, respectively. The accuracy rates of four single-scale entropies are generally low, because the entropy value of some types of faults is similar, and there is no discrimination. Aiming at the problem of bearing fault classification considering different fault locations and severities at the same time, the accuracy of single-scale entropy values is generally low, which obviously cannot meet the requirements. Single-scale entropy is difficult to fully extract the fault characteristic information of the signal.

TABLE 3. Classification rates of single-scale entropy.

Methods	Accuracy /%	Misclassified Sample Number
AAPE	68.5	126
PE	73.5	106
SampEn	82	72
FuzzyEn	75	100

Therefore, multi-scale analysis is performed to calculate the entropy values of different scales of the vibration signal, which are used as a feature vector set and input into the classifier for bearing fault diagnosis. The max scale factor is selected as 20. The results of rolling bearing data analyzed by the RCMAAPE are given in Figure 9(a). Obviously, the entropy of a single scale cannot effectively identify all bearing states. Correspondingly, an initial training data set with a dimension of 100×20 and an initial test data set with a dimension of 400×20 can be obtained. The training data set with dimension 100×20 is used to train PSO-LibSVM. The test data set with dimension 400×20 is used to test the well-trained multi-classifier and the outputs are given in Figure 11(a), from which it can be found that all test samples are correctly classified and the fault identification rate is 100%. It can be observed that the proposed method is effective in detecting different working conditions of rolling bearing.

In order to highlight the advantages of the RCMAAPE, the existing methods (MAAPE, RCMSE, and RCMFE) are also adopted for comparing their diagnosis performance. Similar to the above method, the construction of the training and test data sets is the same as the fault diagnosis method proposed in this paper. The results of rolling bearing data analyzed by the RCMAAPE are given

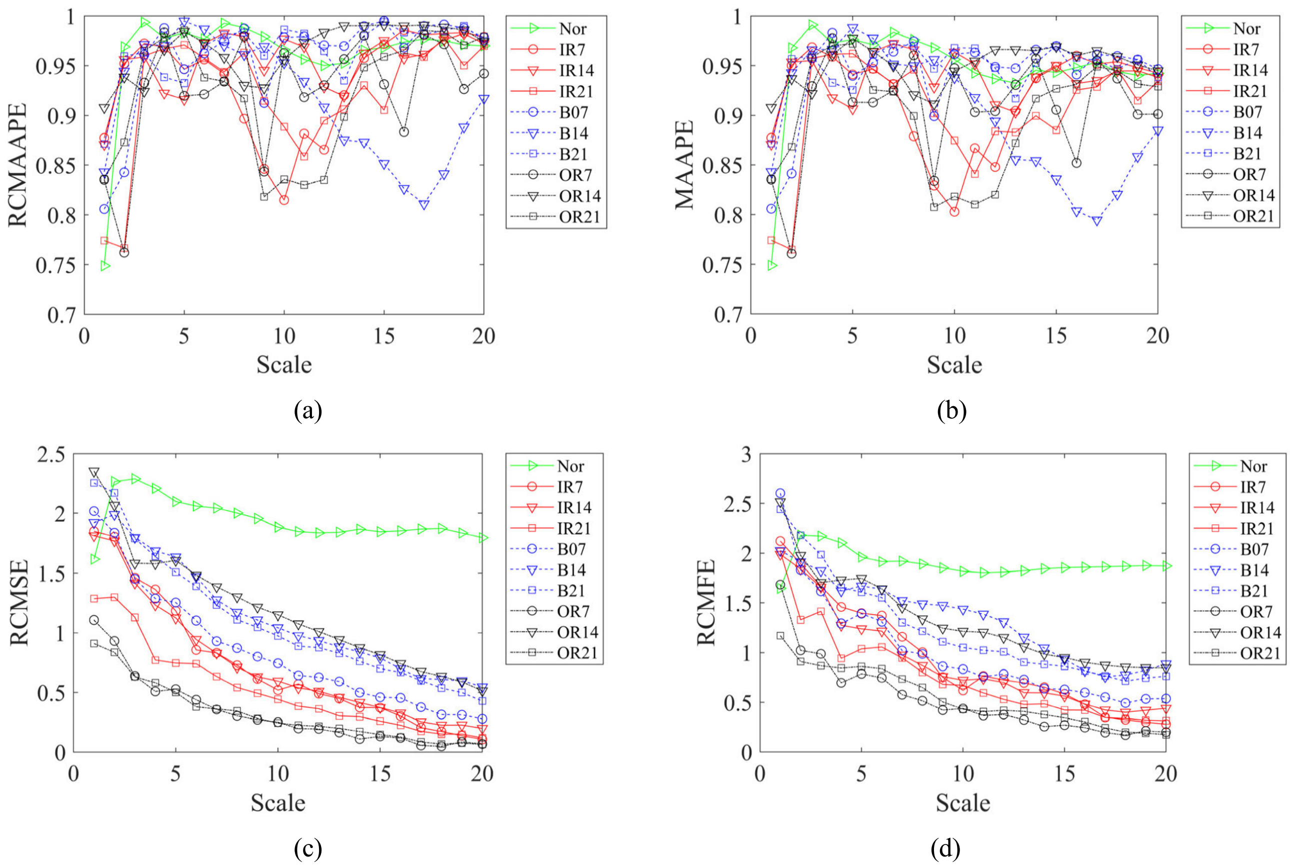

FIGURE 9. **The results of rolling bearing data analyzed by the four entropies.**

in Figure 9(b)-(d). The PSO-LibSVM outputs of three methods are presented in Figure11(b)-(d), and the classification results are also shown in Table 4. The identification accuracy of MAAPE is 96.75%, thirteen inner ring fault samples (IR14) are misclassified to rolling element fault (B21). From Figure 9(a) and (b), the entropy curve distribution of MAAPE is similar to that of RCMAAPE. However, the recognition accuracy of MAAPE is slightly lower than that of RCMAAPE. In order to explore the reason for the higher fault recognition rate of RCMAAPE, this article also uses CV to compare the differences between the two algorithms. It can be seen from Figure 10(a) that in most cases, the difference between the CV of the RCMAAPE value and the CV of the MAAPE value is negative, which shows that the degree of dispersion of RCMAAPE is significantly less than that of MAAPE, and RCMAAPE has better stability and reliability. Therefore, refined composition is a process of averaging multiple sets of data, RCMAAPE has a smaller error in extracting the fault characteristics of the bearing, and the result is more stable.

Similarly, as shown in Figure 9(c) and (d), the entropy curve distribution of RCMSE is also similar to that of RCMFE. It can be seen from Figure 11(c) and (d) that the identification accuracy of RCMSE and RCMFE are 94.25% and 99.75%. The identification accuracy of RCMFE is significantly higher than that of RCMSE. RCMFE uses fuzzy function to measure the complexity of data, better statistical stability can be obtained. Because the fuzzy function has soft and continuous boundaries. It can be seen from Figure 10(b) that in most cases, the difference between the CV of the RCMFE value and the CV of the RCMSE value is also negative, which is consistent with theoretical analysis.

The above analysis shows that the stability of the entropy value has a significant impact on the classification problem including fault category and severity of rolling bearing. The higher the stability, the higher the classification accuracy. The proposed method can get the highest identification accuracy when handling the classification problem including fault category and severity of rolling bearing.

In the above case, we input the entropy values of 20 scales as feature vectors into the classifier, and obtain an accurate fault recognition rate. However, the high feature dimension obviously has information redundancy, which increases the calculation time. Therefore, it is necessary to reduce redundant features to form low-dimensional feature vectors. In this paper, MCFS is used to select the best feature subset from all fault features.

For comparison purposes, we choose RCMAAPE and RCMFE with better classification results to use MCFS for feature selection. After the feature selection, the fault

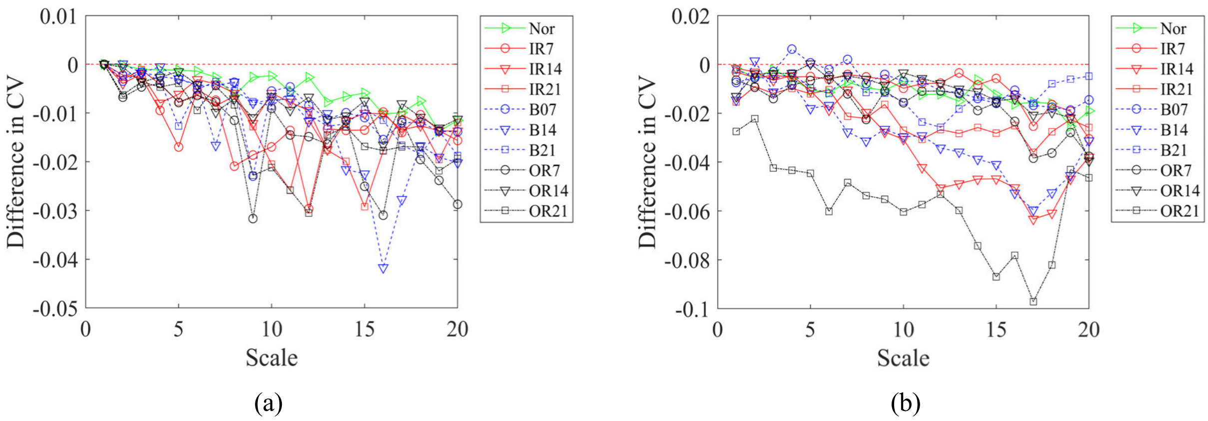

FIGURE 10. (a) CV differences between RCMAAPE and MAAPE, (a) CV differences between RCMFE and RCMSE.

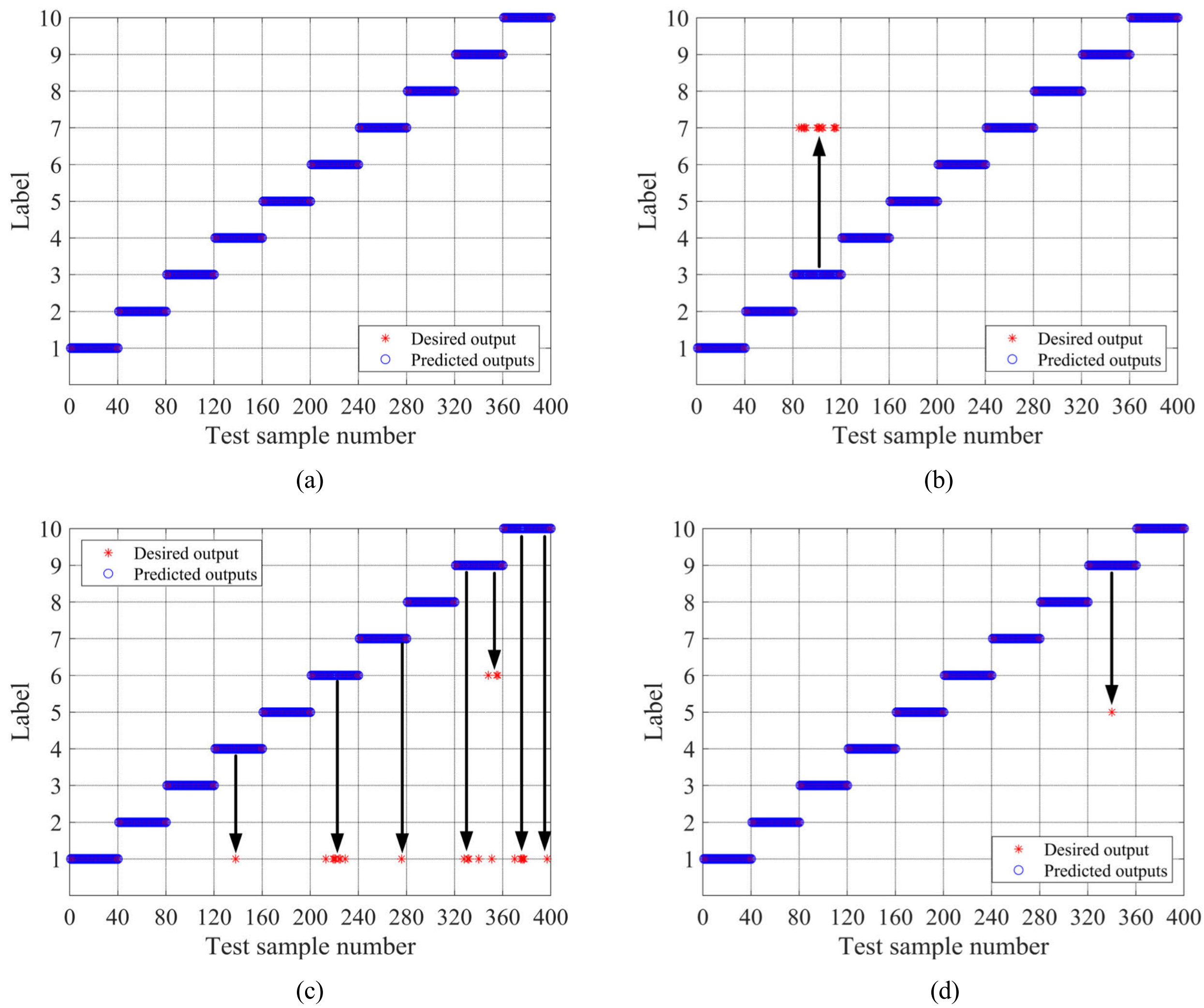

FIGURE 11. The PSO-LibSVM outputs of four entropies: (a) RCMAAPE, (b) MAAPE, (c) RCMSE, (d) RCMFE.

features of different dimensions are sequentially input into the multi-classifier based on PSO-LibSVM. Table 5 and Table 6 show the recognition rates of the above-mentioned different methods for different numbers of new fault feature elements. The comparison of recognition rates is shown in Figure 12. From Figure 12, Table 5 and Table 6, we can find

TABLE 4. The diagnosis results of different methods.

Methods	Accuracy /%	Number of misclassified sample	Computing Time of 20 scale entropy /s
RCMAAPE	100	0	1.076
MAAPE	96.75	13	0.293
RCMSE	94.25	23	2.430
RCMFE	99.75	1	15.893

TABLE 5. Identification accuracy of RCMAAPE and MCFS for different number of features.

Method	Number of features	1	2	3	4	5	6	7	8	9	10
RCMAAPE + MCFS + PSO-LibSVM	Identification accuracy %	68.5	87.5	96	98	99.5	100	99.75	99.5	100	100
	Number of features	11	12	13	14	15	16	17	18	19	20
	Identification accuracy %	100	100	100	100	100	100	100	100	100	100

TABLE 6. Identification accuracy of RCMFE and MCFS for different number of features.

Method	Number of features	1	2	3	4	5	6	7	8	9	10
RCMFE + MCFS + PSO-LibSVM	Identification accuracy %	75.25	95.75	98.5	100	100	100	100	99.5	99.25	99
	Number of features	11	12	13	14	15	16	17	18	19	20
	Identification accuracy %	99.25	99.5	99.5	99	99.75	99.75	99.75	100	100	100

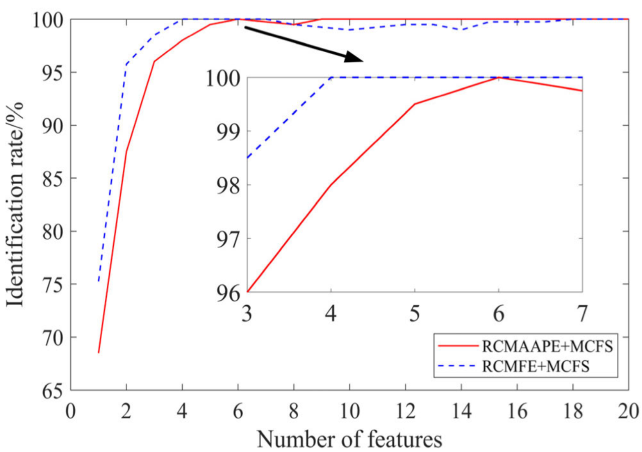

FIGURE 12. Identification accuracy comparison of RCMAAPE and RCMFE versus number of features.

that when the number of fault feature elements is set as 4, the identification accuracy of RCMFE reaches 100%. when the number of fault feature elements is set as 6, the identification accuracy of RCMAAPE reaches 100%. When the number of fault feature elements is larger than 9, the identification accuracy of RCMAAPE is 100%. Consequently, after the feature selection, both RCMAAPE and RCMFE only need a small number of feature vectors to effectively identify the different states of the bearing. This shows that the use of MCFS for feature selection is very necessary and effective.

In general, the overall recognition rates of four single-scale entropies are relatively low. Because the entropy values of different states are similar, the recognition rate is low. Single-scale entropy is difficult to fully extract the fault characteristic information of the signal. After the multi-scale analysis, the identification accuracy is significantly improved. It can be seen from Table 4 that RCMAAPE with the highest identification accuracy reaches 100%. However, the high feature dimension obviously has information redundancy. After feature selection, the recognition accuracy of RCMFE only needs 4 feature vectors to reach 100%, but from Table 4, it can be seen that the computing time of RCMAAPE and RCMFE with 20 scales is 1.076 s and 15.893 s, respectively. The computing efficiency of RCMAAPE is much higher than that of RCMFE. Considering the computing efficiency and recognition accuracy, RCMAAPE shows better performance than other methods.

TABLE 7. Comparison of identification rates of different rolling bearing fault diagnosis methods.

References	Number of bearing states	Recognition accuracy / %	
		Method of comparison	The proposed method
[18]	4	97.78	100
[51]	10	99.25	100
[57]	10	99.67	100
[59]	10	99.10	100

In order to further illustrate the performance of the proposed method in rolling bearing fault diagnosis, the rolling bearing fault diagnosis method proposed in this paper is directly compared with the existing research methods. The results are shown in Table 7. It can be seen from Table 7 that the method proposed in this paper can effectively roll bearing fault type and severity, and the fault identification accuracy reaches 100%. Moreover, the proposed method has higher computational efficiency. Therefore, this is a very effective and promising fault diagnosis method.

V. CONCLUSION

To solve the problems of unclear early fault characteristics and difficult extraction of rolling bearings. A new non-linear signal analysis method called RCMAAPE is introduced in this paper, which is used to extract bearing fault features. Two experimental data are used to verify the effectiveness of the proposed method. The conclusions of this paper are as follows:

(1) It is analyzed by using WGN and 1/ f noise, and compared with RCMSE, RCMFE and MAAPE, which shows that RCMAAPE had very good stability, and also solves the problems of coarse-grained method in multi-scale analysis. RCMAAPE is the most suitable for performance degradation evaluation of rolling bearings.

(2) The proposed method effectively reflects the performance degradation process of rolling bearings. Compared with other methods, it is more sensitive to the early failure of the bearing and can reflect the early degradation process of the bearing earlier. It is an effective index to characterize the performance degradation of the bearing. Refined composition is a process of averaging multiple sets of data, so RCMAAPE has better stability.

(3) RCMAAPE is combined with PSO-LibSVM and applied to experimental data to verify the effectiveness of RCMAAPE in fault classification of rolling bearings. The result shows that the fault recognition rate reaches 100%, which can not only identify different types of faults well, but also can identify different severity of faults. Especially after selecting the features, RCMAAPE only need a small number of feature vectors to effectively identify the different states of the bearing, which not only can reduce information redundancy but also lessen calculative burden.

In summary, this article applies RCMAAPE to rolling performance degradation assessment and fault classification. Both experimental results show that RCMAAPE can effectively extract the fault characteristics of bearings. The computing efficiency of RCMAAPE is also significantly higher than that of RCMSE and RCMFE, and its comprehensive performance is very good. It is a very promising fault feature extraction method. In future research, we will combine it with some intelligent algorithms, and use the proposed method to perform performance degradation assessment and fault diagnosis in other areas such as gears and transmissions.

REFERENCES

[1] C.-W. Fei, Y.-S. Choy, G.-C. Bai, and W.-Z. Tang, "Multi-feature entropy distance approach with vibration and acoustic emission signals for process feature recognition of rolling element bearing faults," *Struct. Health Monitor.*, vol. 17, no. 2, pp. 156–168, Mar. 2018.

[2] J. Tian, L. Liu, F. Zhang, Y. Ai, R. Wang, and C. Fei, "Multi-domain entropy-random forest method for the fusion diagnosis of inter-shaft bearing faults with acoustic emission signals," *Entropy*, vol. 22, no. 1, p. 57, Dec. 2019.

[3] Y. Li, X. Wang, Z. Liu, X. Liang, and S. Si, "The entropy algorithm and its variants in the fault diagnosis of rotating machinery: A review," *IEEE Access*, vol. 6, pp. 66723–66741, 2018.

[4] S. Pincus, "Approximate entropy (ApEn) as a complexity measure," *Chaos: Interdiscipl. J. Nonlinear Sci.*, vol. 5, no. 1, pp. 110–117, Mar. 1995.

[5] S. M. Pincus, "Approximate entropy as a measure of system complexity," *Proc. Nat. Acad. Sci. USA*, vol. 88, pp. 2297–2301, Mar. 1991.

[6] R. Yan and R. X. Gao, "Approximate entropy as a diagnostic tool for machine health monitoring," *Mech. Syst. Signal Process.*, vol. 21, no. 2, pp. 824–839, Feb. 2007.

[7] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *Amer. J. Physiol.-Heart Circulatory Physiol.*, vol. 278, no. 6, pp. 2039–2049, 2000.

[8] J. Liang, J.-H. Zhong, and Z.-X. Yang, "Correlated EEMD and effective feature extraction for both periodic and irregular faults diagnosis in rotating machinery," *Energies*, vol. 10, no. 10, p. 1652, Oct. 2017.

[9] G. Cheng, X. Chen, H. Li, P. Li, and H. Liu, "Study on planetary gear fault diagnosis based on entropy feature fusion of ensemble empirical mode decomposition," *Measurement*, vol. 91, pp. 140–154, Sep. 2016.

[10] L. Zhang, G. Xiong, H. Liu, H. Zou, and W. Guo, "Fault diagnosis based on optimized node entropy using lifting wavelet packet transform and genetic algorithms," *Proc. Inst. Mech. Eng., I, J. Syst. Control Eng.*, vol. 224, no. 5, pp. 557–573, Aug. 2010.

[11] M. Costa, A. L. Goldberger, and C.-K. Peng, "Multiscale entropy analysis of complex physiologic time series," *Phys. Rev. Lett.*, vol. 89, no. 6, Jul. 2002, Art. no. 068102.

[12] H. Liu and M. Han, "A fault diagnosis method based on local mean decomposition and multi-scale entropy for roller bearings," *Mechanism Mach. Theory*, vol. 75, pp. 67–78, May 2014.

[13] N.-K. Hsieh, W.-Y. Lin, and H.-T. Young, "High-speed spindle fault diagnosis with the empirical mode decomposition and multiscale entropy method," *Entropy*, vol. 17, no. 4, pp. 2170–2183, Apr. 2015.

[14] Y. Li, K. Feng, X. Liang, and M. J. Zuo, "A fault diagnosis method for planetary gearboxes under non-stationary working conditions using improved vold-Kalman filter and multi-scale sample entropy," *J. Sound Vib.*, vol. 439, pp. 271–286, Jan. 2019.

[15] A. Verma, S. Sarangi, and M. H. Kolekar, "Stator winding fault prediction of induction motors using multiscale entropy and grey fuzzy optimization methods," *Comput. Electr. Eng.*, vol. 40, no. 7, pp. 2246–2258, Oct. 2014.

[16] G. Cheng, H. Li, X. Hu, X. Chen, and H. Liu, "Fault diagnosis of gearbox based on local mean decomposition and discrete hidden Markov models," *Proc. Inst. Mech. Eng., C, J. Mech. Eng. Sci.*, vol. 231, no. 14, pp. 2706–2717, Jul. 2017.

[17] L. Zhang, G. Xiong, H. Liu, H. Zou, and W. Guo, "Bearing fault diagnosis using multi-scale entropy and adaptive neuro-fuzzy inference," *Expert Syst. Appl.*, vol. 37, no. 8, pp. 6077–6085, Aug. 2010.

[18] B. Ju, H. Zhang, Y. Liu, F. Liu, S. Lu, and Z. Dai, "A feature extraction method using improved multi-scale entropy for rolling bearing fault diagnosis," *Entropy*, vol. 20, no. 4, p. 212, Mar. 2018.

[19] S.-D. Wu, C.-W. Wu, S.-G. Lin, K.-Y. Lee, and C.-K. Peng, "Analysis of complex time series using refined composite multiscale entropy," *Phys. Lett. A*, vol. 378, no. 20, pp. 1369–1374, Apr. 2014.

[20] M. Costa and A. Goldberger, "Generalized multiscale entropy analysis: Application to quantifying the complex volatility of human heartbeat time series," *Entropy*, vol. 17, no. 3, pp. 1197–1203, Mar. 2015.

[21] Y. Jiang, C.-K. Peng, and Y. Xu, "Hierarchical entropy analysis for biological signals," *J. Comput. Appl. Math.*, vol. 236, no. 5, pp. 728–742, Oct. 2011.

[22] K. Zhu, X. Song, and D. Xue, "A roller bearing fault diagnosis method based on hierarchical entropy and support vector machine with particle swarm optimization algorithm," *Measurement*, vol. 47, pp. 669–675, Jan. 2014.

[23] W. Chen, J. Zhuang, W. Yu, and Z. Wang, "Measuring complexity using FuzzyEn, ApEn, and SampEn," *Med. Eng. Phys.*, vol. 31, no. 1, pp. 61–68, Jan. 2009.

[24] J. Zheng, J. Cheng, and Y. Yang, "A rolling bearing fault diagnosis approach based on LCD and fuzzy entropy," *Mechanism Mach. Theory*, vol. 70, pp. 441–453, Dec. 2013.

[25] J. Zheng, "Rolling bearing fault diagnosis based on partially ensemble empirical mode decomposition and variable predictive model-based class discrimination," *Arch. Civil Mech. Eng.*, vol. 16, no. 4, pp. 784–794, Sep. 2016.

[26] J. Zheng, J. Cheng, Y. Yang, and S. Luo, "A rolling bearing fault diagnosis method based on multi-scale fuzzy entropy and variable predictive model-based class discrimination," *Mechanism Mach. Theory*, vol. 78, pp. 187–200, Aug. 2014.

[27] Y. Li, M. Xu, R. Wang, and W. Huang, "A fault diagnosis scheme for rolling bearing based on local mean decomposition and improved multiscale fuzzy entropy," *J. Sound Vib.*, vol. 360, pp. 277–299, Jan. 2016.

[28] H. Zhao, M. Sun, W. Deng, and X. Yang, "A new feature extraction method based on EEMD and multi-scale fuzzy entropy for motor bearing," *Entropy*, vol. 19, no. 1, p. 14, Dec. 2016.

[29] J. Zheng, H. Pan, and J. Cheng, "Rolling bearing fault detection and diagnosis based on composite multiscale fuzzy entropy and ensemble support vector machines," *Mech. Syst. Signal Process.*, vol. 85, pp. 746–759, Feb. 2017.

[30] Y. Li, M. Xu, H. Zhao, and W. Huang, ''Hierarchical fuzzy entropy and improved support vector machine based binary tree approach for rolling bearing fault diagnosis,'' *Mechanism Mach. Theory*, vol. 98, pp. 114–132, Apr. 2016.
[31] C. Bandt and B. Pompe, ''Permutation entropy: A natural complexity measure for time series,'' *Phys. Rev. Lett.*, vol. 88, no. 17, Apr. 2002, Art. no. 174102.
[32] X. An and L. Pan, ''Bearing fault diagnosis of a wind turbine based on variational mode decomposition and permutation entropy,'' *Proc. Inst. Mech. Eng., O, J. Risk Rel.*, vol. 231, no. 2, pp. 200–206, Apr. 2017.
[33] M. Kuai, G. Cheng, Y. Pang, and Y. Li, ''Research of planetary gear fault diagnosis based on permutation entropy of CEEMDAN and ANFIS,'' *Sensors*, vol. 18, no. 3, p. 782, Mar. 2018.
[34] Z. Shi, W. Song, and S. Taheri, ''Improved LMD, permutation entropy and optimized K-Means to fault diagnosis for roller bearings,'' *Entropy*, vol. 18, no. 3, p. 70, Feb. 2016.
[35] J. Liu, Y. Hu, B. Wu, Y. Wang, and F. Xie, ''A hybrid generalized hidden Markov model-based condition monitoring approach for rolling bearings,'' *Sensors*, vol. 17, no. 5, p. 1143, May 2017.
[36] Y. Gao, F. Villecco, M. Li, and W. Song, ''Multi-scale permutation entropy based on improved LMD and HMM for rolling bearing diagnosis,'' *Entropy*, vol. 19, no. 4, p. 176, Apr. 2017.
[37] L.-Y. Zhao, L. Wang, and R.-Q. Yan, ''Rolling bearing fault diagnosis based on wavelet packet decomposition and multi-scale permutation entropy,'' *Entropy*, vol. 17, no. 12, pp. 6447–6461, Sep. 2015.
[38] G. Tang, X. Wang, and Y. He, ''A novel method of fault diagnosis for rolling bearing based on dual tree complex wavelet packet transform and improved multiscale permutation entropy,'' *Math. Problems Eng.*, vol. 2016, May 2016, Art. no. 5432648.
[39] Z. Wang, W. Zhao, W. Du, N. Li, and J. Wang, ''Data-driven fault diagnosis method based on the conversion of erosion operation signals into images and convolutional neural network,'' *Process Saf. Environ. Protection*, vol. 149, pp. 591–601, May 2021.
[40] B. Keshtegar, M. Bagheri, C. Fei, C. Lu, O. Taylan, and D. Thai, ''Multi-extremum-modified response basis model for nonlinear response prediction of dynamic turbine blisk,'' *Eng. Comput.*, 2021.
[41] C.-W. Fei, H. Li, H.-T. Liu, C. Lu, L.-Q. An, L. Han, and Y.-J. Zhao, ''Enhanced network learning model with intelligent operator for the motion reliability evaluation of flexible mechanism,'' *Aerosp. Sci. Technol.*, vol. 107, Dec. 2020, Art. no. 106342.
[42] C. Lu, C.-W. Fei, H.-T. Liu, H. Li, and L.-Q. An, ''Moving extremum surrogate modeling strategy for dynamic reliability estimation of turbine blisk with multi-physics fields,'' *Aerosp. Sci. Technol.*, vol. 106, Nov. 2020, Art. no. 106112.
[43] C.-W. Fei, H. Li, H.-T. Liu, C. Lu, B. Keshtegar, and L.-Q. An, ''Multilevel nested reliability-based design optimization with hybrid intelligent regression for operating assembly relationship,'' *Aerosp. Sci. Technol.*, vol. 103, Aug. 2020, Art. no. 105906.
[44] J. Tian, S.-G. Wang, J. Zhou, Y.-T. Ai, Y.-W. Zhang, and C.-W. Fei, ''Fault diagnosis of intershaft bearing using variational mode decomposition with TAGA optimization,'' *Shock Vib.*, vol. 2021, Jan. 2021, Art. no. 8828317.
[45] Z. Wang, N. Yang, N. Li, W. Du, and J. Wang, ''A new fault diagnosis method based on adaptive spectrum mode extraction,'' *Struct. Health Monitor.*, Jan. 2021, Art. no. 147592172098694.
[46] J. Yu, ''Health condition monitoring of machines based on hidden Markov model and contribution analysis,'' *IEEE Trans. Instrum. Meas.*, vol. 61, no. 8, pp. 2200–2211, Aug. 2012.
[47] X. Zhu, Y. Zhang, and Y. Zhu, ''Bearing performance degradation assessment based on the rough support vector data description,'' *Mech. Syst. Signal Process.*, vol. 34, nos. 1–2, pp. 203–217, Jan. 2013.
[48] Y. Pan, J. Chen, and X. Li, ''Bearing performance degradation assessment based on lifting wavelet packet decomposition and fuzzy c-means,'' *Mech. Syst. Signal Process.*, vol. 24, no. 2, pp. 559–566, Feb. 2010.
[49] L.-K. Song, G.-C. Bai, and C.-W. Fei, ''Probabilistic LCF life assessment for turbine discs with DC strategy-based wavelet neural network regression,'' *Int. J. Fatigue*, vol. 119, pp. 204–219, Feb. 2019.
[50] H. Azami and J. Escudero, ''Amplitude-aware permutation entropy: Illustration in spike detection and signal segmentation,'' *Comput. Methods Programs Biomed.*, vol. 128, pp. 40–51, May 2016.
[51] Y. Chen, T. Zhang, W. Zhao, Z. Luo, and H. Lin, ''Rotating machinery fault diagnosis based on improved multiscale amplitude-aware permutation entropy and multiclass relevance vector machine,'' *Sensors*, vol. 19, no. 20, p. 4542, Oct. 2019.
[52] H. Qiu, J. Lee, J. Lin, and G. Yu, ''Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics,'' *J. Sound Vib.*, vol. 289, nos. 4–5, pp. 1066–1090, Feb. 2006.
[53] Y. Qian, R. Yan, and S. Hu, ''Bearing degradation evaluation using recurrence quantification analysis and Kalman filter,'' *IEEE Trans. Instrum. Meas.*, vol. 63, no. 11, pp. 2599–2610, Nov. 2014.
[54] K. Zhu and X. Song, ''Cross-fuzzy entropy-based approach for performance degradation assessment of rolling element bearings,'' *Proc. Inst. Mech. Eng., E, J. Process Mech. Eng.*, vol. 232, no. 2, pp. 173–185, Apr. 2018.
[55] (Nov. 29, 2017). *Case Western Reserve University Bearings Data Set*. [Online]. Available: http://www.eecs. cwru.edu/laboratory/bearing
[56] D. Cai, C. Zhang, and X. He, ''Unsupervised feature selection for multi-cluster data,'' in *Proc. 16th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD)*, 2010, pp. 333–342.
[57] K. Zhu, L. Chen, and X. Hu, ''Rolling element bearing fault diagnosis based on multi-scale global fuzzy entropy, multiple class feature selection and support vector machine,'' *Trans. Inst. Meas. Control*, vol. 41, no. 14, pp. 4013–4022, Oct. 2019.
[58] C. Chih-Chung and L. Chih-Jen, ''LIBSVM: A library for support vector machines,'' *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, p. 2, 2011.
[59] X. Yan and M. Jia, ''Intelligent fault diagnosis of rotating machinery using improved multiscale dispersion entropy and mRMR feature selection,'' *Knowl.-Based Syst.*, vol. 163, pp. 450–471, Jan. 2019.

YOUSHUO SONG received the Ph.D. degree in engineering from Zhejiang University, China, in 2014. He worked with the University of Shanghai for Science and Technology, where he is currently a Lecturer. His research interests include nonlinear dynamics analysis, signal processing, condition monitoring, and fault diagnosis of rotating machinery.

WEIYU WANG received the B.S. degree in mechatronics from the Jiangsu University of Technology, China, in 2019. He is currently pursuing the M.Sc. degree in mechanical engineering with the University of Shanghai for Science and Technology. His research interests include nonlinear signal processing and fault diagnosis of rolling bearing.