# Refined Composite Multivariate Multiscale Dispersion Entropy and Its Application to Fault Diagnosis of Rolling Bearing

**CONGZHI LI, JINDE ZHENG[ID], HAIYANG PAN[ID], JINYU TONG, AND YIFANG ZHANG**

School of Mechanical Engineering, Anhui University of Technology, Ma'anshan 243032, China

Corresponding author: Jinde Zheng (lqdlzheng@126.com)

**ABSTRACT** Many nonlinear dynamic and statistic methods, including multiscale sample entropy (MSE) and multiscale fuzzy entropy (MFE), have been widely studied and employed to fault diagnosis of the rolling bearing. Multiscale dispersion entropy (MDE) is a powerful tool for complexity measure of time series, and compared with MSE and MFE, it gets much better performance and costs less time for computation. Since single-channel time series analysis will cause information missing, inspired by multivariate multiscale sample entropy (MMSE) and multivariate multiscale fuzzy entropy (MMFE), refined composite multivariate multiscale dispersion entropy (RCMMDE) was proposed in this paper. After that, RCMMDE was compared with MDE, MMSE, and MMFE by analyzing synthetic signals and the results show that the RCMMDE has certain advantages in terms of robustness. A hybrid fault diagnostics approach is proposed for rolling bearing with a combination of RCMMDE, multi-cluster feature selection, and support vector machine. Also, the proposed method is compared with MDE, MMSE, and MMFE, as well as multivariate multiscale dispersion entropy-based fault diagnosis methods by analyzing the experimental data of rolling bearing, and the result shows that the proposed method gets a higher identification rate than the existing other fault diagnosis methods.

**INDEX TERMS** Multiscale entropy, multiscale fuzzy entropy, multivariate multiscale dispersion entropy, refined composite multivariate multiscale dispersion entropy, rolling bearing, fault diagnosis.

## I. INTRODUCTION

Rolling bearing is one of the most common part of rotating machines, which is usually used in industrial and mechanical applications. It is important to study the fault diagnosis technologies of rolling bearings [1]–[3]. In general, the vibration signals of rolling bearing will present strong non-stationary and nonlinear characteristics when some localized failures occur on rolling bearings [4], [5]. Traditional linear and stationary signal analysis methods inevitably have some limitations in analysis of non-stationary and nonlinear signals. In recent decades, the entropy-based parameters that aim at evaluating complexity of raw signals have been extensively applied in fault diagnostics of rotary machines [6]–[9], such as multiscale sample entropy (MSE) [10]–[12],

The associate editor coordinating the review of this manuscript and approving it for publication was Zhixiong Peter Li.

multiscale permutation entropy (MPE) [13] and multiscale fuzzy entropy (MFE) [14]–[16]. MSE was utilized study [17] to extract failure related features from raw signals of rolling bearing. In [16], Zheng et al applied MFE to reveal faulty characteristics of rolling bearing. MPE was applied to fault feature representation of rotary machines and was compared with MSE through analyzing experiment data of rolling bearing. However, the existing nonlinear dynamic tools for complexity or irregularity measure still have some intrinsic limitations when used to real world data analysis. For example, sample entropy (SampEn) is computationally inefficient for long data and has a large variation in similarity measure. In MPE the relationship of adjacent amplitudes is unconsidered. Recently, dispersion entropy (DisEn) was proposed by Rostaghi and Azami [18] to overcome the defects of SampEn and PE. The superiority of DisEn to SampEn and PE is verified by analyzing simulation and biological signals.

Later, DisEn was extended to multiscale dispersion entropy (MDE) [19] and then is enhanced by refined composite multiscale dispersion entropy (RCMDE) in [20].

However, at present, lots of fault diagnosis methods of rotary machines are generally founded based on single channel or one direction vibration signals, while the other channels or directions of vibration information are often ignored. Take gearbox vibration as an example, the common used fault diagnosis methods by many scholars are generally based on vibration signals acquired from vertical body direction of gearbox. Although the information of vibration signals in single direction or channel usually can effectively identifying the diagnostics of local fault, due to the complexity of transmission paths of gearbox, the signals collected from different directions also contain important fault information. Therefore, it is necessary to develop multi-channel data based fault diagnosis method to improve diagnosis effectiveness. The assessment of dynamic relationship between different channels from a multi-channel synchronous data has become predominant in field [21]–[23]. Therefore, the fault information hidden in multi-channel vibration signals can provide much more information related with fault, as well a chance for getting a much higher identification rate of fault modes.

Based on multidimensional embedding reconstruction theory, MSE and MFE were extended to multivariate framework, i.e. multivariate MSE (MMSE) [10] and multivariate MFE (MMFE) [24] were developed to measure complexity of time series in multi-channel data taking into account the mutual predictability. In MMSE and MMFE, the dynamic relationship of multi-channel data was investigated in considering long range correlation, complexity as well as mutual correlation prediction [25], [26]. Also nonlinear internal coupling characteristic of multi-channel data was presented. Since the benefits of multi-channel signal analysis had been recognized, MDE was further upgraded to multivariate MDE (MMDE) [27] and was compared with MMSE and MMFE. The superiority of MMDE was proved by analyzing biomedical times series. In this paper, the refined composite multiscale dispersion entropy (RCMMDE) is proposed to improve the performance of MMDE and then is employed to extract nonlinear fault feature of rolling bearing.

After that, since the extracted RCMMDE in different scales often contain redundant features and it is unreasonable to take all RCMMDE values as inputted features of classifier. It is necessary to refine the fault features into a low dimension space. Multi-cluster feature selection (MCFS) recently proposed in [28] is a dimension reduction tool, which maintains data clustering structure and realize an unsupervised or supervised dimensionality reduction of data. MCFS was utilized to reorder the extracted fault features from vibration signals of rolling bearing according to their significance. Several most sensitive extracted fault features were shortlisted to train and test multi-class classifier in order to perform intelligent fault diagnosis of rolling bearing [29]. Finally, based on RCMMDE, MCFS and support vector machine (SVM) [30], a new fault diagnosis method for rolling bearing is proposed

in this paper. The effectiveness of the proposed fault diagnosis method is verified by comparing with single-channel MDE, multi-channel MMSE, MMFE and MMDE based fault diagnosis ones through experimental data analysis.

The rest of this paper is structured as following. In section 2 multivariate dispersion entropy (MvDE) and MMDE are briefly reviewed. RCMMDE was proposed in section 3 and its comparison with MDE was given by analyzing simulation signals. The proposed fault diagnosis method of rolling bearing and its application analysis in experimental data are given in section 4. Conclusions are discussed in section 5.

## II. MULTIVARIATE MULTISCALE DISPERSION ENTROPY
### A. MULTIVARIATE DISPERSION ENTROPY
To quantify complexity of multivariate time series, Based on multivariate embedding theory, dispersion entropy (DisEn) was extended to multivariate dispersion entropy (MvDE) and the detailed steps are shown as follows.

(1) A multivariate time series $X = \{x_{k.i}\}_{k=1,2,\ldots n}^{i=1,2,\ldots N}$ is mapped to $Y = \{y_{k.i}\}_{k=1,2,\ldots n}^{i=1,2,\ldots N}$ by

$$y_{k,i} = \frac{1}{\sigma_k \sqrt{2\pi}} \int_{-\infty}^{x_{k,i}} e^{\frac{-(t-\mu_k)^2}{2\sigma_k^2}} \, \mathrm{d}t \qquad (1)$$

where $\mu$ is expectation and $\sigma^2$ is variance.

(2) Y is mapped to $Z = \{z_{k.i}\}_{k=1,2,\ldots n}^{i=1,2,\ldots N}$ (from 1 to c) by a linear transform

$$z_{k,i} = \mathrm{R}(c \cdot y_{k,i} + 0.5) \qquad (2)$$

where $c$ represents the class and $\boldsymbol{R}$ is the rounding function.

(3)According to multivariate embedding theory, the time series $Z$ is reconstructed as follows.

$$\begin{aligned} Z_m(j) = [&z_{1,j}, z_{1,j+d_1}, \ldots, z_{1,j+(m_1-1)d_1}, \\ &z_{2,j}, z_{2,j+d_2}, \ldots, z_{2,j+(m_2-1)d_2}, \ldots, \\ &z_{n,j}, z_{n,j+d_n}, \ldots, z_{p,j+(m_n-1)d_n}] \end{aligned} \qquad (3)$$

where $j \in [1, N-(m-1)d]$, $m = [m_1, m_2, \ldots, m_n]$ represents embedding dimension and $\lambda = [\lambda_1, \lambda_2, \ldots, \lambda_n]$ represents time delay.

(4) For every $Z_m(j)$, all combinations of $m$ elements in $Z_m(j)$, termed $\phi_{q,}l(j)$ ($q \in [1, C_m^{mn}]$, $l \in [1, m]$ ), are created, where the $C_m^{mn}$ is the number of all combinations of $mn$ numbers with length $m$.

(5) Each $\phi_q l(j)$ is mapped to a dispersion pattern $\pi_{v_0 v_1 \ldots v_{m-1}}$ ($v = 1, 2, \ldots, c$) where $\phi_{q,1}(j) = v_0$, $\phi_{q,2}(j) = v_1$, $\ldots$, $\phi_{q,}l(j) = v_{m-1}$. Since $\pi_{v_0 v_1 \ldots v_{m-1}}$ consists of $m$ digits and each $m$ has $c$ classes, there are totally $c^m$ dispersion patterns. The total number of combinations of each $Z_m(j)$ is $C_m^{mn}$. Therefore, there are $[N - (m-1)d] C_m^{mn}$ dispersion patterns for all $n$ channel data.

(6) The probability of each dispersion pattern can be computed by

$$p(\pi_{v_0 v_1 \ldots v_{m-1}}) = \frac{Number(\pi_{v_0 v_1 \ldots v_{m-1}})}{(N - (m-1)d)C_m^{mn}} \qquad (4)$$

where *Number* in Eq. (4) denotes the number of $\pi_{v_0 v_1 \ldots v_{m-1}}$ in $\phi_{q,} l(j)$.

(7) Finally, according to the definition of Shannon entropy, MvDE of multivariate data $X$ is calculated by

$$MvDE(x, m, c, d) = -\sum_{\pi=1}^{c^m} p(\pi_{v_0 v_1 \ldots v_{m-1}}) \ln p(\pi_{v_0 v_1 \ldots v_{m-1}}) \tag{5}$$

In [27] three algorithms of MvDE were introduced. However, these algorithms were not used because of their low computational efficiency and the need for a large amount of storage space. The relationship of different channel data is considered in MvDE and it is more reliable and accurate than single-channel DisEn. However, MvDE was limited to the analysis of single-scale data, MvDE in multiscale framework was introduced to overcome this challenge [27].

### B. MULTIVARIATE MULTISCALE DISPERSION ENTROPY

The steps of MMDE is firstly to implement the coarse-grained time series of original multivariate data, then MvDE of each multivariate coarse-grained time series is computed. The steps of MMDE are given as follows.

(1) For $n$ channel data U= $\{u_{k,b}\}$ with length $L$, the coarse-grained multivariate time series for scale factor $\tau$ is calculated by

$$x_{k,i}^{\tau} = \frac{1}{\tau} \sum_{b=(i-1)\tau+1}^{i\tau} u_{k,b}, \; 1 \le i \le L/\tau, \; 1 \le k \le n \tag{6}$$

(2) MvDE of each coarse-grained multivariate time series $\{x_{k,i}^{\tau}\}$ is computed under the same parameters.

MMDE was obtained by expending MvDE from signal-scale to multi-scale and get more information from multivariate coarse-grained time series in different scales. However, in the above coarse graining multivariate time series at scale factor $\tau$ of MMDE, only information of coarse graining multivariate time series starting with $u_{k,1}$ is considered and the remaining $\tau - 1$ multivariate time series are missing. MMDE does not consider the relationship between coarse-grained time series and causes the lack of statistical information.

## III. REFINED COMPOSITE MULTIVARIATE MULTISCALE DISPERSION ENTROPY

### A. RCMMDE ALGORITHM

The detail steps of RCMMDE are given as follows.

(1) For $n$ channel multivariate time series U= $\{u_{k,b}\}$ with length $L(k = 1, 2, \ldots, n, b = 1, 2, \ldots, L)$, the coarse-grained multivariate time series are calculated for a given scale vector $\tau$ and elements of the $a-$th coarse-grained time series $X_a^{\tau} = \{x_{k,i,1}^{\tau}, x_{k,i,2}^{\tau}, \cdots\}$ is calculated by

$$x_{k,i,a}^{\tau} = \frac{1}{\tau} \sum_{b=a+\tau(i-1)}^{a+i\tau-1} u_{k,b} \tag{7}$$

where $1 \le i \le L/\tau$, $1 \le k \le n$, $1 \le a \le \tau$.

(2) The RCMMDE of original multivariate time series U can be calculated by

$$RCMMDE(U, m, c, d, \tau)$$
$$= -\sum_{\pi=1}^{c^m} \bar{p}(\pi_{v_0 v_1 \ldots v_{m-1}}) \ln \bar{p}(\pi_{v_0 v_1 \ldots v_{m-1}}) \tag{8}$$

where $\bar{p}(\pi_{v_0 v_1 \ldots v_{m-1}}) = \frac{1}{\tau} \sum_{1}^{\tau} p_a^{\tau}$ represents the mean frequency of $\pi_{v_0 v_1 \ldots v_{m-1}}$ in all $X_a^{\tau}$ ($p_a^{\tau}$ is the frequency of dispersion pattern $\pi_{v_0 v_1 \ldots v_{m-1}}$ in $a-$th multivariate coarse-grained time series $X_a^{\tau}$).

In RCMMDE algorithm, the parameters including class $c$, embedding dimension $m$ and time delay $d$ need to be preset. It was recommended in [18] to set $m = 2$ or $3$ and $c \in [4]$, [8]. If $c$ is too small, two different amplitudes may be assigned to the same class and if $c$ is too large, a very small error will result in the change of its class. Besides, since too small $m$ will cause RCMMDE insensitive to the dynamic changes of the original signal while too large $m$ will make RCMMDE unable to detect small changes, $m$ usually takes 2 or 3. Last, time delay $\lambda$ is set as 1, as when $\lambda$ is larger than 1, some important information may be lost. Therefore, $c = 6$, $m = 3$ and $\lambda = 1$ were set as recommended by [18].

### B. COMPARISON ANALYSIS OF SYNTHETIC SIGNALS

In this section, the performance of RCMMDE is investigated by analyzing synthetic signals. Also it is compared with other multivariate complex measure methods to verify its effectiveness and superiority. White Gaussian Noise (WGN) and 1/f noise are two kinds of random signals that are widely used in complexity analysis. It has been validated that generally the structure of 1/f noise is more complex than that of WGN, which leads to that WGN gets a large entropy at lower scale and the entropy values will decrease with the increase of scale factor, while the entropy values of 1/f noise will be stable at a constant value and get larger values than that of WGN at most scale factors. Correspondingly, for the multivariate data, if we use WGN and 1/f noise to construct multivariate synthetic signal, the more channels occupied by 1/f noise, the more complex the multivariate signal is. Without loss of generality, WGN and 1/f noise are used to construct different kinds of multivariate synthesis signals with three channels, they are: a) three channel WGN signals, b) two channel WGN signals and 1 channel 1/f noise, c) 1 channel WGN and 2 channel 1/f noises and d) three channel 1/f noises. We take each kind of multivariate synthesis signals 30 samples with length 2048 points.

Next, for comparison purpose, MMDE, MMFE and RCMMDE are computed of the four kind multivariate synthesis signals. The mean standard deviation diagram of MMDE, MMFE and RCMMD for the four kinds of signals are shown in Fig. 1. It can be seen from Fig. 1 that the overall trends of MMDE, MMFE and RCMMDE curves are very similar and the overall relationships of MMDE, MMFE and RCMMDE of four kinds of multivariate signals from second scale to end are as follows: d) > c) > b) > a). Generally, the multivariate
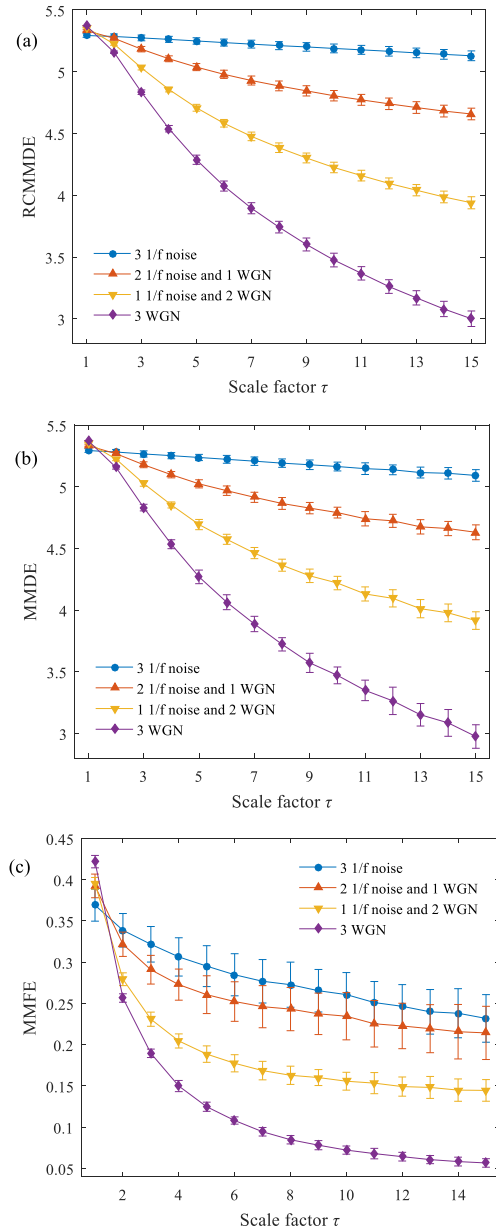
**FIGURE 1.** RCMMDE, MMDE and MMFE of multivariate synthetic signals. (a) RCMMDE; (b) MMDE and (c) MMFE.

signal consisting of three 1/f noises is the most complex one, and the following is the one consists of one channel WGN and 2 channel 1/f noises, then the one with two channel WGN signals and 1 channel 1/f noise, last, the most simple one with three channel WGN signals. The results of MMDE and RCMMDE are very consistent with above fact; however, the standard deviations of RCMMDE of four kinds of multivariate signals are obviously smaller than that of MMDE, which indicates RCMMDE is much more stable and robust than MMDE. The overall trends of MMFE for the four kinds of multivariate signals are same as that of RCMMDE and MMDE, however, the third class multivariate signals (one WGN and two 1/f noises and the forth class (three 1/f noises) are not clearly separated by MMFE and have

overlap at most scales. Therefore, compared with MMDE and MMFE, RCMMDE shows a better separating capacity and higher stability and has more advantages in multivariate signal feature representation than MMDE and MMFE.

## IV. RCMMDE BASED FAULT DIAGNOSIS FOR ROLLING BEARING

### A. MULTI-CLUSTER FEATURE SELECTION

Based on the advantages of RCMMDE, in this paper, it is employed to represent nonlinear complexity information related with fault feature of vibration signals of rolling bearing. However, RCMMDEs at all scale factors may contain redundant information, which will affect fault classification and diagnostic efficiency. It is necessary to refine or select the features related with fault from the obtained RCMMDEs to fulfill a high fault effect. MCFS is a recently proposed dimensionality reduction tool, which keeps the multi-cluster structure of data well while choosing the feature and can realize an unsupervised or supervised learning by using spectral analysis technique. In [28], MCFS is compared with other dimensionality reduction methods, such as Q-a algorithm, maximum variance and Laplacian score and the results show that MCFS has achieved remarkable results in clustering and classification and the performance of MCFS is particularly good when the number of features is less than 50. Therefore, in this paper, MCFS is utilized to refine the extracted features of vibration signals of rolling bearings.

### B. THE PROPOSED FAULT DIAGNOSIS METHOD

Based on the advantages of RCMMDE and MCFS, the proposed fault diagnosis method for rolling bearing is given as follows.

(1) Suppose there are $K$ classes multi-channel vibration data of rolling bearing and each class have $N_k$ samples, i.e. $\{X_{k,n}, n = 1, 2, \ldots, N_k, k = 1, 2, \ldots, K\}$, where $\{X_{k,n}\}$ is the $n-$th $p$-channel vibration data of the $k-$th class. Totally, we have $N = \sum_{k=1}^{K} N_k$ samples and if let $N_1 = N_2 = \ldots = N_K$ and thus $N = KN_1$.

(2) RCMMDE of all $N$ multi-channel vibration data $\{X_{k,n}\}$ is calculated with selected parameters and thus each multi-channel vibration data $\{X_{k,n}\}$ are mapped into the initial 3-D feature sets $\{RCMMDE_{k,n}(\tau)\}$ with dimension $K \times N_1 \times \tau_{\max}$ where $\tau = 1, 2, \ldots, \tau_{\max}$, $\tau_{\max}$ is the preset largest scale factor, $k = 1, 2, \ldots, K$, $n = 1, 2, \ldots, N_k$.

(3) For $N_k$ samples of $K$ classes, $M_k$ samples of each class are randomly selected as training samples while the remaining $(N_k - M_k)$ are seen as testing samples. i.e. the initial 3-D feature sets $\{RCMMDE_{k,n}(\tau)\}$ is divided into training samples $\{T_{k,m_1}(\tau)\}$ and testing samples $\{Q_{k,l_1}(\tau)\}$, where $m_1 = 1, 2, \ldots, M_k$, $l_1 = 1, 2, \ldots, N_k - M_k$, $k = 1, 2, \ldots, K$.

(4) The training samples $\{T_{k,m_1}(\tau)\}$ are reduced into $d$ number of features by using MCFS, i.e. the number of feature elements $\tau_{\max}$ is reduced to $d$: $\{T_{k,m_1}(\tau)\} \rightarrow \{T_{k,m_1}(\tau')\}$, $\tau' = 1, 2, \ldots, d$, $d < \tau_{\max}$. The testing samples $\{Q_{k,l_1}(\tau)\}$ are reduced into $d$ number of features in the $\tau$ direction.

The selection of testing sample is consistent with that of the training sample, i.e. the selected features are the same as the training samples in the same order. i.e. $\{Q_{k,l_1}(\tau)\} \rightarrow \{Q_{k,l_1}(\tau')\}$. The $\{T_{k,m_1}(\tau')\}$ is the new training data sets and $\{Q_{k,l_1}(\tau')\}$ is the new testing data sets.

(5) The new training data sets $\{T_{k,m_1}(\tau')\}$ are input to SVM-based multi-classifier for constructing the diagnosis model. Then new testing data sets $\{Q_{k,l_1}(\tau')\}$ are used to test the trained diagnosis model and the outputs are used to fulfill the diagnosis of fault types and degrees.



**FIGURE 2.** Flowchart of the proposed method.

Fig. 2 illustrates the flowchart of the proposed method.
*CASE 1:*

The effectiveness of the proposed method is verified by experiment data of rolling bearing of Case Western Reserve University (CWRU) bearing data center [31]. The test rig shown in Fig. 3 comprises a 2 hp motor, a torque transducer, a dynamometer and control electronics. The sampling frequency is 12 kHz, motor speed is 1730 r/min and motor load 3 hp. The fault diameters of ball element, inner and outer races of rolling bearing with are 0.1778 mm (labeled as BE1, IR1 and OR1) and 0.5334mm (labeled as BE2, IR2 and OR2) where the outer race fault is located at 6 o'clock. Thus 6 classes vibration signals of rolling bearing with different types and degrees were used as well as the normal bearing (label as Norm). For each class, the synchronous vibration signal of Fan and drive end are used as two-channel data and there are 29 samples with 4096 sampling points are
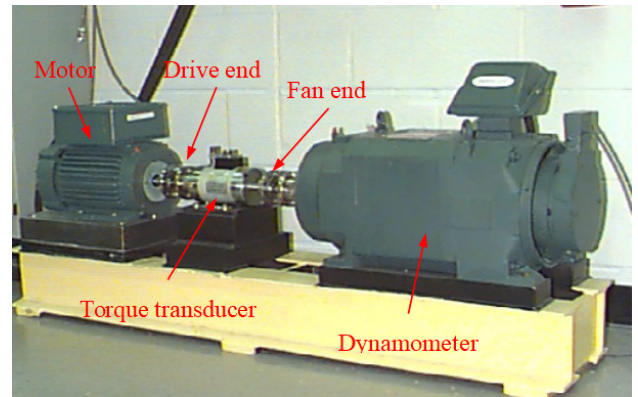


**FIGURE 3.** The test stand of rolling bearing of CWRU.

**TABLE 1.** Label description of experimental data of rolling bearing.

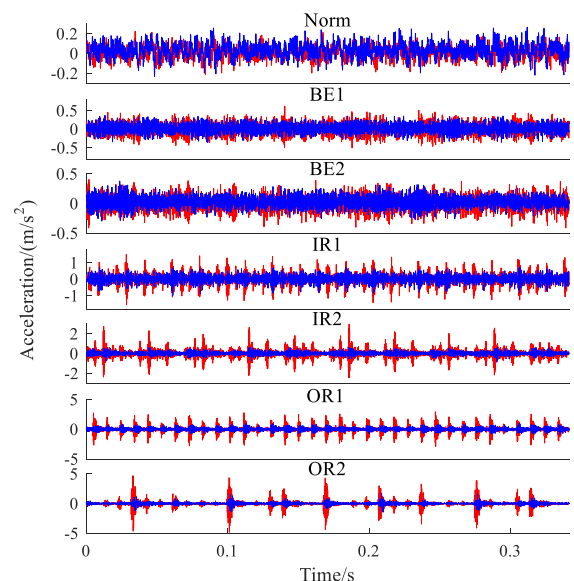| Class label | State | Fault size | No. of training data | No. of testing data |
|---|---|---|---|---|
| 1 | BE1 | 0.1778mm | 15 | 14 |
| 2 | BE2 | 0.5334mm | 15 | 14 |
| 3 | IR1 | 0.1778mm | 15 | 14 |
| 4 | IR2 | 0.5334mm | 15 | 14 |
| 5 | OR1 | 0.1778mm | 15 | 14 |
| 6 | OR2 | 0.5334mm | 15 | 14 |
| 7 | Norm | 0 | 15 | 14 |



**FIGURE 4.** Waveform of different classes of rolling bearing of CWRU's data; where red represents data of drive end and blue represents fan end.

used and totally 203 samples are under our consideration. A detail description of label information of experimental data is shown in Table 1. The waveform of vibration signal of rolling bearing of each class are shown in Fig. 4.

The proposed fault diagnosis approach is examined by the above experimental data of rolling bearing. RCMMDEs in the first 15 scales (i.e. $\tau_{\max} = 15$) of 203 samples of all classes are computed and taken as initial fault features. The means
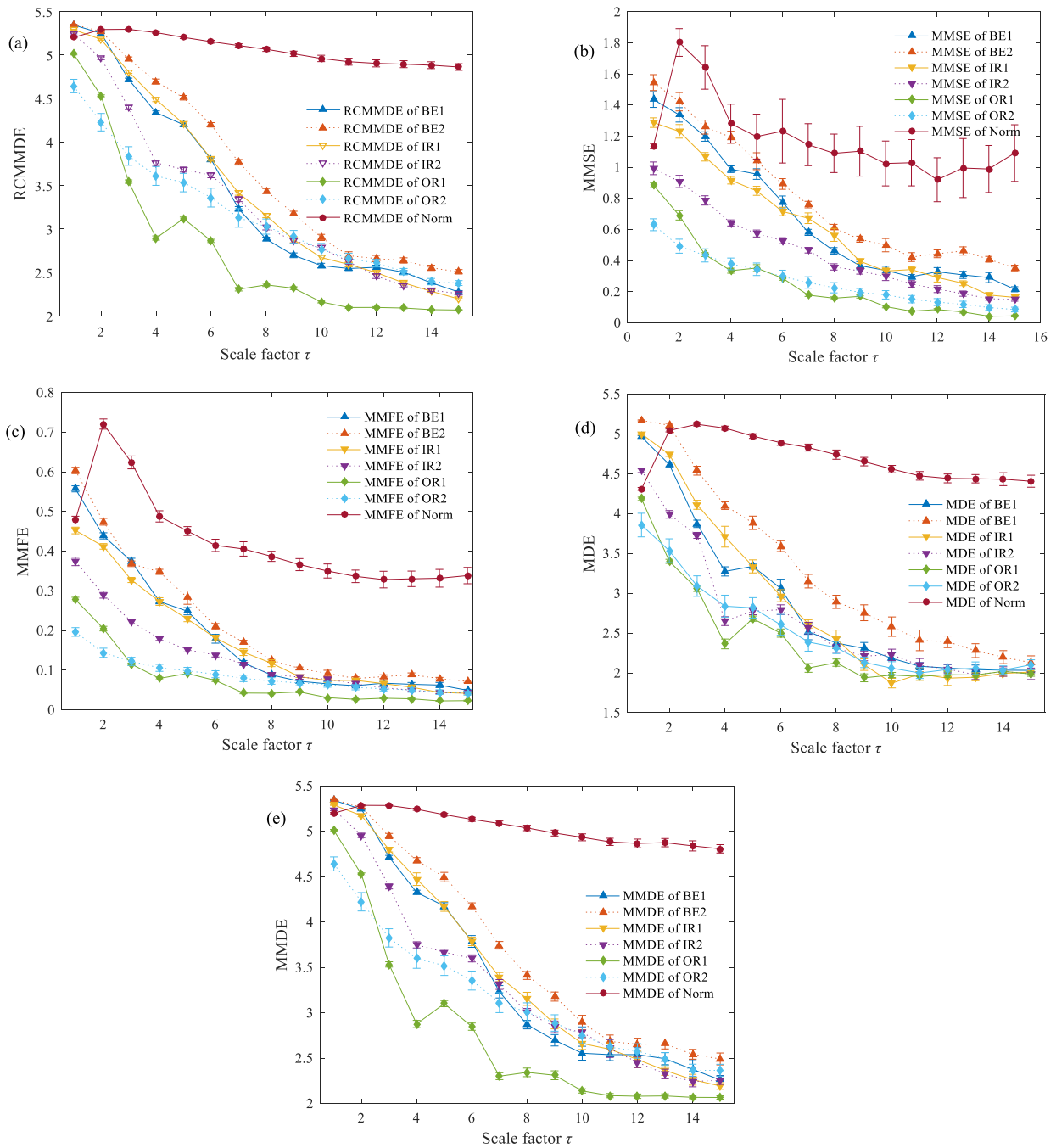
**FIGURE 5.** The entropy results of rolling bearing data analyzed by using five methods. (a) RMMDE; (b) MMSE; (c) MMFE; (d) MDE and (e) MMDE.

and standard deviations of RCMMDE for each class are demonstrated in Fig. 5(a). From Fig. 5(a), it can be seen that the RCMMDE extracted from vibration signal of normal bearing varies slowly and nearly stabilize at 5 with increase of scale factor, while that of rolling bearing with local fault generally decrease monotonically. Besides, the DisEn values of vibration signal of normal bearing at most scales (larger than 2) are larger than that of vibration signals of faulty bearing and this can be explained that generally, the vibration signal of healthy bearing is irregular and similar to 1/f

noise. And the vibration signals will present obvious periodic impulse components that modulates the natural frequency ones, once local faults appear in the rolling bearing. Despite this fact, it is noted that bearing vibrations with different faults show periodic impulses at different frequencies. They will have different DisEns at different scales and thus can be distinguished by RCMMDE. For comparison purpose, MMSE, MMFE and MMDE of the above multi-channel experiment signals also are computed, as well as MDE of single-channel signal from the Drive end and the corresponding results are

given in Fig. 5(b-e). First, by observing the Fig. 5(b, c), it can be found that compared with MMSE and MMFE of normal and faulty rolling bearings, RCMMDE is more stable and its error is much smaller than other methods. Second, the MMSE curves of OR1 and OR2 are very similar at most scales and it is difficult to differ from each other. The fault characteristics curves extracted by MMSE and MMFE are much closer and both differ greatly from RCMMDE. Though different fault classes can be distinguished by MDE, it is different to differ IR1 and IR2, OR1 and OR2. Also, it is hard for MMDE to distinguish IR1 and BE1. By observing the Fig. 5 (a, d), it can be seen that the separation degree of MDE curve is smaller than that of RCMMDE. By observing Fig. 5(a, d, e), it can be seen that the trends of MDE, MMDE and RCMMDE are nearly the same, however, MDE has larger standard deviation than MMDE, when RCMMDE gets the smallest one. This indicates that multichannel signal analysis is more stable than traditional single-channel data analysis one. Therefore, the above analysis indicates that RCMMDE is superior to MMSE, MMDE and MDE in feature extraction and show much stronger distinguish effect for different fault classes and degrees.
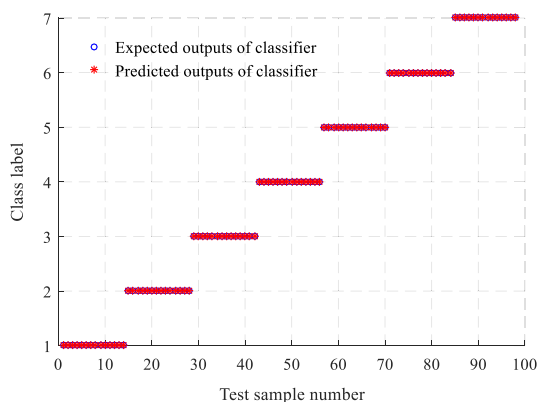


**FIGURE 6.** Outputs of the proposed method.

Next, RCMMDEs of seven classes of rolling bearing are divided into training samples and test samples. in each class, 15 samples are randomly selected as training samples and the remaining 14 are seen as testing ones. Then MCFS is used to select the fault features by using training data. Namely, RCMMDEs of 105 training samples with 15 scales are input the MCFS for feature selection. The first 4 elements of training fault features are used to instruct the new fault features of the 98 testing samples i.e. the selected scales (scale at 6, 8, 7 and 11) of testing samples' features are same as that of training samples. The training labels were created, i.e. BE1, BE2, IR1, IR2, OR1, OR2 and Norm are labeled as 1, 2, 3, 4, 5, 5, 6 and 7 respectively. Finally, the training data and training labels are used to train the SVM based multi-classifier, which we use is Lin's LIBSVM tool and the kernel function is radial basis function (RBF) [29]. After that, all 98 new fault features of testing data are input to the trained classifier for testing and the predictive results are shown in Fig. 6. From Fig. 6 it can
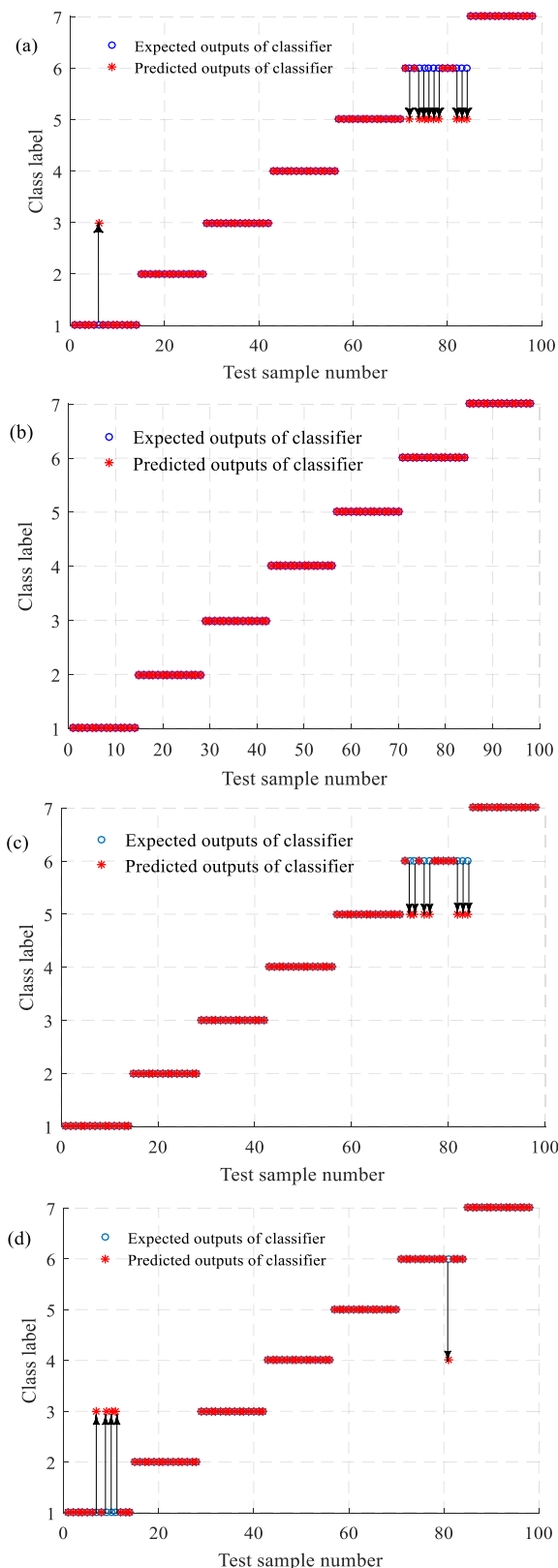


**FIGURE 7.** Outputs of MMSE, MMFE, MDE, and MMDE based fault diagnosis methods. (a) MMSE; (b) MMFE; (c) MDE and (d) MMDE.

be found that all testing samples are identified correctly and the recognition rate of fault locations and degrees is 100 %, which verified the effectiveness of the proposed method.

**TABLE 2.** Comparison of identifying rate of the proposed method with existing ones.

| Method | Misclassified sample number | Identifying rate |
|---|---|---|
| MMSE-MCFS-SVM | 10 | 89.80% |
| MMFE-MCFS-SVM | 0 | 100% |
| MDE-MCFS-SVM | 7 | 92.86% |
| MMDE-MCFS-SVM | 5 | 94.90% |
| RCMMDE-MCFS-SVM | 0 | 100% |

**TABLE 3.** Selected features with different number of RCMMDE.

| No. of feature | Selected feature scale |
|---|---|
| 2 | 15, 11 |
| 3 | 7, 8, 15 |
| 4 | 6, 8, 7, 11 |
| 5 | 6, 2, 7, 11, 15 |
| 6 | 6, 11, 2, 15, 7, 14 |
| 7 | 6, 2, 11, 7, 15, 14, 4 |
| 8 | 11, 2, 14, 7, 15, 6, 8, 13 |
| 9 | 11, 14, 2, 7, 6, 8, 15, 12, 4 |
| 10 | 11, 14, 2, 7, 6, 8, 15, 12, 4, 13 |
| 11 | 11, 14, 7, 8, 2, 15, 6, 12, 4, 1, 13 |
| 12 | 11, 14, 8, 2, 7, 15, 4, 12, 6, 1, 5, 13 |
| 13 | 11, 8, 14, 2, 7, 15, 4, 6, 1, 5, 13, 12, 9 |
| 14 | 11, 8, 14, 2, 7, 4, 1, 15, 6, 5, 13, 12, 3, 9 |
| 15 | 11, 8, 14, 2, 7, 4, 15, 1, 6, 5, 13, 3, 12, 9, 10 |

For comparison purpose, in the proposed fault diagnosis method, RCMMDE is replaced by MMSE, MMFEE, MMDE and MDE then combing MCFE and SVM. The MMSE, MMFEE, MMDE and MDE of same training and testing samples are computed. Similar to the above training and testing process, MCFS is used to sort the order of fault features and the first four ones are selected as new fault features for training and testing. The outputs of testing samples are shown in Fig. 7(a-d) and corresponding recognition rate are shown in Table 2. From Table 2 and Fig. 7, it can be found that there are 10 testing samples are misclassified, i.e. one sample of BE1 class is misclassified to IR1 and 9 sample of OR2 is misclassified to OR1 and the fault recognition rate is 89.80%. The fault recognition rate of MMFE based fault diagnosis method is 100%. For the MDE based fault diagnosis method there are 7 samples are misclassified while for MMDE based method 5 samples are misclassified and the identifying rates of MDE and MMDE based fault diagnosis methods are 92.86% and 94.90%. The above analysis indicates that the superiority and effectiveness of RCMMDE to MDE, MMDE and MMSE. Also the advantages of multi-channel analysis to traditional single-channel analysis was verified in rolling bearing fault diagnosis by this case.

In order to investigate the influence of feature number on identifying accuracy, the number of features $d$ ranging from 2 to 15 are selected by using MCFS from initial features.
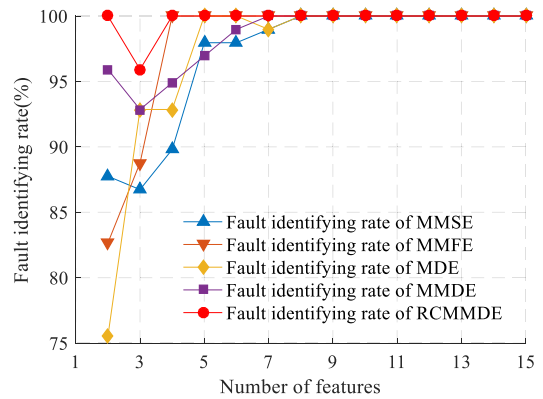


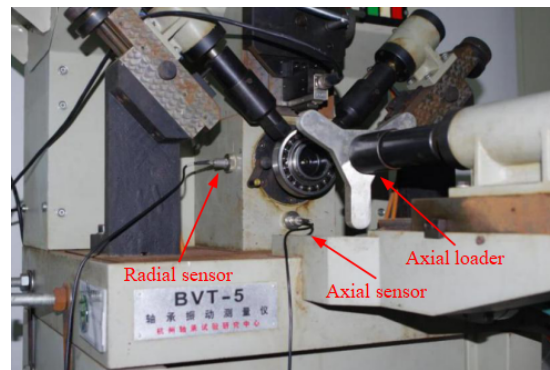**FIGURE 8.** Identifying rate comparison for different numbers of features.

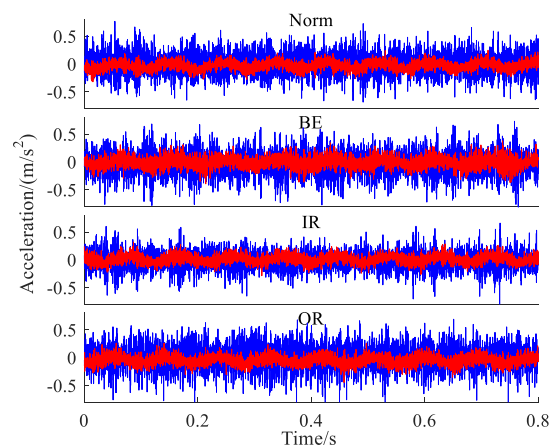

**FIGURE 9.** The rolling bearing test rig of AHUT.



**FIGURE 10.** The waveforms of experiment data; where red is radial and blue is axial signals.

Correspondingly, the selected different numbers of RCMMDE features are shown in Table 3. The fault recognition rates of RCMMDE, as well as MDE, MMFE, MMSE and MMDE based fault diagnosis method with different number of features, are given in Fig. 8. From Fig. 8 it can be found that RCMMDE and MCFS based fault diagnosis method has higher fault identifying rate than other methods when the number of features are smaller than 8. In fact, for RCMMDE method, when the number of feature is larger than 4, the fault recognition rate will reach 100%, while the other methods may need more features. Therefore, the above analysis
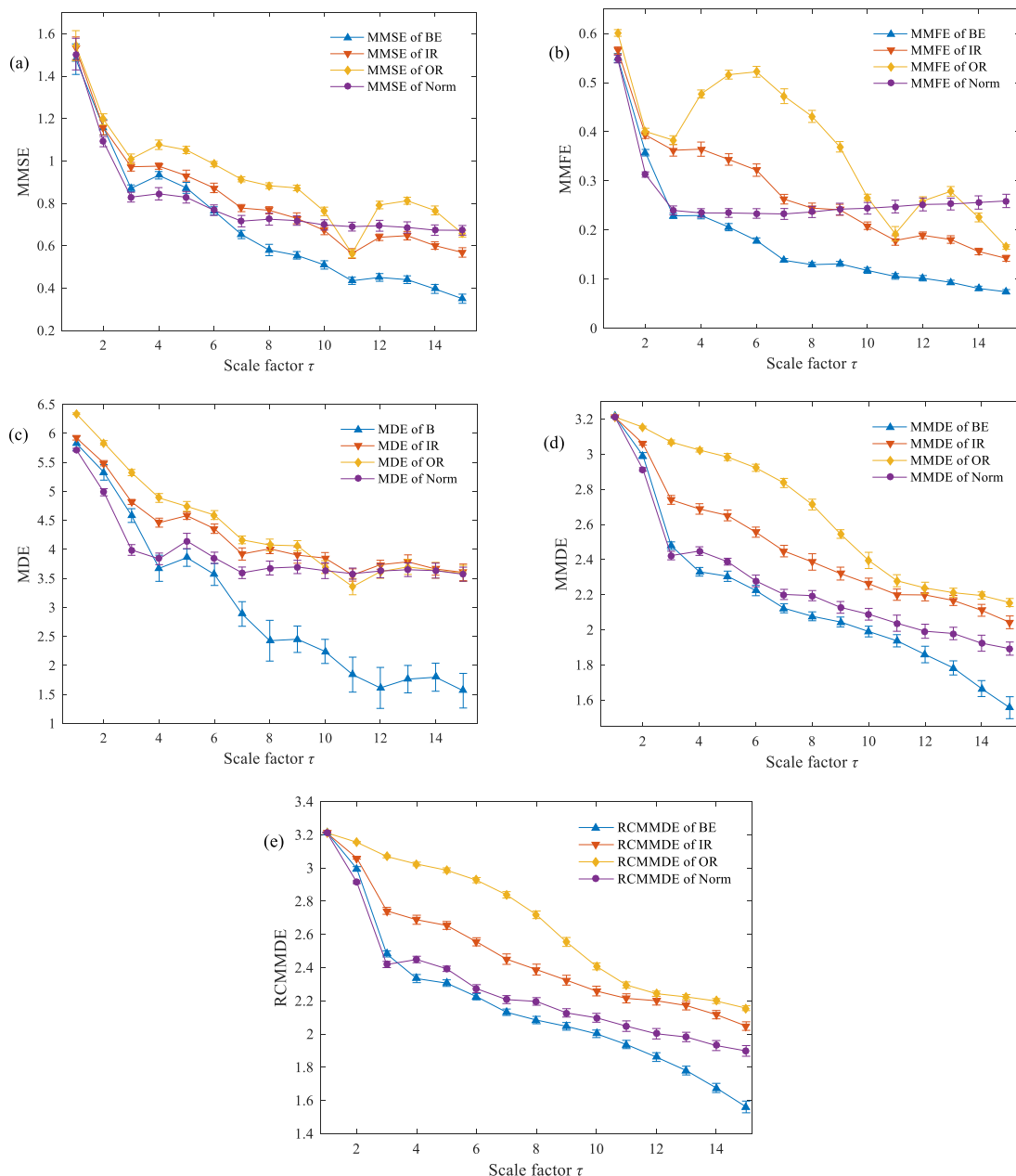
**FIGURE 11.** Results of our rolling bearing data analyzed by the five methods. (a) MMSE; (b) MMFE; (c) MDE; (d) MMDE and (e) RCMMDE.

indicates the effectiveness of RCMMDE and SVM based fault diagnosis method for rolling bearing.

*CASE 2:*

In this part, the experimental data of Anhui University of Technology (AHUT) was used to test and verify the effectiveness of proposed fault diagnosis method. The test bearing is 1210 self-aligning ball bearing, and the local faults were seeded by metal electro-engraving machine. In the test, the outer race of rolling bearing is fixed and the inner race rotates with the shaft with a speed of 1, 800 r/min and load 100 N. The experiment data are collected from normal bearing (label as Norm) and the faulty rolling bearings where the local faults located at ball element, inner race and outer race (label

as BE, IR and OR). The vibration signals of bearing at axial and radial channels are synchronously collected by sensors with sampling frequency 5120 Hz and sampling time 120 s. The test rig of AHUT is shown in Fig. 9. The waveforms of vibration signal of rolling bearing in two channels (with 4096 sampling points) are shown in Fig. 10.

For the experiment data of rolling bearing, four classes, i.e. Norm, BE, IR and OR have 29 samples with length 4096 and thus 116 samples are obtained. Then MMSE, MMFE, MMDE and RCMMDE of all 116 samples were calculated with scale factor 15, as well as MDE of vibration signal at axial. The mean and standard deviations of the extracted features of rolling bearing are shown in Fig. 11, from which it can
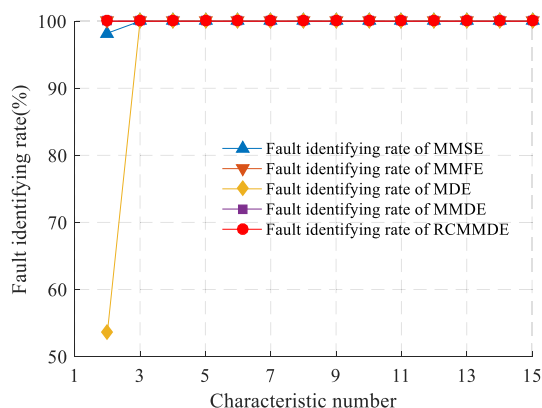
**FIGURE 12.** Identifying rate comparison for different number of features of AHUT experiment data.

be seen that multivariate fault extraction methods have much stronger separation capability than the single scale MDE based fault diagnosis one. For each class, 15 training samples are randomly selected and the remaining 14 are seen as the testing ones. Then MCFS is used to select $d$ ($d$ ranging 2 from 15) values from initial fault features of training samples to construct new fault training feature sets. Also the selected orders of training samples are used to select testing samples and form fault testing feature sets. Then the new fault training feature sets are used to train the SVM based multi-classifier and the identifying rates for different numbers of features are given in Fig. 12. For comparison purpose, the fault recognition rate of the MMSE, MMFE, MDE, MMDE based fault diagnosis methods with the different number of $d$ ($d$ ranging 2 from 15) are also shown in Fig. 12. It can be seen from Fig. 12 that the identifying rates of all fault diagnosis methods are 100 % when the number of fault feature larger than 3. The results show that the rolling bearing fault diagnosis method proposed in this paper can diagnose the fault locations and degrees of rolling bearings effectively.

## V. CONCLUSIONS

In this paper, refined composite multivariate multiscale dispersion entropy (RCMMDE) is proposed as a new nonlinear dynamic method for measuring correlation and complexity of multi-channel data. The proposed RCMMDE method is compared with MMSE, MMFE, MMDE and MDE through analyzing multichannel synthetic signals. The results show that the proposed RCMMDE has advantages in feature extraction stability and accuracy. The new fault diagnosis method for rolling bearing was proposed based on RCMMDE for feature extraction, MCFS for feature selection and SVM for mode classification. Also the proposed method is compared with the single channel MDE and multi-channel MMSE, MMFE and MMDE methods through analyzing rolling bearing experimental data analysis of CWRU and AHUT. The analysis result shows that the proposed fault diagnosis method for rolling bearing has higher fault recognition rate than the existing fault diagnosis methods.

## REFERENCES

[1] Z. Li, K. Goebel, and D. Wu, "Degradation modeling and remaining useful life prediction of aircraft engines using ensemble learning," *J. Eng. Gas Turbines Power*, vol. 141, no. 4, Nov. 2019, Art. no. 041008.

[2] Z. Li, D. Wu, C. Hu, and J. Terpenny, "An ensemble learning-based prognostic approach with degradation-dependent weights for remaining useful life prediction," *Rel. Eng. Syst. Saf.*, vol. 184, pp. 110–122, Apr. 2019.

[3] Z. Li, Y. Jiang, Q. Guo, C. Hu, and Z. Peng, "Multi-dimensional variational mode decomposition for bearing-crack detection in wind turbines with large driving-speed variations," *Renew. Energy*, vol. 116, pp. 55–73, Feb. 2018.

[4] H. Rahnejat and R. Gohar, "The vibrations of radial ball bearings," *Proc. Inst. Mech. Eng., Part C, J. Mech. Eng. Sci.*, vol. 199, no. 3, pp. 181–193, Jul. 1985.

[5] S. Vafaei, H. Rahnejat, and R. Aini, "Vibration monitoring of high speed spindles using spectral analysis techniques," *Int. J. Mach. Tools Manuf.*, vol. 42, no. 11, pp. 1223–1234, Sep. 2002.

[6] R. Yan, R. X. Gao, and X. Chen, "Wavelets for fault diagnosis of rotary machines: A review with applications," *Signal Process.*, vol. 96, pp. 1–15, Mar. 2014.

[7] R. Yan, Y. Liu, and R. X. Gao, "Permutation entropy: A nonlinear statistical measure for status characterization of rotary machines," *Mech. Syst. Signal Process.*, vol. 29, pp. 474–484, May 2012.

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

[9] 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.

[10] V. V. Nikulin and T. Brismar, "Comment on 'multiscale entropy analysis of complex physiologic time series'," *Phys. Rev. Lett.*, vol. 89, no. 6, pp. 705–708, Feb. 2004.

[11] D. Xiang and S. Ge, "Method of fault feature extraction based on EMD sample entropy and LLTSA," *J. Aerosp. Power*, vol. 29, no. 7, pp. 1535–1542, 2014.

[12] J. L. Lin, J. Y.-C. Liu, C.-W. Li, L.-F. Tsai, and H.-Y. Chung, "Motor shaft misalignment detection using multiscale entropy with wavelet denoising," *Expert Syst. Appl.*, vol. 37, no. 10, pp. 7200–7204, Oct. 2010.

[13] J. Zheng, H. Pan, S. Yang, and J. Cheng, "Generalized composite multiscale permutation entropy and Laplacian score based rolling bearing fault diagnosis," *Mech. Syst. Signal Process.*, vol. 99, pp. 229–243, Jan. 2018.

[14] 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. 296–311, Feb. 2017.

[15] W. Chen, Z. Wang, H. Xie, and W. Yu, "Characterization of surface EMG signal based on fuzzy entropy," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 15, no. 2, pp. 266–272, Jun. 2007.

[16] 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," *Mech. Mach. Theory*, vol. 78, no. 16, pp. 187–200, 2014.

[17] L. Zhang, L. Zhang, G. Xiong, J. Zhou, N. Wang, and M. Wang, "Rolling bearing fault diagnosis based on multiscale entropy and elman neural network," *Mech. Sci. Technol. Aerosp. Eng.*, vol. 33, no. 12, pp. 1854–1858, 2014.

[18] M. Rostaghi and H. Azami, "Dispersion entropy: A measure for time-series analysis," *IEEE Signal Process. Lett.*, vol. 23, no. 5, pp. 610–614, May 2016.

[19] H. Azami, E. Kinney-lang, A. Ebied, A. Fernández, and J. Escudero,, "Multiscale dispersion entropy for the regional analysis of resting-state magnetoencephalogram complexity in Alzheimer's disease," in *Proc. 39th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Seogwipo, South Korea, Jul. 2017, pp. 3182–3185.

[20] H. Azami, M. Rostaghi, D. Abásolo, and J. Escudero, "Refined composite multiscale dispersion entropy and its application to biomedical signals," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 12, pp. 2872–2879, Dec. 2017.

[21] L. Cao, A. Mees, and K. Judd, "Dynamics from multivariate time series," *Phys. D, Nonlinear Phenomena*, vol. 121, nos. 1–2, pp. 75–88, Oct. 1998.

[22] Z.-K. Gao *et al.*, "Multivariate multiscale entropy analysis of horizontal oil–water two-phase flow," *Phys., A Stat. Mech. Appl.*, vol. 417, pp. 7–17, Jan. 2015.

[23] Z.-K. Gao, Y.-X. Yang, L.-S. Zhai, M.-S. Ding, and N.-De Jin, "Characterizing slug to churn flow transition by using multivariate pseudo Wigner distribution and multivariate multiscale entropy," *Chem. Eng. J.*, vol. 291, pp. 74–81, May 2016.

[24] M. Costa and A. L. Goldberger, and C.-K. Peng, "Multiscale entropy analysis (MSE)," *Int. J. Numer. Methods Eng.*, vol. 59, no. 9, pp. 1147–1166, 2008.

[25] X. Zou, L. Feng, and H. Sun, "Robust compressive sensing of multichannel EEG signals in the presence of impulsive noise," *Inf. Sci.*, vol. 429, pp. 120–129, Mar. 2018.

[26] L. Márton, L. Bakó, S. T. Brassai, and L. Losonczi, "Multichannel EEG signal recording analysis based on cross frequency coupling method," *Procedia Technol.*, vol. 12, no. 1, pp. 133–140, 2014.

[27] H. Azami, A. Fernandez, and J. Escudero. (2017). "Multivariate multiscale dispersion entropy of biomedical times series." [Online]. Available: https://arxiv.org/abs/1704.03947?context=q-bio

[28] D. Cai, C. Zhang, and X. He, "Unsupervised feature selection for multicluster data," in *Proc. 16th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD)*, New York, NY, USA, Jul. 2010, pp. 333–342.

[29] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, Apr. 2011, Art. no. 27.

[30] Y. Li, Y. Yang, X. Wang, B. Liu, and X. Liang, "Early fault diagnosis of rolling bearings based on hierarchical symbol dynamic entropy and binary tree support vector machine," *J. Sound Vib.*, vol. 428, pp. 72–86, Aug. 2018.
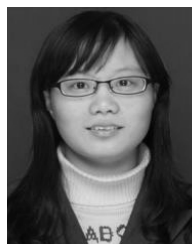
[31] Y. Li, X. Wang, S. Si, and S. Huang, "Entropy based fault classification using the case western reserve university data: A benchmark study," *IEEE Trans. Rel.*, to be published.
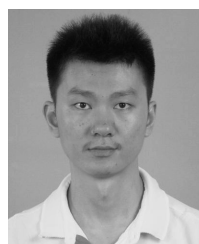
**JINDE ZHENG** received the B.S. degree in mathematics from Anhui Normal University, Wuhu, China, in 2009, and the Ph.D. degree in mechanical engineering from Hunan University, Changsha, China, in 2014. From 2014 to 2016, he was a Lecturer, and since 2017, he has been an Associate Professor with the Mechanical Engineering Department, Anhui University of Technology. He has authored more than 50 articles. His research interests include machinery health monitoring and fault diagnosis, statistical signal processing, and complexity theory.

**HAIYANG PAN** was born in Suzhou, China, in 1989. He received the master's degree in mechanical engineering from Hunan University, Changsha, China, in 2014, where he is currently pursuing the Ph.D. degree. His research interests include fault diagnosis, signal processing, and pattern recognition.

**JINYU TONG** was born in Huainan, China. She received the B.S. degree from the Anhui University of Technology, Ma'anshan, China, in 2011. She is currently pursuing the Ph.D. degree with Hohai University, Nanjing, China. Her research interests include statistical signal processing, vibration signal analysis, and measurement.

**CONGZHI LI** was born in Lu'an, China. He received the B.S. degree from the Anhui University of Technology, Ma'anshan, China, in 2016, where he is currently pursuing the master's degree. His research interests include fault diagnosis of gearbox and rolling bearings.

**YIFANG ZHANG** was born in Anqing, China. He received the Ph.D. degree from the University of Science and Technology Beijing, Beijing, China, in 2015. His main research interests include mechanical vibration and system dynamics analysis.

• • •