Springer

KSME

Journal of Mechanical Science and Technology 33 (1) (2019) 109~120
www.springerlink.com/content/1738-494x(Print)/1976-3824(Online)
DOI 10.1007/s12206-018-1211-8

Refined composite multiscale fuzzy entropy: Localized defect detection of rolling element bearing †

Yongjian Li[1,2,3], Bingrong Miao[2,*], Weihua Zhang[2], Peng Chen[4], Jihua Liu[1] and Xiaoliang Jiang[5]

[1]*School of Railway Tracks and Transportation, Wuyi University, Jiangmen, 529020, China*
[2]*State Key Laboratory of Traction Power, Southwest Jiaotong University, Chengdu 610031, China*
[3]*Key Laboratory of Automotive Measurement, Control and Safety, Xihua University, Chengdu 610039, China*
[4]*School of Mechanical Engineering, Southwest Jiaotong University, Chengdu 610031, China*
[5]*College of Mechanical Engineering, Quzhou University, Quzhou, 324000, China*

(Manuscript Received March 20, 2018; Revised July 28, 2018; Accepted August 28, 2018)

Abstract

We proposed an appealing method based on refined composite multiscale fuzzy entropy (RCMFE), infinite feature selection (Inf-FS) algorithm, and support vector machine (SVM) for implementing localized defect detection to keep the downtime and extended damage caused by incipient failure of bearing at a minimum. As a useful approach, multiscale fuzzy entropy (MFE) was utilized to measure the complexity and dynamic changes of signals. However, an inaccurate entropy value would be yielded with the increase of scale factor. Here, as an improvement version of MFE, the RCMFE was proposed to address the shortcomings in the case of short time series. For this novel method, we conducted a full investigation of the effects and robustness by comparing the proposed method with two other entropy-based approaches using synthetic signals and real data. Results indicate that the proposed algorithm outperforms the other approaches in terms of reliability and stability. The RCMFE values of bearing signals from one healthy condition and seven fault states are calculated as diagnostic information. Moreover, an intelligent fault identification method was constructed by combining the Inf-FS algorithm and SVM classifier. Experimental results show the usefulness of the proposed strategy.

Keywords: Defect detection; Infinite feature selection; Multiscale fuzzy entropy; Refined composite technique; Rolling bearing condition

1. Introduction

Rolling element bearing (REB) is the most common part of a mechanical system in engineering. A running REB is likely to break down due to overload capacity or complicated condition, thereby causing catastrophic system failure and massive financial loss. Therefore, accurate health monitoring for running bearings is a critical and ongoing challenge. Defective bearings lead to considerably typical phenomena, including periodic noise and vibration signals [1-3], whereas healthy REBs show no such regular recordings. Thus, effective methodologies that are affordable and efficient are needed to detect incipient defect by signal processing during operation.

In the past decades, a number of traditional techniques are presented, such as time- and frequency-domain methods [4, 5], to perform fault detection. The traditional approaches to capture the fault feature of collected recordings may fail to yield for nonlinear and nonstationary signals. To resolve these drawbacks, advanced concepts based on time-frequency analysis, known as ensemble empirical mode decomposition and wavelet transform [1, 6-8], have been developed as a popular method for identifying the characteristic frequency of fault REB from complex frequency components. However, some limitations occur in noise removal and fault feature extraction. For example, the procedure of diagnosis remains complicated, and its recognition rate is relatively low [9].

Significant effort has been devoted to finding an efficient nonlinear dynamic approach that can characterize fault information from the perspective of statistical parameter estimation. Important advances in feature extraction have been conducted, with a large and increasing studies on techniques, such as multiscale permutation (MPE) [10-12], Lempel-Ziv complexity [13], multiscale entropy (MSE) [14-19], multiscale fuzzy entropy (MFE) [20-22]. Recent studies have employed these methods to quantify the complexity of mechanical signals and biomedical recordings. The results reveal important signatures of working state or disease and a potentially novel way for monitoring and detecting abnormal condition.

As a powerful alternative, fuzzy entropy (FuzzyEn) [20] was introduced to overcome the issue of the sensitivity of the estimation of sample entropy (SampEn) to a threshold value.

*Corresponding author. Tel.: +86 28 87600868
E-mail address: brmiao@home.swjtu.edu.cn
† Recommended by Associate Editor Gyuhae Park

To address the limitations of a single time scale, the MSE and MFE, which are combined multiscale concept and entropy-based approach, were developed to utilize their relative advantages. In other words, we can implement entropy measures over a range of scales instead of only one. Some instances may further illustrate the importance of the multiscale concept [16, 22]. For example, the entropy value of 1/f noise is theoretically higher than that of a white Gaussian noise (WGN) signal because of long-range correlations [17]. Although analysis results agree with the expectation in theory when the MSE method can be used, contradictory conclusions at one time scale may still be achieved. Moreover, the MSE algorithm can be time-consuming and have the disadvantage of coarse-graining, which results in an imprecise measurement of entropy for short data points. Therefore, Wu et al. [18] proposed an improved version of the MSE algorithm to yield an accurate value and enhance computational efficiency.

In the present work, an improved MFE method termed as refined composite MFE (RCMFE) is developed based on the superiority of the MFE [22] and refined composite multiscale algorithm [18] to address the aforementioned issues. We evaluate the capability of the proposed approach using simulated time series and apply the technique to investigate in rolling bearing fault detection. Thus, the characteristics obtained in normal and faulty conditions should feed into a useful classifier to identify the defect automatically. However, the feature vector that contains the defect information is inevitably of high-dimensional data and with redundant information, which can also adversely affect computational resources and recognition accuracy. Roffo et al. [23, 24] introduced a powerful tool called infinite feature selection (Inf-FS) to refine the features according to scores. In the present work, we use the Inf-FS algorithm to select the important features by assessing the importance of each feature. The classifier is utilized to discriminate the bearing status after reconstructing important and low-dimensional features. Support vector machine (SVM) [25, 26], a well-known technique that is successfully applied in different research fields, such as pattern recognition and data mining, is utilized to identify the working condition of REBs.

This paper is organized as follows. Sec. 2 briefly describes and evaluates the RCMFE algorithm. Sec. 3 reviews the Inf-FS method. Sec. 4 presents the SVM algorithm. The proposed fault diagnosis strategy is validated using experimental data in Sec. 5. Finally, Sec. 6 concludes this study.

2. Refined MFE

2.1 Entropy methods

To overcome the bias problem in approximate entropy, SampEn was developed in 2000 by Richman and Moorman to measure the complexity of dynamical systems [14]. In the present work, we briefly describe the key steps of SampEn statistics procedure as follows [14, 16, 17].

(1) Given a discrete time series with N data points $u(s)=\{u(1),u(2),\cdots,u(N)\}$, we form $N-m+1$ vectors as follows:

$$\begin{aligned} &U_s^m=\{u(s),u(s+1),\cdots,u(s+m-1)\} \\ &s=1,2,\cdots,N-m,N-m+1 \end{aligned} \tag{1}$$

where m is the dimensionality of vectors.

(2) The distance between two vectors, namely, template U_s^m and template match U_t^m, is calculated as

$$\begin{aligned} &d\left[U_s^m,U_t^m\right]=\max\left\{\left|u(s+k)-u(t+k)\right|\right\} \\ &0\le k\le m-1;t=1,2,\cdots,N-m+1. \end{aligned} \tag{2}$$

A match exists when $d\left[U_s^m,U_t^m\right]<r$, where r denotes tolerance.

(3) To obtain the probability of the matched vectors in $\left[1,N-m+1\right]$, the function $B_s^m(r)$ should be defined as

$$\begin{aligned} &B_s^m(r)=\frac{\psi(d_{st}<r)}{N-m+1},s,t=1,2,\cdots N-m,s\ne t \\ &\psi(d)=\begin{cases}1 & d>0\\ 0 & d\le 0\end{cases}. \end{aligned} \tag{3}$$

The function can be calculated as

$$B^m(r)=(N-m)^{-1}\sum_{s=1}^{N-m}B_s^m(r). \tag{4}$$

Similarly, we define $B^{m+1}(r)$ in case of $m+1$ embedding dimension.

(4) Finally, SampEn is expressed as follows:

$$SampEn(m,r,N)=-\left(\ln B^{m+1}(r)-\ln B^m(r)\right). \tag{5}$$

The SampEn concept based on Heaviside function is of seriously limited use when applied in engineering time series. The function, which leads to a type of conventional two-state classifier, cannot adapt to the ambiguousness of the boundaries between classes. To address this shortcoming, Chen et al. [20] developed FuzzyEn by introducing fuzzy sets, that is, the Heaviside function was replaced by an exponential function to form a fuzzy function. The key steps are expressed as follows.

(1) For a random sample of N observations, $u(s)=\{u(1),u(2),\cdots,u(N)\}$, with parameters including embedding dimension m, power n, and tolerance r, a sensible vector can be formed as

$$\begin{aligned} &U_s^m=\{u(s),u(s+1),\cdots,u(s+m-1)\}-u_0(s) \\ &s=1,2,\cdots,N-m,N-m+1 \end{aligned} \tag{6}$$

where $u_0(s)$ is the mean value of the u values for baseline removal.

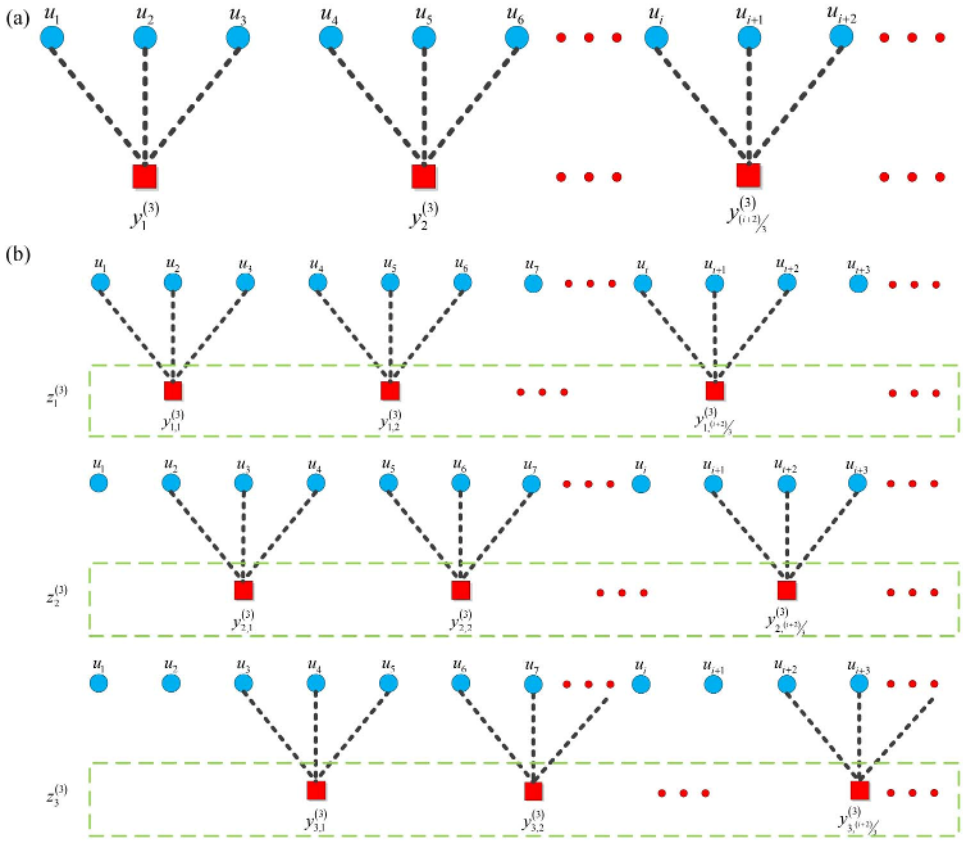


Fig. 1. Schematic of coarse-graining for τ = 3: (a) Traditional process; (b) refined composite process.

$$u_0(s)=\sum_{k=0}^{m-1}\frac{u(s+k)}{m}. \tag{7}$$

(2) The distance between U_s^m and U_t^m is calculated similar to SampEn, as shown as follows:

$$\begin{aligned} d_{st} &= d\left[U_s^m, U_t^m\right] \\ &= \max\left\{\left|[u(s+k)-u_0(s)]-[u(t+k)-u_0(t)]\right|\right\} \\ 0 &\le k \le m-1; s,t=1,2,\cdots,N-m. \end{aligned} \tag{8}$$

(3) The similarity degree D_{st} is defined via a fuzzy function.

$$D_{st}=\mu\left(d_{st},n,r\right)=\exp\left(-\ln 2(d_{st}/r)^n\right). \tag{9}$$

(4) The function ϕ^m can be expressed as follows:

$$\phi^m(n,r)=(N-m)^{-1}\sum_{s=1}^{N-m}\left((N-m-1)^{-1}\sum_{t=1,s\neq t}^{N-m}D_{st}^m\right). \tag{10}$$

Similarly, we define the ϕ^{m+1} in case of $m+1$ dimension.

(5) Finally, we can define the FuzzyEn writing as the negative natural logarithm of the ratio of $\phi^m(n,r)$ and $\phi^{m+1}(n,r)$, as shown as follows:

$$FuzzyEn(m,n,r,N)=-\ln\frac{\phi^{m+1}}{\phi^m}. \tag{11}$$

2.2 MSE using coarse-graining

To measure the information inherent in multiscale time series, Costa et al. [16, 17] imported the concept of coarse-graining and combined it with SampEn to propose a novel approach called MSE. Later, Zheng et al. [22] introduced MFE, which is similar to MSE. Coarse-graining is implemented as follows.

(1) A time series is represented as $u(s)=\{u(1),u(2),\cdots,u(N)\}$, which can be preprocessed at a scale factor by computing the mean of consecutive samples to form a new vector, as shown as follows:

$$y_\tau(i)=\frac{1}{\tau}\sum_{(i-1)\tau}^{i\tau}u(i),1\le i\le\left\lfloor N/\tau\right\rfloor \tag{12}$$

where τ is the scale factor. $\lfloor b\rfloor$ denotes the floor function that is the largest integer but is smaller than b.

(2) The FuzzyEn of each coarse-grained time series is calculated according to the steps described in Sec. 2.1.2 to achieve the MFE results.

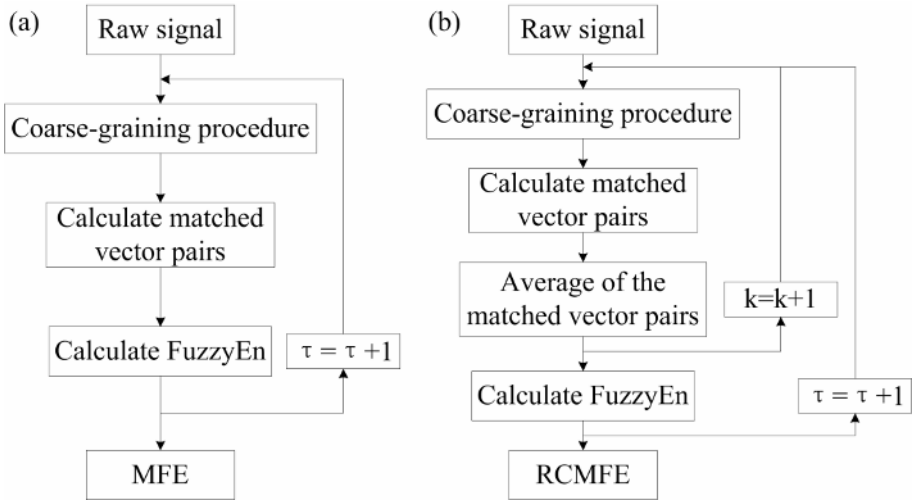


Fig. 2. Flowcharts of the MFE and RCMFE methods.

2.3 RCMSE

Notably, the MSE algorithm takes a coarse-grained concept to compute the entropy values in multiscale; however, this approach has some drawbacks [18]. First, MSE is heavily dependent on the size of data samples. The size of data points divided in the coarse-grained step at a large scale will be small, in which case the entropy values with large error have been derived from the original time series. Second, for some dataset, coarse-graining is inconsistent. That is, the calculation procedure of u_1, u_2 and u_3 should remain the same with u_2, u_3 and u_4 (Fig. 1); however, separation occurs between u_3 and u_4 in the traditional MSE method. In view of the relative advantages of FuzzyEn approach and multiscale concept, the RCMFE is developed in this study to address the aforementioned issues. The flowcharts of the conventional MFE and RCMFE methods are shown in Fig. 2. The estimation using RCMFE is achieved in two steps.

(1) For a time series $u(s)=\{u(1),u(2),\cdots,u(N)\}$, the improved coarse-grained forms can be obtained as follows:

$$y_\tau(i,k)=\frac{1}{\tau}\sum_{(i-1)\tau+k}^{i\tau+k-1}u(i),1\le i\le\left\lfloor N/\tau\right\rfloor,1\le k\le\tau\,. \tag{13}$$

Here, τ coarse-grained time series will be generated; however, only one time series is received in the MFE algorithm.

(2) For a given scale factor τ, the two defined functions $\phi_{k,\tau}^m$ and $\phi_{k,\tau}^{m+1}$ are calculated. Then, the mean of $\phi_{k,\tau}^m$ and $\phi_{k,\tau}^{m+1}$ denoted as $\overline{\phi}_{k,\tau}^m$ and $\overline{\phi}_{k,\tau}^{m+1}$ on $1\le k\le\tau$, respectively, is computed. Finally, RCMFE can be expressed as

$$RCMFE(x,\tau,m,n,r)=-\ln\frac{\overline{\phi}_{k,\tau}^{m+1}}{\overline{\phi}_{k,\tau}^{m}}\,. \tag{14}$$

The difference between RCMFE and MFE is that the former is conducted by two critical steps, that is, improving coarse-graining and the other averaging the value in FuzzyEn calculation, such that stable entropy statistics can be obtained.

3. Defect diagnosis strategy

3.1 Inf-FS

In preliminary analysis, the RCMFE offers a new method for revealing the changes of a mechanical system via feature extraction. However, redundant information can be extracted along with the useful characteristics, and a FuzzyEn-based feature vector can be considered as high-dimensional dataset, thereby resulting in long time-consumption and decrease in the accuracy rate of pattern recognition. Inf-FS algorithm [23] is used for selecting the most important characteristics with effective defective features from the calculated scales to overcome the problem of dimensional disaster and enhance the diagnostic accuracy of REBs.

As filter-based methodology, the Inf-FS is mainly founded on a simple cross-validation strategy to select the best *k* features in an unsupervised manner. For a feature dataset $E=\{f_1,\cdots,f_n\}$, $x\in R$ indicates a sample of *f*. Let s_i denote the final energy scores of the *i*th feature. The major process can be expressed as follows [23].

(1) An undirected fully connected graph *G* is constructed, which is then represented as adjacency matrix *A*. The weighted edges can be specified between two pairwise measures that link f_i and f_j.

(2) The energy term can be expressed as follows:

$$\begin{aligned}&A(i,j)=\alpha\sigma_{ij}+(1-\alpha)\varepsilon_{ij}\\&\sigma_{ij}=\max(\sigma_i,\sigma_j)\\&\varepsilon_{ij}=1-\left|Spearman(f_i,f_j)\right|\end{aligned} \tag{15}$$

where α is a loading coefficient [0, 1], σ_i is the standard deviation (SD) of the samples $\{x\}$, and *Spearman* denotes the

Spearman's rank correlation coefficient.

(3) The geometric matrix power series can be obtained as follows:

$$\breve{S} = \left(\mathrm{I} - rA\right)^{-1} - \mathrm{I} \tag{16}$$

where r is a regularization factor.

(4) Finally, the energy scores for each feature are defined as follows:

$$\breve{s}\left(i\right) = \left[Se\right]_i . \tag{17}$$

Similarly, the candidate features can be ranked in a descending order according to the scores. The larger the score is, the more important the information is. In this study, the featured approach should be applied to evaluate the importance of FuzzyEn-based features. Furthermore, the most relevant features are selected to feed into the classifier.

3.2 SVM classifier

On the basis of the features received by the aforementioned approaches, a classifier is used to study the intelligent diagnosis of the bearing condition. In terms of pattern recognition and prediction, the SVM algorithm is frequently used for classification in various engineering fields [25, 26], such as biomedicine, imaging, and mechanical study. As previously stated, various kernel functions, known as Gaussian (or RBF), polynomial, and sigmoid kernel, exist. Here, the RBF is utilized due to its promising capability. However, the recognition results may be affected considerably by two parameters, namely regularization parameter and kernel width. We should tune the values of the two vital parameters based on previous studies on rotation machinery fault diagnosis [26]. Thus, particle swarm optimization (PSO) algorithm is used to determine the two values [27].

3.3 Proposed method

Given the superiorities of RCMFE, Inf-FS, and PSO-SVM, we propose an intelligent bearing fault diagnosis approach to sort multiple and normal types of faulty bearing. The overall step of the proposed method is shown in Fig. 3. The further details of the diagnosis approach are summarized as follows:

(1) Samples of raw data are taken by accelerometer at a particular sampling frequency under various operating conditions.

(2) The RCMFE values are calculated for different vibration signals with parameter selection and scale factor $\tau_{\max} = 20$. Twenty features, including the defect information of the 1D signals of rolling bearing, are measured.

(3) The 20 fault features are sorted by Inf-FS algorithm in ascending order in terms of their importance and relationship.

(4) The first five features can be selected by the least scores that describe the bearing condition to form the new feature vector for fault diagnosis.

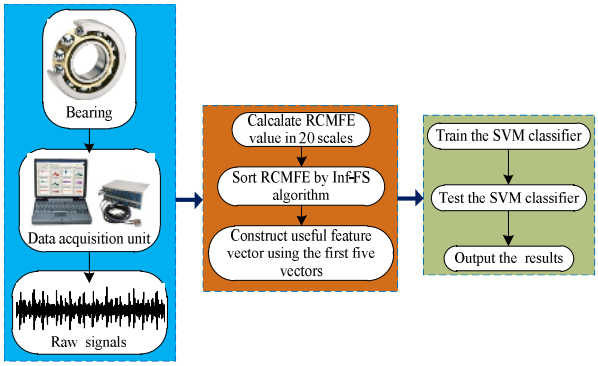


Fig. 3. Framework of the proposed method.

(5) The training and testing samples are fed into the classifier called PSO-SVM to implement automatic fault diagnosis.

4. Evaluation of synthetic signal

4.1 Parameter selection

Here, the key parameters used in the entropy algorithm should be determined. Time lag δ, embedding dimension m, fuzzy power n, threshold r, and scale factor τ exist. The time lag δ is a key parameter for coarse-graining. However, standard criteria for selecting this parameter are unavailable; thus, for simplicity of computation, we fix time lag as 1 [16, 22]. The embedding dimension m determines how much information the reconstructed time series contains. An overly large m will need a large number of data points, which results in increased time consumption; by contrast, an overly small m will cause insufficient information. For a trade-off between computational efficiency and performance, a good overall selection of embedding dimension is 2. The fuzzy power n is a parameter that controls the boundary gradient of similar tolerance. Usually, it should be sufficiently large to capture the detailed information of interest in the boundary. However, an extremely large n will create a large gradient, resulting in the lack of detailed information. Hence, a reasonable value of n is 2. Threshold r is one of the dominant parameters of the entropy-based approach because it controls the similarity between the two matched vector components. For larger values of r, fewer vectors can be used for matching, which results in difficulty in recognition. A smaller value of r will affect the entropy measure with insufficient accuracy. Thus, a good value is 0.15 SD, where SD is the standard deviation of the given signal. A scale factor is determined by the trade-off between complexity and reliability result; thus, the value of this parameter is set experimentally as 20 [22].

4.2 Comparative study of MFE, RCMSE and RCMFE

In this section, the performances of the conventional MFE, RCMSE, and the proposed RCMFE are evaluated based on the two typical types of synthetic time series, namely, WGN

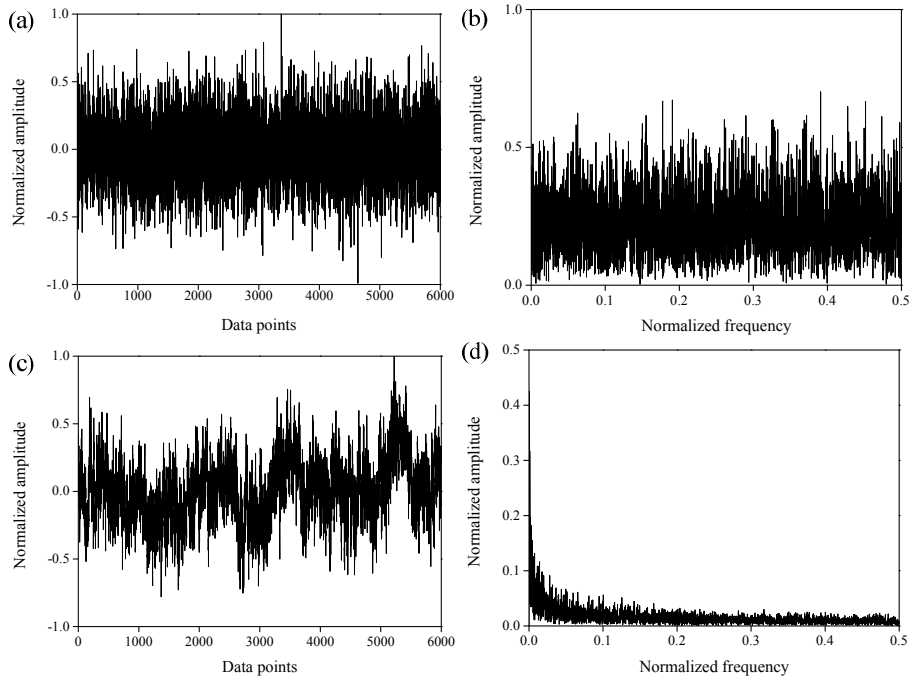


Fig. 4. (a) Time domain of white noise; (b) frequency spectra of white noise; (c) time domain of 1/*f* noise; (d) frequency spectra of 1/*f* noise.

and 1/*f* noise [12, 16, 22]. The waveforms and frequency spectra are shown in Fig. 4. First, the relationship between the length and the accuracy of the MFE, RCMSE and RCMFE algorithms should be investigated. By generating 100 independent samples with different data points (N = 1024, 2048, 4096, 8192), we calculate the entropy values for the two noise signals. The SDs of entropy measures are shown in Fig. 5. As expected, in the case of WGN and 1/*f* noise, the SDs of the entropy measures of the three entropy-based methods decreases at scales of 1 to 20 with the increase in data length. In addition, the fluctuation amplitude of the SDs in each scale is effectively reduced. Hence, the accuracy of entropy statistics is affected by the size of data sample. In other words, the longer the length of the time series, the higher the accuracy of the calculation. However, a large value of N results in increased computational cost. A reasonable data length with 2048 data points is used in this study considering the computational efficiency and accuracy.

In the case of the 2048 dataset, the entropy estimation of WGN noise for the three approaches is depicted in Fig. 6(a). From the figure, the entropy values exhibit a monotonically decreasing trend at each scale. The SampEn-based value obtained with RCMSE is larger than that of the FuzzyEn-based result. Notably, the curve of the entropy value acquired by RCMFE algorithm is smoother than that acquired using the traditional MFE.

Then, the SDs of the three methods are comparatively analyzed. As shown in Fig. 7(a), the RCMFE has smaller SDs in every scale factor compared with RCMSE and MFE. Moreover, the SD of MFE reaches the maximum among the three algorithms. Thus, the RCMSE, which is the improved version of MSE, estimates more stable values than MFE, whereas the MFE outperforms the conventional MSE. Furthermore, as expected theoretically, RCMFE can generate the most stable results. Notably, for a certain number of data points, the SDs render a monotonically increasing trend when the scale factor ranges from 1 to 20. This result confirms our theoretical analysis about the unstable values of entropy, which results in an extremely short coarse-grained sequence.

A similar trend is observed in the 1/*f* signal, as shown in Figs. 6(b) and 7(b). Unlike the WGN noise, the curves of entropy-based values have small variations and are approximately constant along with the increased scale. Hence, intuitively, the 1/*f* noise signals are more intrinsically complex than WGN signals; accordingly, higher scales can contain additional important information. Therefore, for WGN and 1/*f* noise signals, the RCMFE values has promising advantages that can improve the stability of entropy measures.

5. Experimental validation

5.1 Evaluation of real vibration signals

As shown in the above analysis, the RCMFE method assesses the complexity of simulated signals efficiently and suitably. The MFE, RCMSE and RCMFE are compared in terms of their capability to reveal structural differences between WGN and 1/*f* noise. We further analyze these techniques when the time series have different numbers of data points. Here, we evaluate the potential advantages of the proposed approach for vibration signals recorded from REBs. The vibration data of the bearings are collected from the CWRU laboratory [28]. As shown in Fig. 8, the experimental

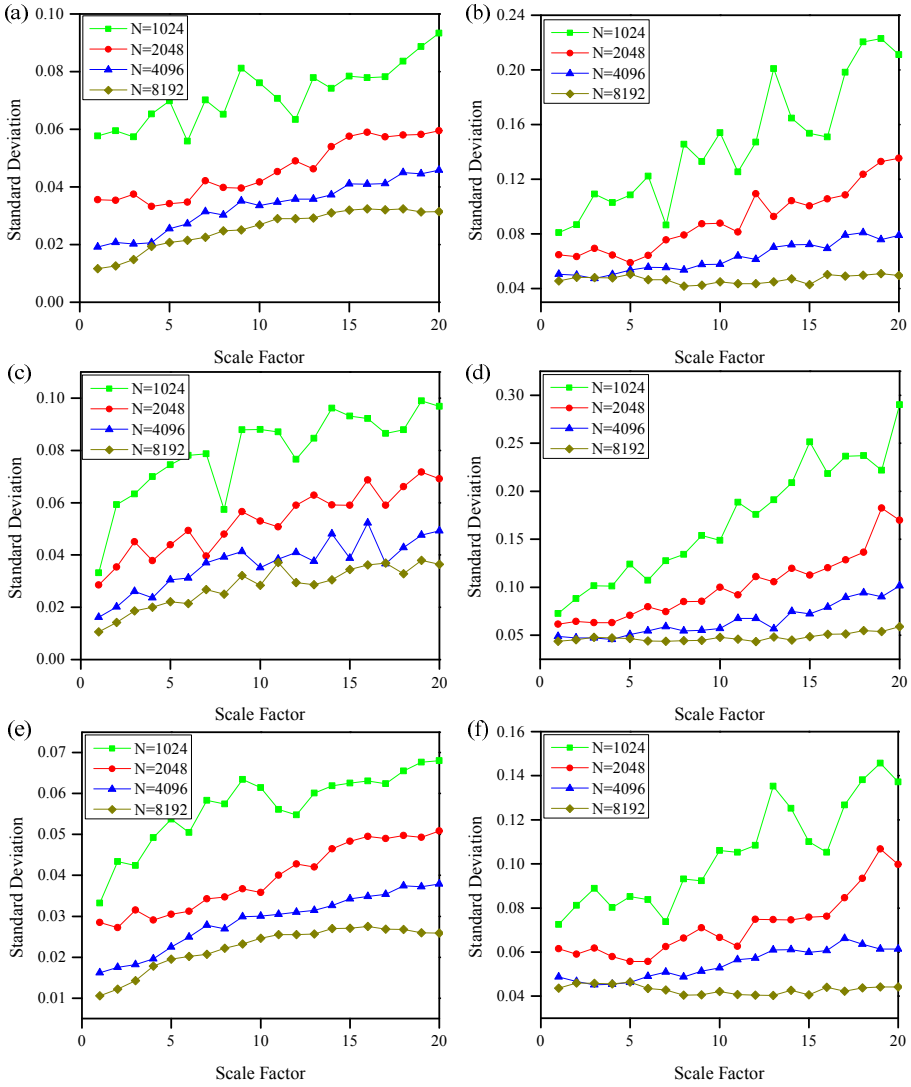


Fig. 5. SD of simulated noise with different data points: (a) RCMSE of white noise; (b) RCMSE of 1/*f* noise; (c) MFE of white noise; (d) MFE of 1/*f* noise; (e) RCMFE of white noise; (f) RCMFE of 1/*f* noise.

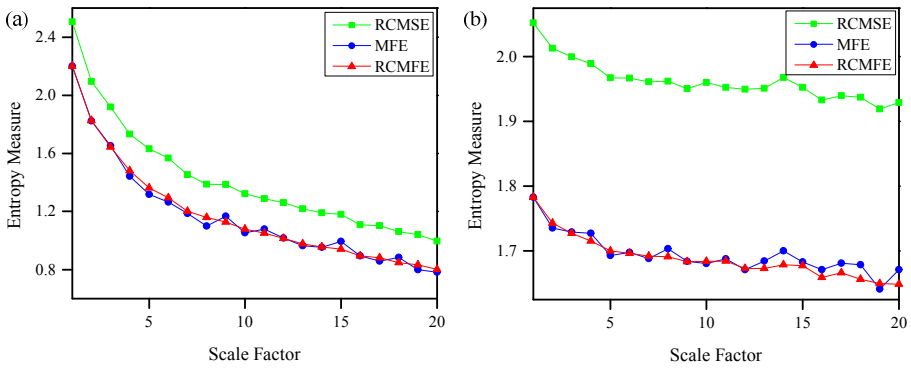


Fig. 6. RCMSE, MFE and RCMFE analyses for the time series of simulated noise: (a) White noise; (b) 1/*f* noise.

device comprises an induction motor, a dynamometer, a torque transducer (center position), and a control system (not shown). The test bearing is a deep groove ball bearing 6205-2RS JEM SKF. Here, normal and defect bearings in different positions are used for testing. The defect bearing is made using electro-discharge machining with fault dimension of 21 mil. An accelerometer is attached on the motor housing at the right position of the motor, with magnetic bases to finish the

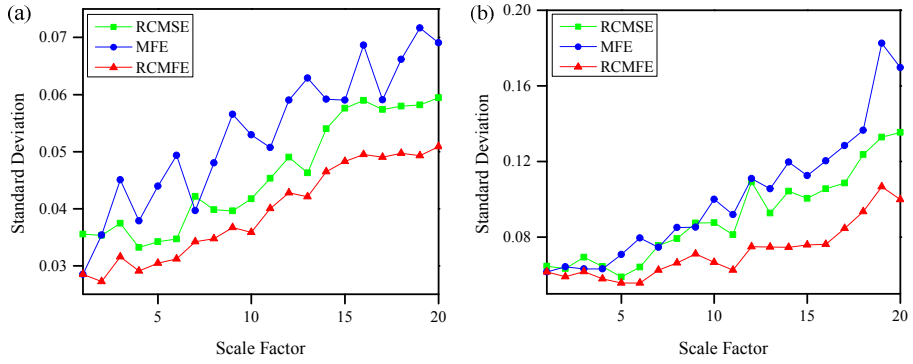


Fig. 7. SD of simulated noise: (a) White noise; (b) 1/*f* noise.

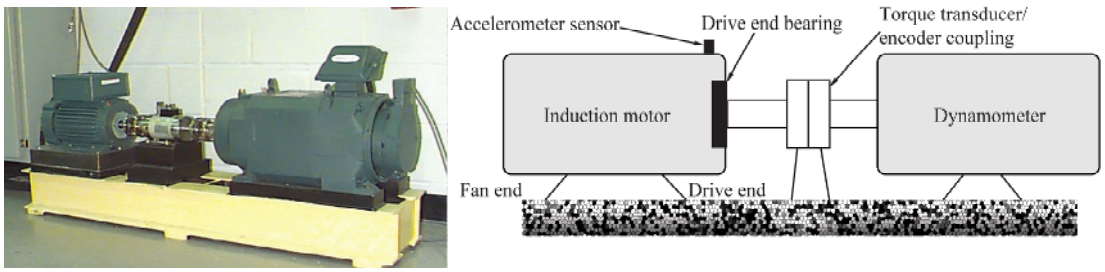


Fig. 8. Experimental setup in the Bearing Data Center of Case Western Reserve University (CWRU) [28].

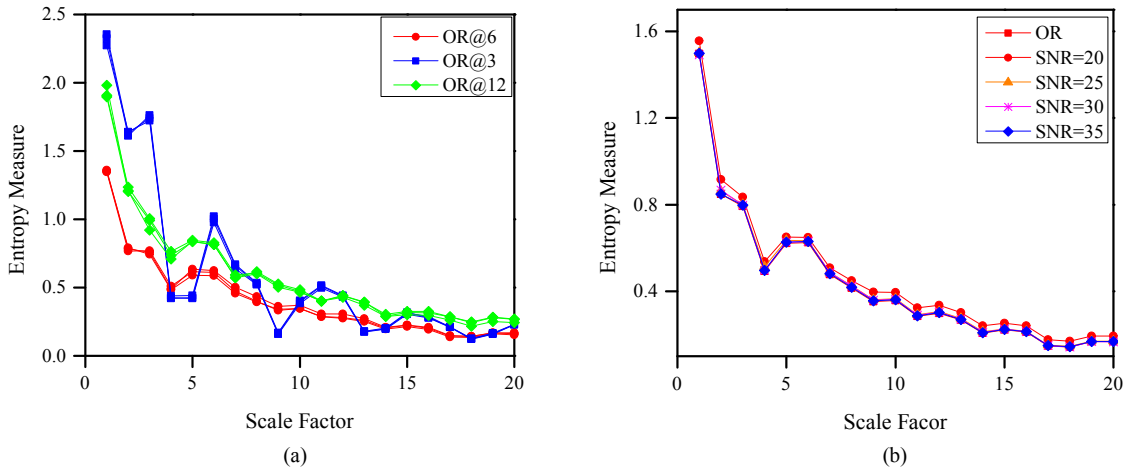


Fig. 9. RCMFE results: (a) Entropy values of OR recordings collected in different locations; (b) entropy values of OR recording with different SNRs.

raw signal gathering process. The sampling frequency of the data recorder is 12 kHz, and the rotation speed frequency is 29 Hz. Additional information about the experiment is available in the literature [28].

We study the signals of the rolling bearing of the outer race fault, from which the dataset is collected in different locations of the bearing housing, such as 6:00, 3:00 and 12:00. We focus on the load zone centered at 6:00. In the ordinary course of events, the signals recorded at the load zone should be clearer than the other positions. To analyze the differences of the recordings, feature extraction is implemented using the RCMFE model to characterize the dynamic change. The entropy measure for each location is shown in Fig. 9(a). The entropy cure of recordings obtained at the load zone maintains the highest stability over most of the temporal scales in comparison with the three locations (Fig. 9(a)), which agrees with our previous expectation. For the signals collected at 3:00, the entropy results for the scale of 1-20 perform serious fluctuations. In the case of the opposite load zone, that is, 12:00, the entropy values at each time scale exhibit more smoothness than the orthogonal position but greater than the values obtained at the load zone. Notably, the trend for the three positions from a scale of 1-20 is monotonically decreasing. This behavior shows the complexity of the outer race fault, which

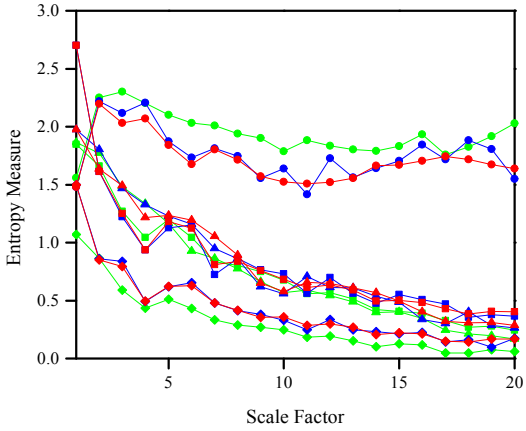


Fig. 10. Entropy-based results of REB recording under four conditions (Normal-○, IR-△, B-□, OR-◇; RCMSE-green, MFE-blue, RCMFE-red).

decreases with the increase in time scale.

To test the robustness of the given techniques to noise, we implement the comparison by adding WGN with different signal-to-noise ratio (SNR) levels. The vibration data with outer race fault are used as the original signal. The noise levels with SNRs are selected as 20, 25, 30 and 35. We then apply the RCMFE algorithm to represent the difference between the original data and the signal with additive noise. The entropy values for each signal are plotted in a scale of 1-20, as shown in Fig. 9(b). Significant differences are not observed at each scale. Notably, the results computed from data with SNR = 20 are slightly higher than those from raw data. These results demonstrate that the RCMFE has good performance, even with large noise.

For a healthy state and the other three defects in different locations, including inner race, roller element, and outer race, we characterize each vibration data with three straightforward methodologies (that is, MFE, RCMSE and RCMFE) to assess the performance of the four conditions. The differences in the results exist between normal and defective conditions, as shown in Fig. 10. Evidently, the four conditions of bearing are depicted by the three approaches with a similar trend. Similar to the previous discussion in Sec. 2, the FuzzyEn-based entropy of random signal is lower than that of SampEn-based values. However, only RCMFE-based entropy shows smoothness and stability.

The entropy values of the normal time series are larger than those of faulty signals over all temporal scales. Hence, important information of the bearing system is widely distributed in the entire scale factors. This situation illustrates that the healthy bearing recording similar to 1/*f* noise has a linear trend, thereby exhibiting more complexity and irregularity than the faulty ones. This result agrees with previous findings [22]. For each defect of different positions, the general trends of entropy measures considerably decrease with the increase in time scale (Fig. 10). This phenomenon suggests that the time series containing fault component assigns less complexity and more regularity due to the fault information induced by defective parts. The load zone position inside the bearing generates shocks between these components. The vibration signals induced by the faulty part can dominate the signals of the entire bearing system because of the major frequency with powerful energy. In addition, the periodicity of faulty signals can create low complexity and regularity. Notably the entropy estimation on the outer race defect is smaller than that on the inner race and roller faults. At the same time, the entropy of recordings derived from the roller fault is larger than the signals collected from the bearing with inner race fault. This dynamic change is observed in different motion patterns of the three components. For most of the working conditions, the outer race is fixed on an axle box instead of rotating around the axle. The vibration contains single-component failure frequency. By contrast, the inner race rotates around the axle because of interference fit, and the roller rotates around the axle and on its own axis, which may cause more complicated frequency components compared with the outer race fault. Thus, the roller fault assigns larger entropy values at most of time scales compared with the inner race defect. Therefore, the RCMFE outperforms the other algorithms in this field and achieves excellent performance in distinguishing the state differences of bearings.

5.2 Experimental results and discussion

Here, we focus on distinguishing the healthy condition of the bearings using important and meaningful characteristics. Table 1 presents the details of the experimental data. The time domain analysis of vibration signals with health and different defect types is shown in Fig. 11. On the basis of these different signals and the multiscale FuzzyEn approach, a total of 400 samples are computed for classification. One-fifth of the total 400 samples is randomly selected as the training dataset, and the remaining 320 samples are the testing dataset. The experiment is an eight-class identification problem.

However, the original RCMFE-based vector is not taken as the diagnostic feature set due to redundant information with high dimension of extracted RCMFE values at time scales between 1 and 20. The high-dimensional fault feature vector will cause computational errors and affect the diagnostic results. To alleviate these problems, the Inf-FS algorithm is used to select the significantly independent characteristics by feature importance to depict the difference of the vibration signals. After ranking the scores, the reordered RCMFE can be expressed as follows:

$$FS_6 > FS_2 > FS_1 > FS_4 > FS_{11} > FS_5 > FS_{15} > FS_8 > FS_{12} > FS_7 > FS_9 > FS_{19} > FS_{18} > FS_{10} > FS_{16} > FS_{20} > FS_{13} > FS_{14} > FS_{17} > FS_3 .$$

The first five ranked features are selected to compose the new feature set. Then, a multi-fault classification algorithm named PSO-SVM with RBF kernel is used to complete the REB health condition identification.

Table 1. Description of the testing bearing data.

Fault type	Fault size (inch)	Rotation speed (rpm)	Number of training data	Number of testing data	Class label
Normal	0	1750	10	40	1
IRF	0.007	1750	10	40	2
REF	0.007	1750	10	40	3
ORF	0.007	1750	10	40	4
IRF	0.014	1750	10	40	5
IRF	0.021	1750	10	40	6
REF	0.021	1750	10	40	7
ORF	0.021	1750	10	40	8

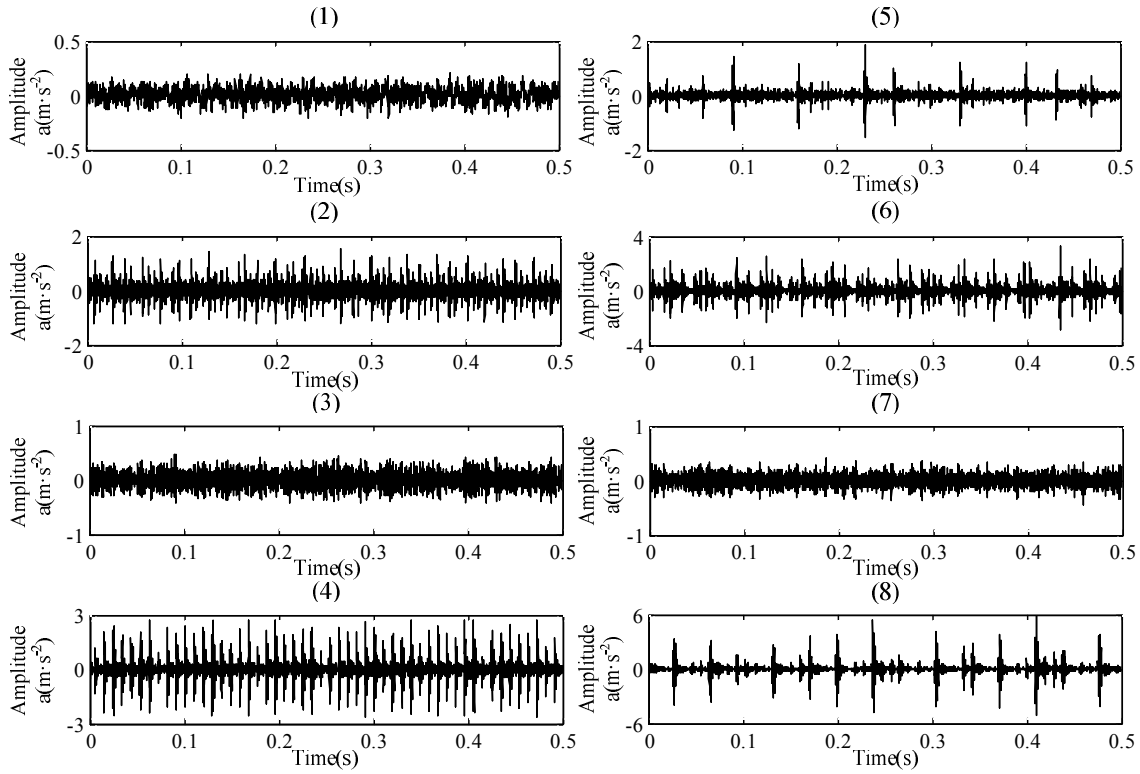


Fig. 11. Waveforms of bearing vibration signals.

The average recognition rate of classification for the proposed method is presented in Fig. 12, which indicates the training and testing outputs. As shown by the results of the PSO-SVM classifier, the proposed method can sufficiently classify different rolling bearing health conditions with 100 % accuracy. Thus, the refined composite MFE method shows good performance to discriminate the health of bearings. Moreover, we investigate the necessity of feature selection algorithm using the Inf-FS approach and the original feature vector. The number of training and testing samples is the same as above. The diagnostic results with different feature numbers are shown in Fig. 13. The accuracy rate of the Inf-FS is higher than that of raw feature vectors. Furthermore, the classification accuracy is up to 100 %, and the five most distinguishable features are selected using the Inf-FS method.

Despite the outstanding results obtained in terms of diagnostic accuracy, some limitations still exist. The achieved fault classification of the proposed method and the other studies is 100 % [21, 22]. This result may be caused by the dependence

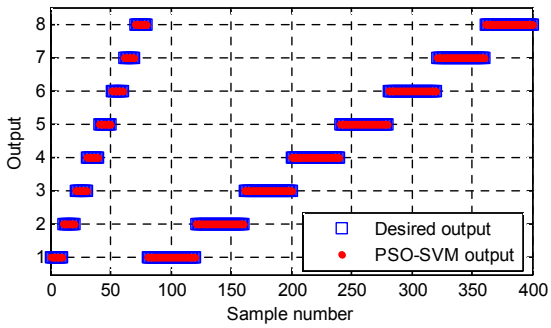


Fig. 12. Classification results using the proposed method.

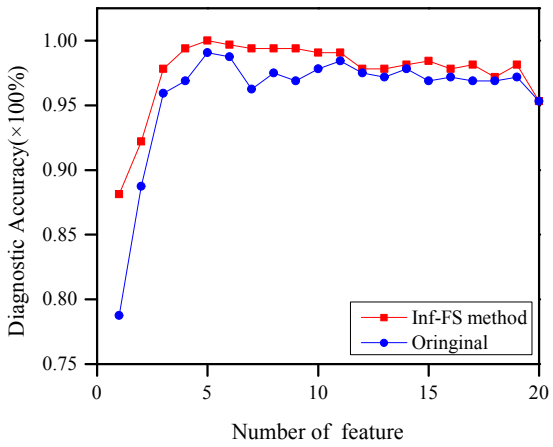


Fig. 13. Diagnostic accuracy of different feature numbers.

of the classification accuracy on the sample size of different fault conditions. Moreover, the proposed method is verified using a small number of samples, which lead to an easy 100 % recognition rate. The experimental data taken from CWRU are used for the proposed method. The fault features are extremely evident to reveal the advantages of this method. Therefore, investigations with large samples under engineering conditions must be used in future works.

6. Conclusion

In this study, we introduce a novel nonlinear statistical technique, namely, RCMFE algorithm, by addressing several shortcomings consisted in the traditional MFE approach to quantify the complexity of 1D signals. The performance of RCMFE is evaluated using synthetic time series, including uncorrelated (WGN) and correlated (1/*f*) signals in comparison with other two entropy-based methods, namely, RCMSE and MFE. The results further conform that the proposed technique has better capability in short-signal analysis than the other two algorithms. Furthermore, when the proposed method is applied to vibrational signals recorded from bearings, the results show improvement in the stability and reliability of entropy values. The fault features based on RCMFE are calculated from the different recordings of REBs. The important and distinguishable features are ranked using the Inf-FS algorithm. Moreover, an excellent bearing multi-defect diagnosis approach is proposed by combining PSO with SVM classifier. The experiments verify the usefulness of the proposed method. We expect that our findings will be applied in gear fault detection in future research.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Nos. 51775456 and 51375405), the Open Research Subject of Key Laboratory of Automotive Measurement, Control, and Safety, Xihua University (szjj2018-132), the National Key Research and Development Program (2016YFB1200401), the Science and Technology Research and Development Program of China Railway Corporation (2016J007B), and the Zhejiang Provincial Natural Science Foundation of China (No. LQ18F010007).

References

[1] N. Upadhyay and P. K. Kankar, Diagnosis of bearing defects using tunable Q-wavelet transform, *Journal of Mechanical Science and Technology*, 32 (2) (2018) 549-558.

[2] H. Zhang, S. Lu, Q. He and F. Kong, Multi-bearing defect detection with trackside acoustic signal based on a pseudo time-frequency analysis and Dopplerlet filter, *Mechanical Systems and Signal Processing,* 70-71 (2016) 176-200.

[3] G. He, K. Ding and H. Lin, Fault feature extraction of rolling element bearings using sparse representation, *Journal of Sound and Vibration*, 366 (2016) 514-527.

[4] C. Pachaud, R. Salvetas and C. Fray, Crest factor and Kurtosis contributions to identify defects inducing periodical impulsive forces, *Mechanical Systems and Signal Processing,* 11 (6) (1997) 903-916.

[5] P. D. Mcfadden and J. D. Smith, Model for the vibration produced by a single point defect in a rolling element bearing, *Journal of Sound and Vibration*, 96 (1) (1984) 69-82.

[6] H. Talhaoui, A. Menacer, A. Kessal and A. Tarek, Experimental diagnosis of broken rotor bars fault in induction machine based on Hilbert and discrete wavelet transforms, *The International Journal of Advanced Manufacturing Technology*, 95 (1-4) (2018)1399-1408.

[7] Y. Fu, L. Jia, Y. Qin, J. Yang and D. Fu, Fast EEMD based AM-correntropy matrix and its application on roller bearing fault diagnosis, *Entropy*, 18 (2016) 242.

[8] N. Saravanan and K. I. Ramachandran, Incipient gear box fault diagnosis using discrete wavelet transform (DWT) for feature extraction and classification using artificial neural network (ANN), *Expert Systems with Applications*, 37 (6) (2010) 4168-4181.

[9] A. Rai and S. H. Upadhyay, A review on signal processing techniques utilized in the fault diagnosis of rolling element bearings, *Tribology International*, 96 (2016) 289-306.

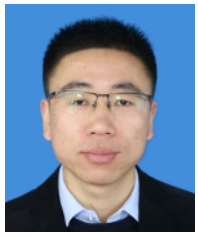
[10] W. Aziz and M. Arif, Multiscale permutation entropy of physiological time series, *Proceedings of the 9th International Multitopic Conference (INMIC '05)*, December (2005) 1-6.

[11] S. D. Wu, P. H. Wu, C. W. Wu, J. J. Ding and C. C. Wang, Bearing fault diagnosis based on multiscale permutation entropy and support vector machine, *Entropy*, 14 (8) (2012) 1343-1356.

[12] Y. Li et al., A rolling bearing fault diagnosis strategy based on improved multiscale permutation entropy and least squares SVM, *Journal of Mechanical Science and Technology*, 31 (6) (2017) 2711-2722.

[13] A. Ibanez-Molina, S. Iglesias-Parro, M. F. Soriano and J. I. Aznarte, Multiscale Lempel-Ziv complexity for EEG meas-

ures, *Clinical Neurophysiology,* 126 (3) (2015) 541-548.
[14] J. S. Richman and J. R. Moorman, Physiological time-series analysis using approximate entropy and sample entropy, *American Journal of Physiology-Heart and Circulatory Physiology*, 278 (6) (2000) H2039-H2049.
[15] Y. Pan, Y. Wang, S. Liang and K. Lee, Fast computation of sample entropy and approximate entropy in biomedicine, *Computer Methods and Programs in Biomedicine*, 104 (3) (2011) 382-396.
[16] M. Costa, A. L. Goldberger and C. K. Peng, Multiscale entropy analysis of complex physiologic time series, *Physical Review Letters*, 89 (6) (2002) 4, Article ID 068102.
[17] M. Costa, A. L. Goldberger and C. K. Peng, Multiscale entropy analysis of biological signals, *Physical Review E*, 71 (2) (2005) 18, Article ID 021906.
[18] 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, *Physics Letters A*, 378 (20) (2014) 1369-1374.
[19] J. Escudero, E. Acar, A. Fernandez and R. Bro, Multiscale entropy analysis of resting-state magnetoencephalogram with tensor factorisations in Alzheimer's disease, *Brain Research Bulletin*, 119 (2015) 136-144.
[20] W. Chen et al., Characterization of surface EMG signal based on fuzzy entropy, *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 15 (2) (2007) 266-272.
[21] 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 and Machine Theory*, 98 (2016) 114-132.
[22] 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 and Machine Theory*, 78 (2014) 187-200.
[23] G. Roffo, S. Melzi and M. Cristani, Infinite feature selection, *2015 IEEE International Conference on Computer Vision (ICCV)* (2015) 4202-4210.
[24] S. Obertino, G. Roffo, C. Granziera and G. Menegaz, Infinite feature selection on shore-based biomarkers reveals connectivity modulation after stroke, *2016 International Workshop on Pattern Recognition in Neuroimaging (PRNI)* (2016) 1-4.
[25] R. Jegadeeshwaran and V. Sugumaran, Fault diagnosis of automobile hydraulic brake system using statistical features and support vector machines, *Mechanical Systems and Signal Processing*, 52-53 (2015) 436-446.
[26] Y. Li, W. Zhang, Q. Xiong, T. Lu and G. Mei, A novel fault diagnosis model for bearing of railway vehicles using vibration signals based on symmetric alpha-stable distribution feature extraction, *Shock and Vibration*, 12 (2016) 13, Article ID 5714195.
[27] J. Kennedy and R. Eberhart, Particle swarm optimization, *Proceedings IEEE International Conference Neural Networks,* 4 (1995) 1942-1948.
[28] K. A. Loparo, *Bearing data center*, Case Western Reserve University.



Yongjian Li received his Ph.D. degree in Mechanical Engineering from the Southwest Jiaotong University, Chengdu, China in 2017. He is currently a lecturer at the School of Railway Tracks and Transportation, Wuyi University, China. His research interests include signal processing and data mining for machine health monitoring and fault diagnosis.

Bingrong Miao received his Ph.D. in Vehicle Engineering from Southwest Jiaotong University in 2007. He is currently working as an Associate Professor in the State Key Laboratory of Traction Power at Southwest Jiaotong University. His research interests include vehicle dynamics, structural fatigue strength, load identification, and damage identification. He has officially published three books and more than 70 articles.