entropy

MDPI

Article

Fault Diagnosis Method for Rolling Bearings Based on Composite Multiscale Fluctuation Dispersion Entropy

Xiong Gan [1], Hong Lu [1] and Guangyou Yang [2,3,*]

1 School of Mechanical and Electronic Engineering, Wuhan University of Technology, Wuhan 430070, China; ganxiong001@126.com (X.G.); landzh@whut.edu.cn (H.L.)
2 Institute of Agricultural Machinery, Hubei University of Technology, Wuhan 430068, China
3 Hubei Province Engineering Technology Research Center for Intelligent Agricultural Machinery, Wuhan 430068, China
* Correspondence: pekka@126.com; Tel.: +86-027-59750857

Received: 17 February 2019; Accepted: 14 March 2019; Published: 18 March 2019



Abstract: This paper proposes a new method named composite multiscale fluctuation dispersion entropy (CMFDE), which measures the complexity of time series under different scale factors and synthesizes the information of multiple coarse-grained sequences. A simulation validates that CMFDE could improve the stability of entropy estimation. Meanwhile, a fault recognition method for rolling bearings based on CMFDE, the minimum redundancy maximum relevancy (mRMR) method, and the k nearest neighbor (kNN) classifier (CMFDE-mRMR-kNN) is developed. For the CMFDE-mRMR-kNN method, the CMFDE method is introduced to extract the fault characteristics of the rolling bearings. Then, the sensitive features are obtained by utilizing the mRMR method. Finally, the kNN classifier is used to recognize the different conditions of the rolling bearings. The effectiveness of the proposed CMFDE-mRMR-kNN method is verified by analyzing the standard experimental dataset. The experimental results show that the proposed fault diagnosis method can effectively classify the conditions of rolling bearings.

Keywords: CMFDE; mRMR; rolling bearings; fault diagnosis

1. Introduction

The working performance of rolling bearings directly affects the safety, reliability, and stability of rotating machinery. When the rolling bearings fail, the collected vibration signals often exhibit complex nonlinear characteristics [1–4]. In recent years, time–frequency analysis methods have been exploited to decompose vibration signals by using empirical mode decomposition (EMD) [5,6], local mean decomposition (LMD) [7,8], and intrinsic timescale decomposition (ITD) [9,10]. However, these signal processing methods have some weaknesses such as end effect and mode mixing, which often fail to extract fault features effectively and reduce the fault recognition accuracy. Furthermore, mechanical fault diagnosis methods based on thermodynamic entropy [11–14] have been proposed. Nevertheless, thermodynamic data are difficult to collect online and are sensitive to the environment, which is not suitable for the early fault diagnosis of rolling bearings. Therefore, it is very necessary to study a novel fault diagnosis method for rolling bearings. With different faults, the vibration signal complexity of rolling bearings could be different. Hence, the nonlinear dynamics theory can be directly applied to extract the fault features of rolling bearings without signal decomposition or transformation.

In 1991, Pincus et al. proposed approximate entropy (APE) to measure the complexity of nonlinear time series and applied it to analyze physiological time series [15]. Subsequently, APE was utilized for extracting the fault features of rolling bearings [16]. The results show that APE can effectively

characterize the types of bearings. However, APE has the disadvantage of producing inaccurate entropy values and of having a tedious computational time. Based on the research on APE, sample entropy (SE) was created by Richman et al. in 2000 [17], which could be more immune to the noise interference, and was applied in the fault diagnosis method for the rolling bearings by Han et al. [18]. However, the shortcomings of a longer calculation time and being vulnerable to mutation signal still existed in SE. In 2002, Bandt et al. [19] proposed the permutation entropy (PE), which possessed a fast computational speed, a simple calculation process, and the capacity to resist disturbance. Considering the advantages of PE, Cheng et al. [20] applied PE to the fault recognition of rolling bearings. Meanwhile, in consideration of the deficiency of single scale, multiscale sample entropy (MSE) [21] and multiscale permutation entropy (MPE) [22–24] were proposed, which were used to extract the fault features for rolling bearings [25] and reflect the characteristics of milling force signals [26], respectively. However, PE also failed to consider the signal amplitude. In 2016, to overcome the problems of APE, SE, and PE, Hamed Azami et al. [27] proposed a nonlinear time complexity evaluation method based on dispersion entropy (DE), which would not create unreliable entropy values as well as was insensitive to noise interference, had high computational efficiency, and could solve the equivalence problem. Thereafter, Mostafa Rostaghi et al. [28] applied DE to classify the conditions of rolling bearings and gears. In addition, Hamed Azami et al. [29] proposed fluctuation dispersion entropy (FDE), which considered the fluctuation of nonlinear time series on the basis of DE. However, FDE still had some space to be improved; for example, it only evaluated the complexity of nonlinear time series from a single scale factor.

In this work, multiscale fluctuation dispersion entropy (MFDE) was originally proposed to analyze the complexity of nonlinear time series under various scale factors. To overcome the shortcomings of the coarsening method in the multiscale process for MFDE, the composite multiscale fluctuation dispersion entropy (CMFDE) was further proposed, which could synthesize the information of multiple coarse-grained sequences and could reduce the standard deviation (SD) of entropy estimation. Meanwhile, a novel fault diagnosis method based on CMFDE, minimum redundancy maximum relevancy (mRMR) [30], and k nearest neighbor (kNN) (CMFDE-mRMR-kNN) is developed for the rolling bearings. As the core of this method, the function of CMFDE is to extract the nonlinear fault features of rolling bearings. Then, mRMR is applied to extract the sensitive fault features to reduce feature dimension and to improve the efficiency of fault diagnosis method. Moreover, the kNN classifier is chosen to recognize the fault conditions of the rolling bearings. The experimental results show that the proposed method can effectively extract the fault features of rolling bearings and can achieve a high recognition accuracy.

2. FDE, MFDE, and CMFDE

2.1. FDE

Input: Nonlinear time series $X = \{\mathrm{x}(1), \mathrm{x}(2), \cdots, \mathrm{x}(N)\}$, embedding dimension m, time delay τ, and number of classes c

Output: *FDE*(X,m,τ,c)

Step 1. $X = \{\mathrm{x}(1), \mathrm{x}(2), \cdots, \mathrm{x}(N)\}$ is mapped to $Y = \{\mathrm{y}(1), \mathrm{y}(2), \cdots, \mathrm{y}(N)\}$ by using the normal cumulative distribution function (NCDF), $\mathrm{y}(j) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\mathrm{x}(j)} e^{\frac{-(t-\mu)}{2\sigma^2}} dt$. Here, μ and σ are the mean and SD of X, respectively.

Step 2. $y(j)$ is mapped to $z(j)$, which can be defined as $z(j)$=$R(c \cdot y(\mathrm{j})+0.5)$, where c is an integer and $R(\cdot)$ represents the rounding function.

Step 3. For Z = {z(1),z(2), $\cdots$, z(N)}, with embedding dimension m and time delay τ, the embedding vector $Z^{m.\tau,c} = \{Z^{m.\tau,c}(1), Z^{m.\tau,c}(2), \cdots, Z^{m.\tau,c}(\mathrm{N} - (\mathrm{m}-1)\tau)\}$ is defined as

$$\begin{cases} Z^{m.\tau,c}(1) = \{z(1), z(1+\tau), \cdots, z(1+(m-1)\tau)\} \\ Z^{m.\tau,c}(2) = \{z(2), z(2+\tau), \cdots, z(2+(m-1)\tau)\} \\ \vdots \\ Z^{m.\tau,c}(N-(m-1)\tau) = \{z(N-(m-1)\tau), z(N-(m-2)\tau), \cdots, z(N)\} \end{cases} \tag{1}$$

Step 4. Transform $Z^{m.\tau,c} = \{Z^{m.\tau,c}(1), Z^{m.\tau,c}(2), \cdots, Z^{m.\tau,c}(N-(m-1)\tau)\}$ into $F^{m.\tau,c} = \{F^{m.\tau,c}(1), F^{m.\tau,c}(2), \cdots, F^{m.\tau,c}(N-(m-1)\tau)\}$ by the following equation

$$\begin{cases} F^{m.\tau,c}(1) = \{z(1+\tau) - z(1) + c, \cdots, z(1+(m-1)\tau) - z(1+(m-2)\tau) + c\} \\ F^{m.\tau,c}(2) = \{z(2+\tau) - z(2) + c, \cdots, z(2+(m-1)\tau) - z(2+(m-2)\tau + c)\} \\ \vdots \\ F^{m.\tau,c}(N-(m-1)\tau) = \{z(N-(m-2)\tau) - z(N-(m-1)\tau) + c, \cdots, z(N) - z(N-\tau) + c\} \end{cases} \tag{2}$$

Step 5. $F^{m.\tau,c}(j) = \{z(j+\tau) - z(j) + c, \cdots, z(j+(m-1)\tau) - z(j+(m-2)\tau) + c\}$ is assigned to the dispersion pattern $\pi_{v_0 v_1 \cdots v_{(m-2)}}$, where $v_0 = z(j+\tau) - z(j) + c, v_1 = z(j+2\tau) - z(j+\tau) + c, \cdots, v_{(m-2)} = z(j+(m-1)\tau) - z(j+(m-2)\tau) + c$. For each possible dispersion pattern $\pi_{v_0 v_1 \cdots v_{(m-2)}}$, its relative frequency $p(\pi_{v_0 v_1 \cdots v_{(m-2)}})$ is given by

$$p(\pi_{v_0 v_1 \cdots v_{(m-2)}}) = \frac{number\left\{ j \middle| j \le N-(m-1)\tau, F^{m.\tau,c}(j) \text{ has type } \pi_{v_0 v_1 \cdots v_{(m-2)}} \right\}}{N-(m-1)\tau} \tag{3}$$

Step 6. $FDE(X, m, \tau, c)$ can be calculated by

$$FDE(X, m, \tau, c) = -\sum_{\pi=1}^{(2c-1)^{m-1}} p(\pi_{v_0 v_1 \cdots v_{(m-2)}}) \ln\left(p(\pi_{v_0 v_1 \cdots v_{(m-2)}})\right) \tag{4}$$

2.2. MFDE

Input: Nonlinear time series $X = \{x(1), x(2), \cdots, x(N)\}$, embedding dimension m, time delay τ, number of classes c, and maximum scale factor s_{max}

Output: *MFDE*

Initialization: *MFDE* = Ø and $s = 1$

Step 1. $X = \{x(1), x(2), \cdots, x(N)\}$ is divided to $y^s = \{y^s(j)\}_{j=1}^{[N/s]}$, and $y^s(j)$ is defined as

$$y^s(j) = \frac{1}{s} \sum_{i=(j-1)s+1}^{js} x(i) \tag{5}$$

where $j = 1, 2, \cdots, [N/s]$, $[\cdot]$ represents the rounding function.

Step 2. $MFDE(X, m, \tau, c, s)$ can be obtained by

$$MFDE(X, m, \tau, c, s) = FDE(y^s, m, \tau, c) \tag{6}$$

Step 3. $MFDE = MFDE \cup \text{MFDE}(X, m, \tau, c, s)$, $s = s + 1$. Keep running Steps 1–3 if s is smaller than s_{max}.

2.3. CMFDE

Input: Nonlinear time series $X = \{x(1), x(2), \cdots, x(N)\}$, embedding dimension m, time delay τ, number of classes c, and maximum scale factor s_{max}

Output: *CMFDE*

Initialization: *CMFDE* = Ø and $s = 1$

Step 1. $X = \{x(1), x(2), \cdots, x(N)\}$ is divided to $\mathrm{y}^{s,q} = \{\mathrm{y}^{s,q}(j)\}_{j=1}^{[(N+1)/\mathrm{s}]-1}$, and $\mathrm{y}^{s,q}(j)$ is defined as

$$\mathrm{y}^{s,q}(j) = \frac{1}{s} \sum_{i=(j-1)\mathrm{s}+q}^{j\mathrm{s}+q-1} x(i) \tag{7}$$

where $j = 1, 2, \cdots, [(N+1)/s] - 1$, $q = 1, 2, \cdots, s$.

Step 2. For each scale factor s, calculate FDE of $\mathrm{y}^{s,1}$, $\mathrm{y}^{s,2}$, ..., $\mathrm{y}^{s,s}$, and then, average these FDE values as $CMFDE(X, m, \tau, \mathrm{c}, s)$, which can be obtained by

$$CMFDE(X, m, \tau, \mathrm{c}, s) = \frac{1}{s} \sum_{q=1}^{s} FDE(y^{s,q}, m, \tau, \mathrm{c}) \tag{8}$$

Step 3. $CMFDE = CMFDE \cup \mathrm{CMFDE}(X, m, \tau, \mathrm{c}, s)$, $s = s + 1$. Keep running Steps 1–3 if s is smaller than s_{max}.

Parameters analysis of CMFDE: If m is too small, CMFDE cannot accurately observe the dynamic behavior of a nonlinear time series. Conversely, if m is too large, CMFDE cannot detect small variations. τ has a small effect on CMFDE. When c is too large, CMFDE is sensitive to noise. When c is too small, two very different amplitude values are assigned to the same category. If s_{max} is too small, CMFDE cannot fully extract the fault features of a nonlinear time series. If s_{max} is too large, CMFDE will generate unstable entropy values. In addition, larger m, c, and s_{max} will lead to low computational efficiency. Here, m, τ, c, and s_{max} are selected as 2, 1, 5, and 20, respectively.

Figure 1 shows the flow charts of MFDE and CFMDE.

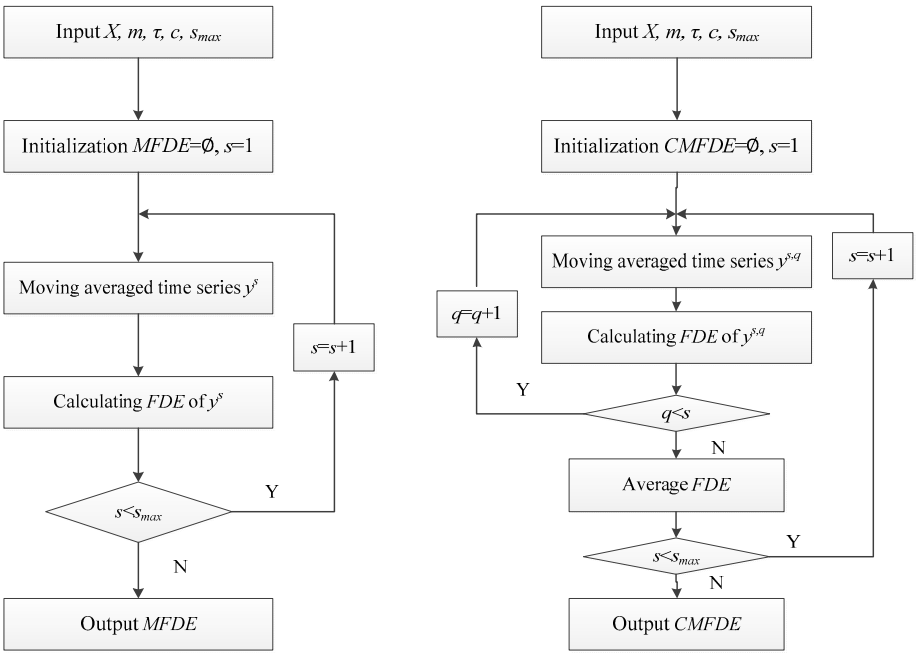


Figure 1. The flow charts of multiscale fluctuation dispersion entropy (MFDE) and composite multiscale fluctuation dispersion entropy (CMFDE).

2.4. Comparison between CMFDE and MFDE

To verify the advantages of CMFDE, the white noise and pink noise are adopted to perform the comparison between MFDE and CMFDE. Here, 100 groups of white noise and pink noise with data lengths of 3000 points are generated. Figures 2–5 show the time domain waveform and frequency spectrum of two simulated signals, respectively. Figure 6 shows the mean and SD values of MFDE and CMFDE for two simulated signals. In this simulation, the embedding dimension is m = 2, the time delay is τ = 1, the number of classes is c = 5, and the largest number of scale factor is s_{max} = 20.

As shown in Figure 6, under the same scale factor, the difference between the mean value of MFDE and the mean value of CMFDE is unobvious. As the scale factor s increases, for white noise, MFDE and CFMDE decrease gradually. However, for pink noise, MFDE and CFMDE tend to be constant values. In addition, with a large-scale factor, the MFDE and CFMDE values of white noise are less than the corresponding entropy values of pink noise. The reason is that the information of white noise is mainly located in the small-scale factor and that the white noise is more irregular than the pink noise. Especially, the SD value of CMFDE is smaller than that of MFDE for both the white noise and pink noise, which indicates that CMFDE displays a better stability than MFDE for the entropy evaluation of nonlinear time series.

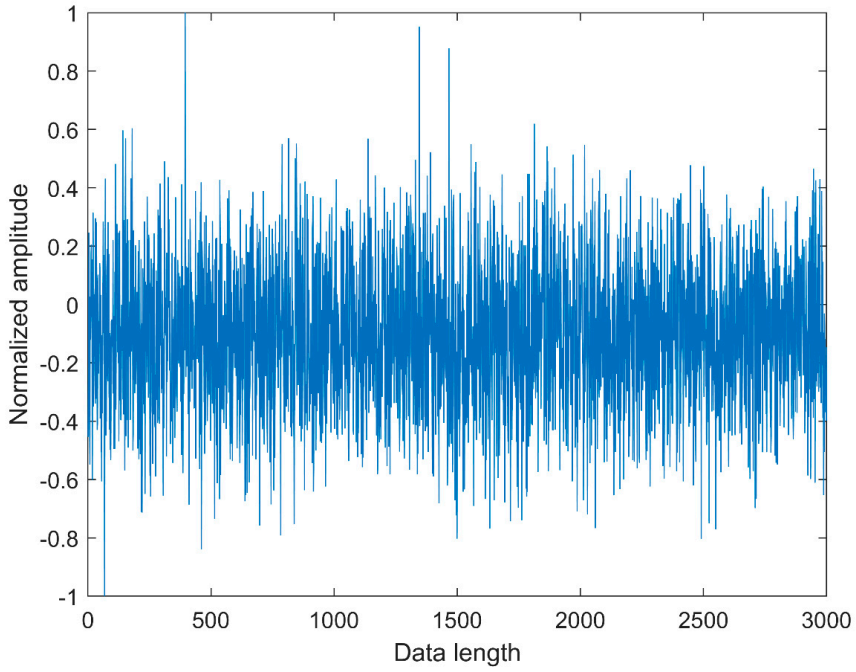


Figure 2. The waveform of white noise.

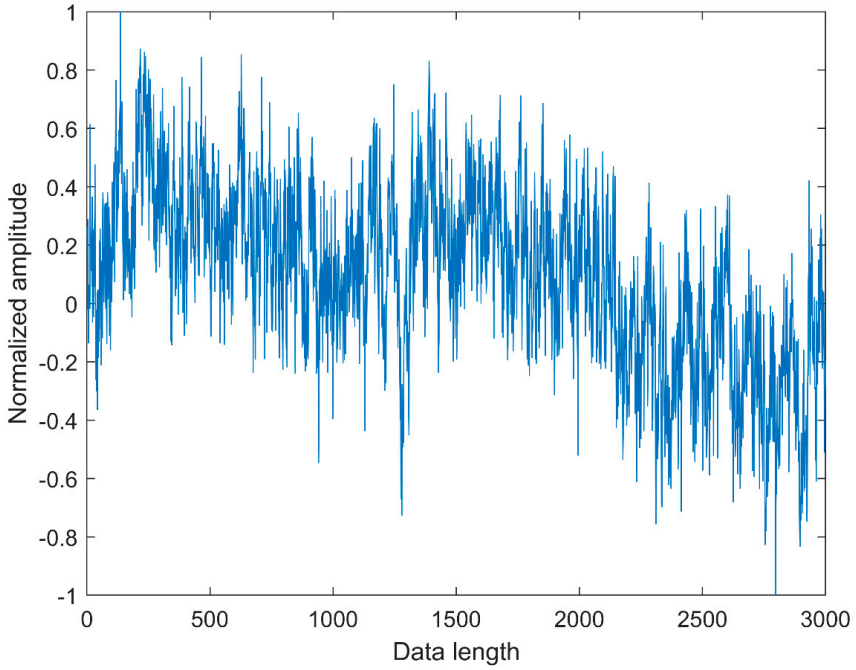


Figure 3. The waveform of pink noise.

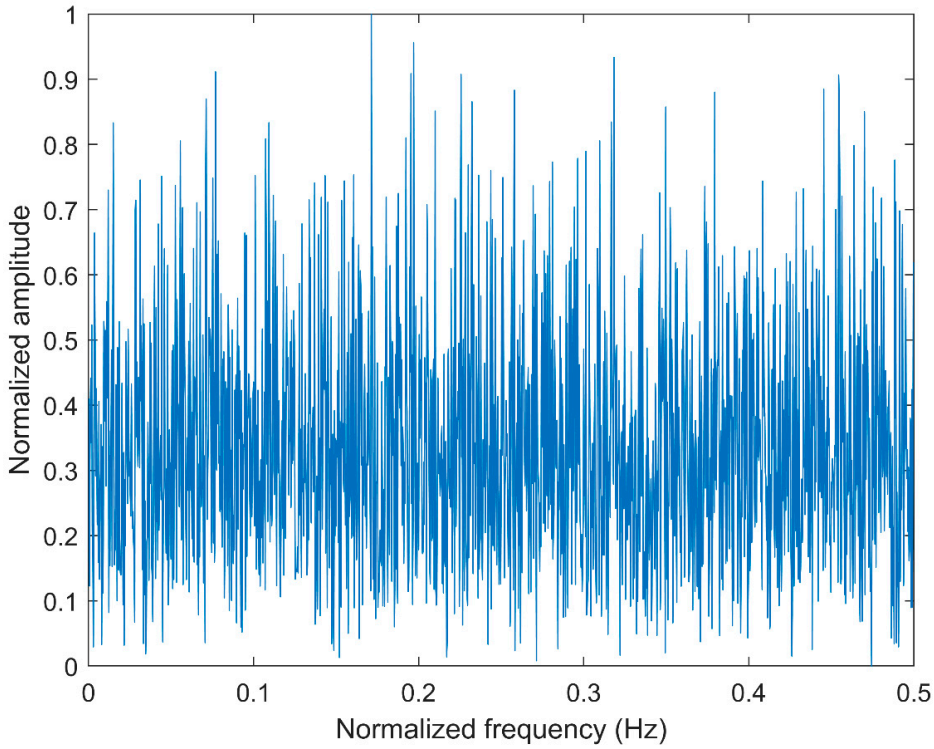


Figure 4. The frequency spectrum of white noise.

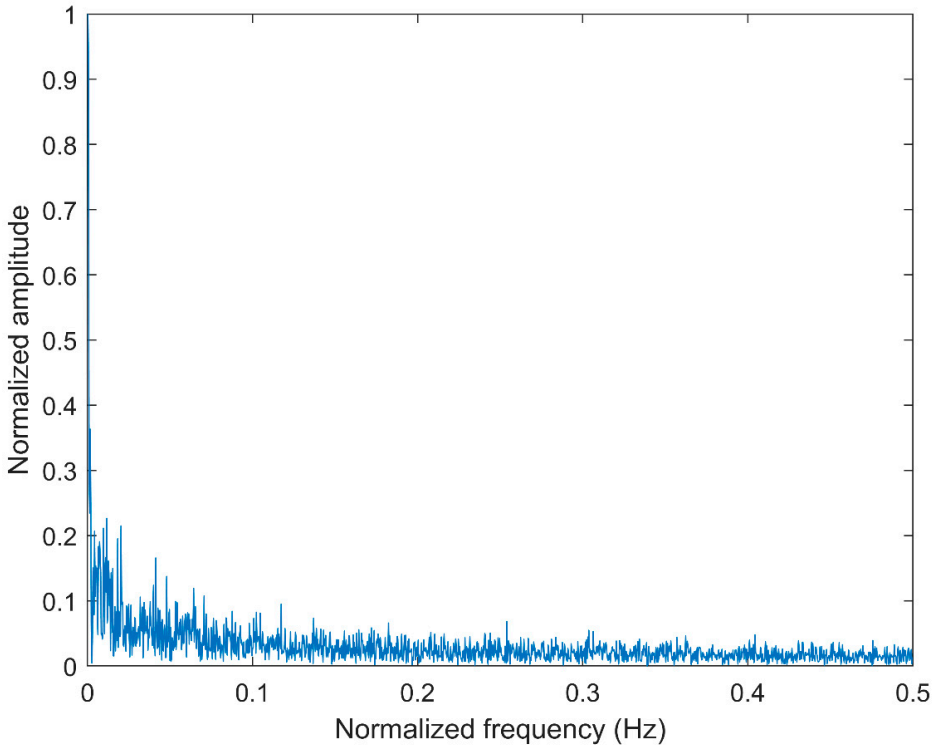


Figure 5. The frequency spectrum of pink noise.

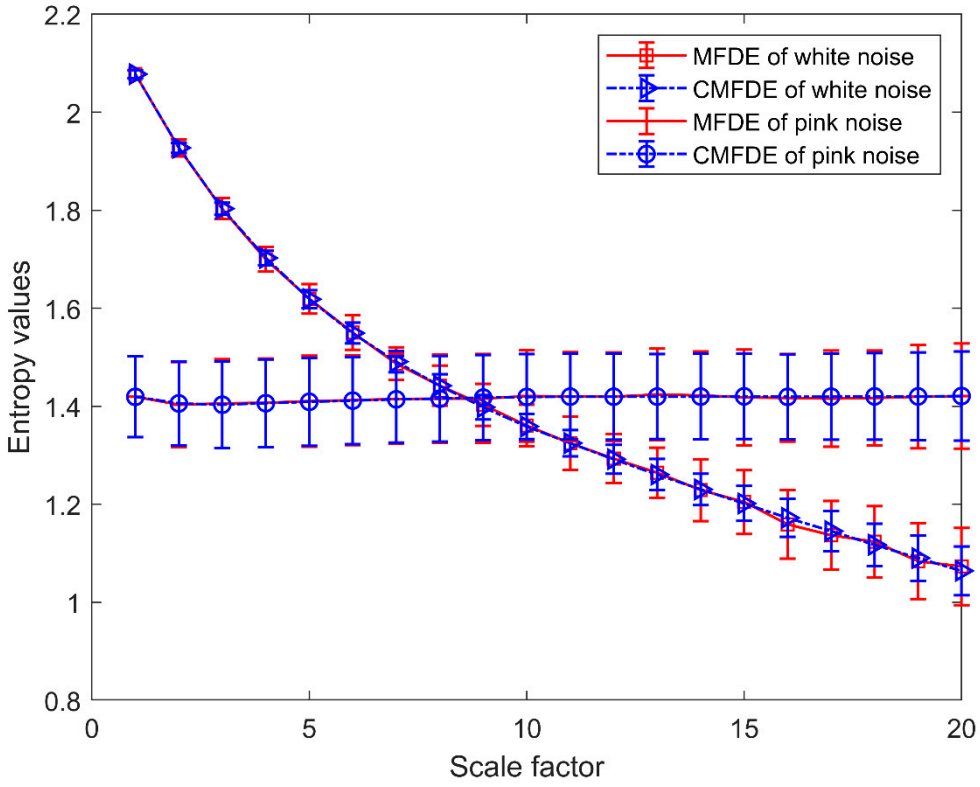


Figure 6. The entropy values of the two simulated signals.

3. Fault Diagnosis Method for Rolling Bearings Based on CFMDE, mRMR, and KNN

mRMR Feature Selection

Sequence the forward search strategy to get the sensitive features from original feature set OF, assume the k features have been chosen to set up the sensitive feature set SF, and choose the (k+1)-th feature from the remaining feature set $\{OF - SF\}$ according to the following criteria

$$sf_{k+1} = \text{argmax}_{of_j \in OF-SF}\left\{ \mathrm{I}(of_j; L) - \frac{1}{k} \sum_{of_i \in OF-SF} \mathrm{I}(of_i; of_j) \right\} \tag{9}$$

where $\mathrm{I}(of_j; L)$ is the mutual information between feature of_j and class L and where $\mathrm{I}(of_i; of_j)$ is the mutual information between feature of_i and feature of_j.

The concrete steps of the mRMR method can be described as follows:

Input: Original feature set $OF = \{of_1, of_2, \cdots, of_s\}$, label information L, and the number of sensitive feature subset p

Output: Sensitive feature subset $SF = \{sf_1, sf_2, \cdots, sf_p\}$

Initialization: $SF = \emptyset$ and $p = 0$

Step 1. Choose the first sensitive feature sf_1 according to $sf_1 = argmax_{i=1,2,\cdots,s}\{I(of_i; L)\}$, $SF = \{sf_1\}$, and $k = 1$.

Step 2. Choose the (k+1)th sensitive feature sf_{k+1} from the remaining feature set $\{OF - SF\}$ according to $sf_{k+1} = \text{argmax}_{of_j \in OF-SF}\left\{ \mathrm{I}(of_j; L) - \frac{1}{k} \sum_{of_i \in OF-SF} \mathrm{I}(of_i; of_j) \right\}$, $SF = SF \cup \{sf_{k+1}\}$, and $k = k + 1$.

Step 3. Keep running Step 2 until the number of sensitive feature subset $|SF|$ is equal to p.

The CMFDE-mRMR-kNN method includes the following three steps.

Input: Input the training vibration signals of K different conditions $VS_{train} = [VS^1_{train}, VS^2_{train}, \cdots, VS^K_{train}]$, testing vibration signals VS_{test}, number of sample points N, embedding

dimension m, number of classes c, time delay τ, maximum scale factor s_{max}, and number of sensitive features p

Output: Testing label L_{test}

Step 1. Vibration signals collection and samples design: The training vibration signals VS_{train} under each condition are divided into nonoverlapping data samples, each sample contains N points, each condition contains M_{train} samples, and the training label is L_{trian}. Meanwhile, splitting the testing vibration signals VS_{test} into M_{test} testing data samples without overlapping, each sample also contains N points.

Step 2. Feature extraction and feature selection: CMFDE method is adopted to extract the fault features and to obtain the training matrix T_{train}. The mRMR method is employed to choose the first p features to achieve the sensitive training matrix $T_{train,\mathrm{mRMR}}$. Meanwhile, according to the ranking results in the training samples, the sensitive testing matrix $T_{test,\mathrm{mRMR}}$ can be obtained.

Step 3. Pattern recognition: To classify the conditions of the rolling bearings, input $T_{train,\mathrm{mRMR}}$, L_{trian}, and $T_{test,\mathrm{mRMR}}$ into the kNN classifier and output the testing label L_{test}.

Figure 7 shows the flowchart of the CMFDE-mRMR-kNN method.

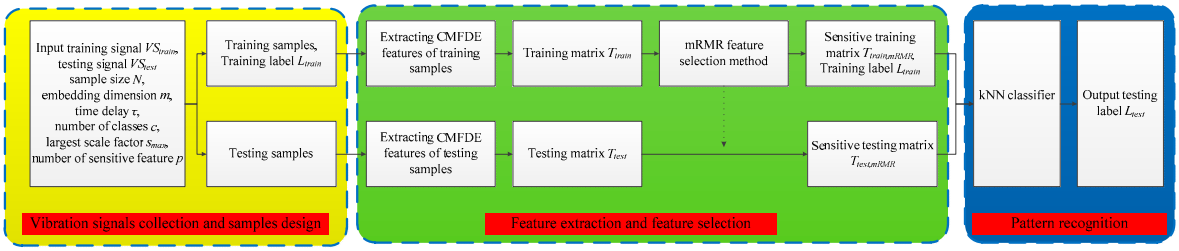


Figure 7. Flowchart of the CMFDE-mRMR-kNN method.

4. Experimental Verification and Analysis

The case university bearings dataset is utilized to validate the effectiveness of the proposed CMFDE-mRMR-kNN method. The test bench is shown in Figure 8; more system details are described in Reference [31]. The experimental bearing is 6203–2RS JEM SKF (SKF, Gothenburg, Sweden) at the fan end. The accelerometer is used for gathering the vibration signals. Vibration signals under 10 different conditions are collected, which are shown in Figure 9. The sampling frequency is 12,000 Hz. A detailed description is shown in Table 1. It is hard to recognize the specific fault condition of bearings based on the time domain waveform. Thus, the proposed CMFDE-mRMR-kNN method is employed for classifying different conditions of rolling bearings.

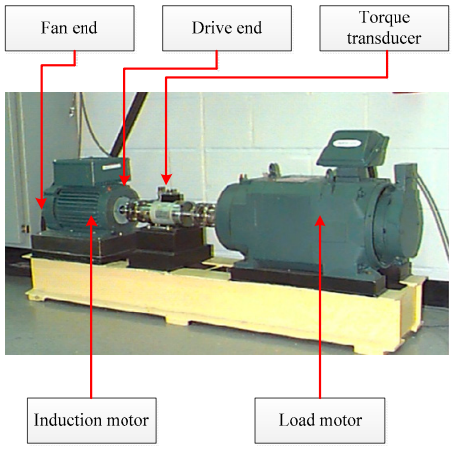


Figure 8. The test bench.

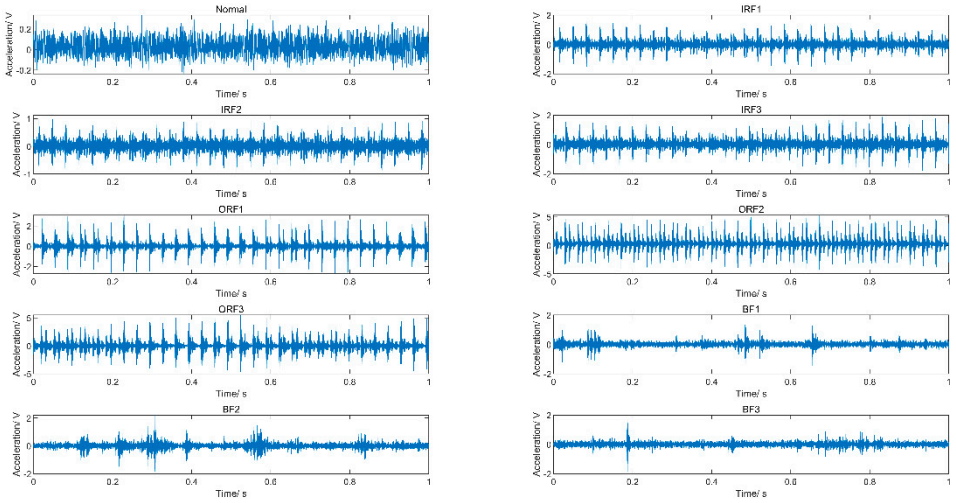


Figure 9. The time domain waveform under 10 different conditions.

Table 1. The detailed descriptions for 10 different conditions.

Different Conditions	Abbreviation	Fault Diameter (mm)	Motor Load (HP)	Motor Speed (rpm)	Label
Normal	Normal	0	0	1797	1
Inner race fault	IRF1	0.1778	0	1797	2
Inner race fault	IRF2	0.3556	0	1797	3
Inner race fault	IRF3	0.5334	0	1797	4
Outer race fault	ORF1	0.1778	0	1797	5
Outer race fault	ORF2	0.3556	0	1797	6
Outer race fault	ORF3	0.5334	0	1797	7
Ball fault	BF1	0.1778	0	1797	8
Ball fault	BF2	0.3556	0	1797	9
Ball fault	BF3	0.5334	0	1797	10

The vibration signals under 10 different conditions with the number of sample points $N = 1024$, the embedding dimension $m = 2$, the time delay $\tau = 1$, the number of classes $c = 5$, the maximum scale factor $s_{max} = 20$, and the number of sensitive features $p = 10$ are inputted. Splitting the vibration signals into multiple data samples without overlapping, each condition contains $M = 100$ data samples. For each condition, $M_{train} = 25$ samples are randomly selected as the training samples, the training label was L_{trian}, and the remaining $M_{test} = 75$ samples are used as testing samples. Then, the training matrix $T_{train} \in \Re^{20\times250}$ is computed, and the first 10 features to achieve a sensitive training matrix $T_{train,mRMR} \in \Re^{10\times250}$ by adopting the mRMR method are chosen. According to the feature selection results, the sensitive testing matrix $T_{test,mRMR} \in \Re^{10\times750}$ is obtained. Finally, $T_{train,\mathrm{mRMR}}$, L_{trian}, and $T_{test,\mathrm{mRMR}}$ are inputted to the kNN classifier($k = 1$), and the testing label L_{test} is outputted.

To verify the superiority of CMFDE-mRMR-kNN, the recognition accuracy of the MSE-mRMR-kNN, MPE-mRMR-kNN, MDE-mRMR-kNN, MFDE-mRMR-kNN, CMDE-mRMR-kNN, and CMFDE-mRMR-kNN are compared. For MSE-mRMR-kNN, MPE-mRMR-kNN, MDE-mRMR-kNN, MFDE-mRMR-kNN, and CMDE-mRMR-kNN, the fault features by MSE, MPE, MDE, MFDE, and CMDE are extracted, separately. Then, mRMR method is employed to choose the sensitive features. Finally, the sensitive features are input to the kNN classifier and output the testing label. For MSE-mRMR-kNN, $m = 2$, $r = 0.15$std, and $s_{max} = 20$. For MPE-mRMR-kNN, $m = 5$, $\tau = 1$, and $s_{max} = 20$. For MDE-mRMR-kNN, $m = 2$, $\tau = 1$, $c = 5$, and $s_{max} = 20$. For MFDE-mRMR-kNN, $m = 2$, $\tau = 1$, $c = 5$, and $s_{max} = 20$. For CMDE-mRMR-kNN, $m = 2$, $\tau = 1$, $c = 5$, and $s_{max} = 20$. Here, m is the embedding dimension, τ is the time delay, c is the number of classes, r is the tolerance, and s_{max} is the largest scale factor. Figure 10 and Table 2 show the identification accuracy of different fault diagnosis methods within 50 run times. In Figure 11 and Table 2, the proposed CMFDE-mRMR-kNN

method attains the highest identification accuracy (96.53%–93.06%). As demonstrated, the effectiveness of the CMFDE-mRMR-kNN method is validated, and the advantage of CMFDE in extracting sensitive features is also highlighted. Then, the SD of other methods is significantly larger than that of the proposed method, which shows that the proposed CMFDE-mRMR-kNN method has a better stability.

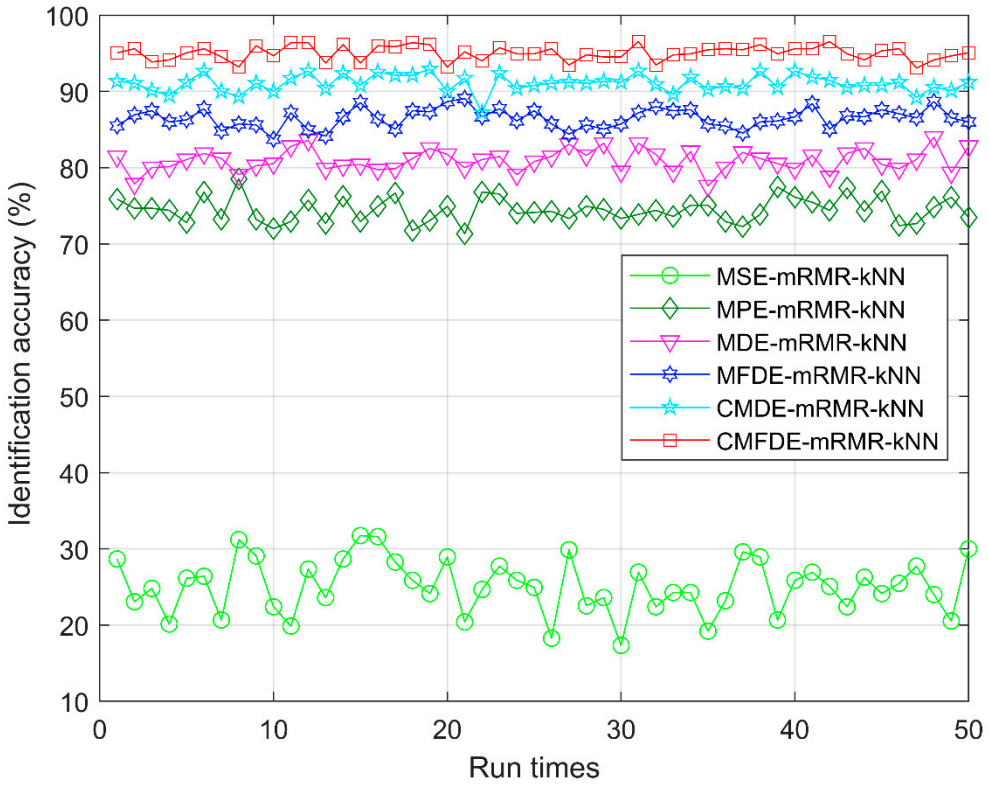


Figure 10. The identification accuracy with the minimum redundancy maximum relevancy (mRMR) method.

Table 2. The identification results with the mRMR method.

Different Methods	Accuracy (%)			
	Maximum	Minimum	Mean	SD
CMFDE-mRMR-kNN	96.53	93.06	95.02	0.95
CMDE-mRMR-kNN	92.93	87.20	91.04	1.19
MFDE-mRMR-kNN	89.06	83.73	86.47	1.25
MDE-mRMR-kNN	84.00	77.60	80.93	1.43
MPE-mRMR-kNN	78.53	71.33	74.47	1.65
MSE-mRMR-kNN	31.73	17.33	25.11	3.56

To investigate the effectiveness of the mRMR method, Figure 11 shows the distribution of the first two initial features without using the mRMR method. Afterwards, Figure 12 shows the distribution of the first two sensitive features by employing the mRMR method. As seen in Figure 11, the clustering centers of different types are mixed together, which makes it difficult to distinguish the different conditions. However, the clustering ability in Figure 12 is superior to the ability in Figure 11, and different conditions can be easily recognized. In order to further verify the necessity of the mRMR method in MSE-mRMR-kNN, MPE-mRMR-kNN, MDE-mRMR-kNN, MFDE-mRMR-kNN, CMDE-mRMR-kNN, and CMFDE-mRMR-kNN, the first 10 scale factors of the original features are selected instead of applying the mRMR method and the new MSE-kNN, MPE-kNN, MDE-kNN, MFDE-kNN, and CMDE-kNN, CMFDE-kNN methods can be obtained. In order to reduce the impact of randomness, 50 run times are conducted under the same parameters used in Figure 10 and Table 2.

The identification results are shown in Figure 13 and Table 3. As demonstrated, all the identification accuracies of the five methods are lower than that combined with the mRMR method, which verify the superiority of mRMR in selecting the sensitive features. It can be also found that the CMFDE-kNN method still achieves the highest diagnosis accuracy, which further demonstrates the advantage of the CMFDE for feature extraction.

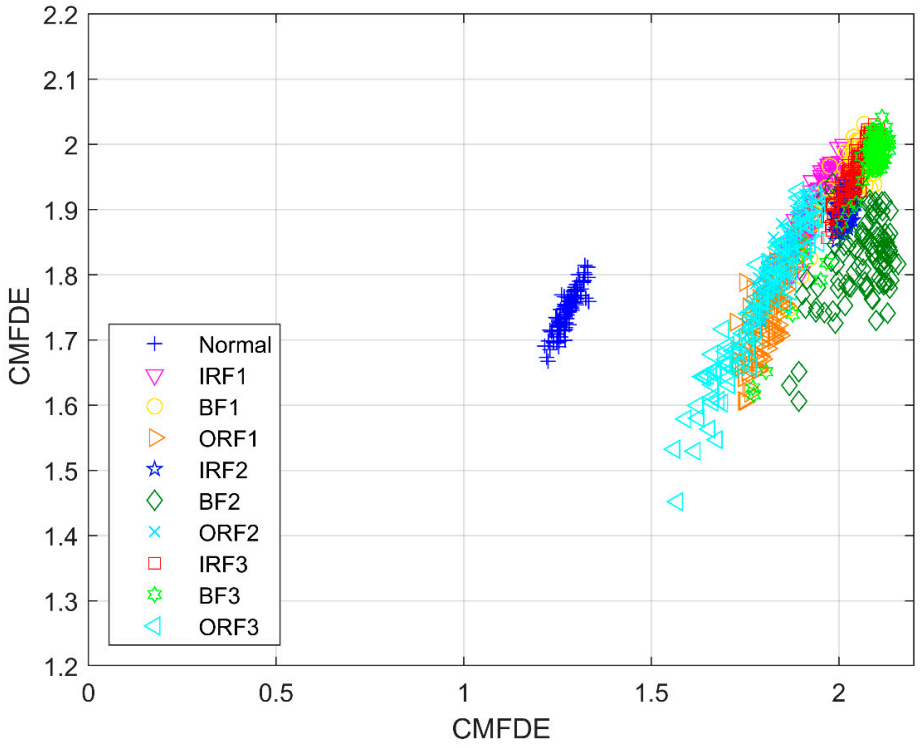


Figure 11. The distribution of the first two CMFDE features without using the mRMR method.

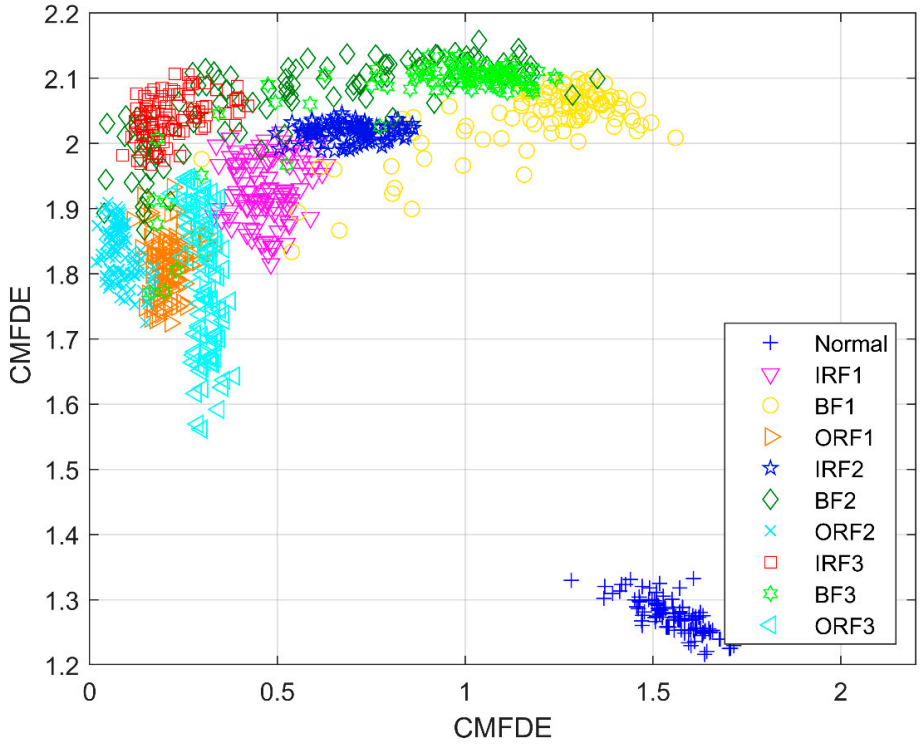


Figure 12. The distribution of the first two CMFDE features with the mRMR method.

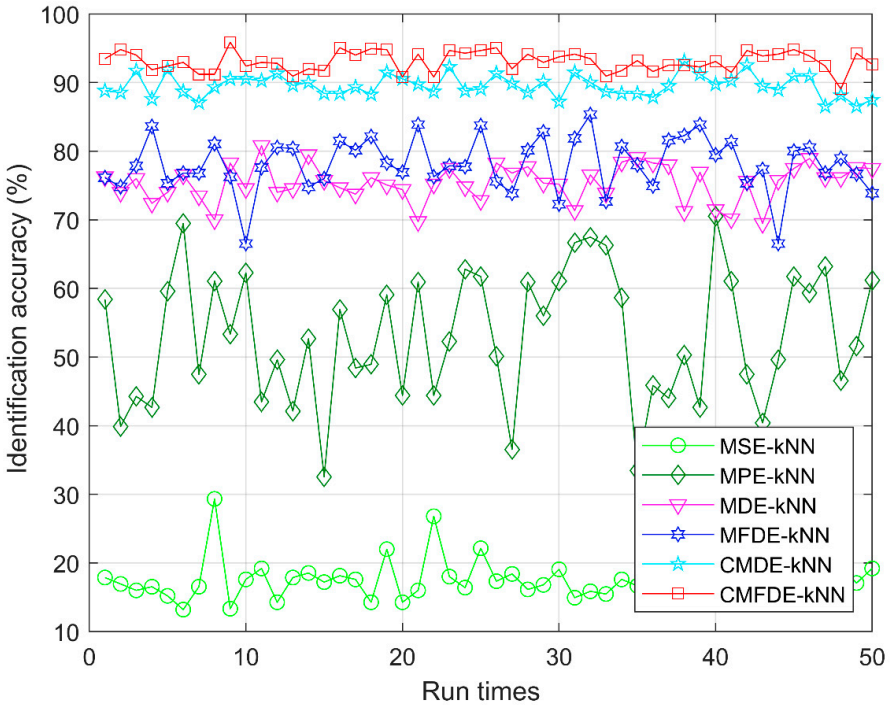


Figure 13. The identification accuracy without using the mRMR method.

Table 3. The identification results without using the mRMR method.

Different Methods	Accuracy (%)			
	Maximum	Minimum	Mean	SD
CMFDE-kNN	95.87	89.06	93.06	1.45
CMDE-kNN	93.06	86.53	89.55	1.56
MFDE-kNN	85.33	66.53	78.11	3.98
MDE-kNN	80.80	69.46	75.34	2.68
MPE-kNN	70.53	32.53	53.02	9.63
MSE-kNN	29.33	13.20	18.27	3.49

To analyze the relationship between the identification accuracy and the number of training/testing samples, we divide the number of training/testing samples into five situations (5/95, 25/75, 50/50, 75/25, and 95/5) and design separately to compute the average identification accuracy of different diagnosis methods within 50 run times. Figure 14 shows the average classification accuracy under different sizes of training/testing samples. As shown in Figure 14, even though the number of training/testing samples is different, the proposed CMFDE-mRMR-kNN still obtains the highest average classification accuracy. Meanwhile, if the number of training samples is far less than that of testing samples (such as 5/95), the different methods obtain the least average identification accuracy. It can be found that, if there are sufficient training samples (such as 25/75, 50/50, 75/25, and 95/5), five different methods can achieve a higher average identification accuracy. However, excessive training samples could result in longer training computation time. Therefore, in order to achieve the balance between identification accuracy and computational efficiency, the number of training/testing samples is 25/75.

To analyze the relationship between the identification accuracy and the size of sensitive features, an average identification accuracy can be achieved by conducting 50 run times. The corresponding average identification result is shown in Figure 15. It can be found that the CMFDE-mRMR-kNN method can achieve much a higher identification accuracy, which verifies the superiority of the proposed method. Moreover, with the increasing size of the selected features, the identification accuracy is not always increasing, a too large or too small size of sensitive fault features

will lead to a decline in the identification accuracy. The reason is that a too small size of sensitive features contains less fault information. On the contrary, a too large size of sensitive features will lead to a redundancy of fault information and a reduction of the identification accuracy. For MSE-mRMR-kNN, MPE-mRMR-kNN, MDE-mRMR-kNN, MFDE-mRMR-kNN, CMDE-mRMR-kNN, and CMFDE-mRMR-kNN, the optimal sizes of sensitive features are 4, 6, 15, 9, 17, and 12, respectively. The corresponding highest average identification accuracies are 62.34%, 76.26%, 81.49%, 86.70%, 91.71% and 95.20%, respectively.

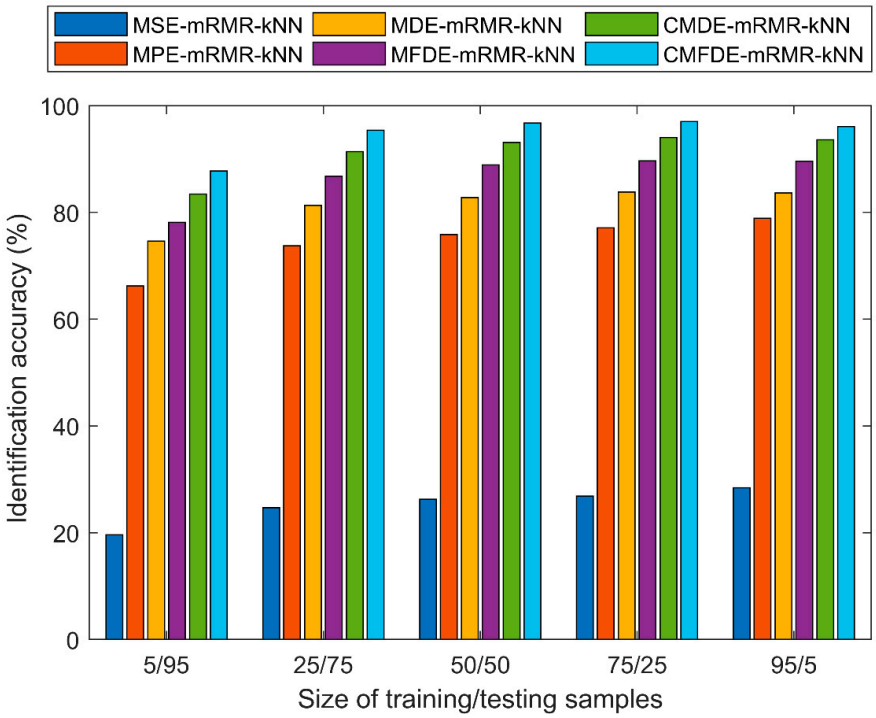


Figure 14. The average recognition accuracy under different sizes of training/testing samples.

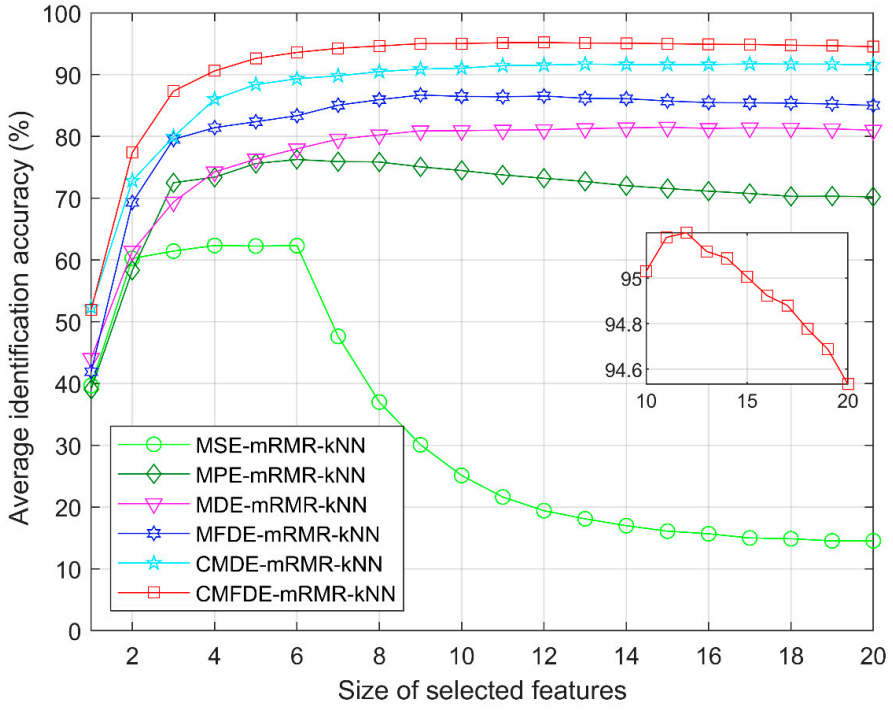


Figure 15. The average identification accuracy under different sizes of sensitive features.

5. Conclusions

In this paper, a nonlinear time series complexity evaluation method based on CFMDE was proposed. Compared with MFDE, CMFDE improved the stability of the entropy evaluation. Then, a fault diagnosis method for rolling bearings based on CMFDE, mRMR, and kNN was proposed. Through analyzing a standard experimental dataset, the effectiveness of the proposed CMFDE-mRMR-kNN method was validated. Meanwhile, the superiority of CMFDE in extracting sensitive fault features was highlighted, and the necessity of mRMR feature selection was also illustrated.

Author Contributions: X.G. implemented the algorithm, analyzed the data, and wrote the manuscript. H.L. and G.Y. investigated the project and conceived and revised the manuscript. All authors have read and approved the final manuscript.

Funding: This work has been supported by National Key Research and Development Program of China (No. 2017YFD07006 and No. 2018YFB0105300).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lou, X.; Loparo, K.A. Bearing fault diagnosis based on wavelet transform and fuzzy inference. *Mech. Syst. Signal Process.* **2004**, *18*, 1077–1095. [CrossRef]
2. Peng, Z.K.; Tse, P.W.; Chu, F.L. A comparison study of improved Hilbert–Huang transform and wavelet transform: Application to fault diagnosis for rolling bearing. *Mech. Syst. Signal Process.* **2005**, *19*, 974–988. [CrossRef]
3. Duan, Z.; Wu, T.; Guo, S.; Shao, T.; Malekian, R.; Li, Z. Development and trend of condition monitoring and fault diagnosis of multi-sensors information fusion for rolling bearings: A review. *Int. J. Adv. Manuf. Technol.* **2018**, *96*, 803–819. [CrossRef]
4. De Azevedo, H.D.M.; Araújo, A.M.; Bouchonneau, N. A review of wind turbine bearing condition monitoring: State of the art and challenges. *Renew. Sustain. Energy Rev.* **2016**, *56*, 368–379. [CrossRef]
5. Xu, Z.; Liu, K.; Zhang, H.; Wang, D.; Zhang, M.; Wu, X. A fault diagnosis method for rolling bearings based on empirical mode decomposition and principal component analysis. *J. Vib. Shock* **2014**, *33*, 133–139.
6. Lei, Y.; Lin, J.; He, Z.; Zuo, M.J. A review on empirical mode decomposition in fault diagnosis of rotating machinery. *Mech. Syst. Signal Process.* **2013**, *35*, 108–126. [CrossRef]
7. Cheng, J.; Yang, Y.; Yang, Y. A rotating machinery fault diagnosis method based on local mean decomposition. *Digit Signal Process.* **2012**, *22*, 356–366. [CrossRef]
8. Wang, Z.; Wang, J.; Kou, Y.; Zhang, J.; Ning, S.; Zhao, Z. Weak Fault Diagnosis of Wind Turbine Gearboxes Based on MED-LMD. *Entropy* **2017**, *19*, 277. [CrossRef]
9. Bo, L.; Lu, C.; Zhao, X. Fault diagnosis method for rolling bearings based on ITD and ICA. *J. Vib. Shock* **2015**, *34*, 153–156.
10. Yang, Y.; Pan, H.; Ma, L.; Cheng, J. A roller bearing fault diagnosis method based on the improved ITD and RRVPMCD. *Measurement* **2014**, *55*, 255–264. [CrossRef]
11. Basaran, C.; Nie, S. An Irreversible Thermodynamics Theory for Damage Mechanics of Solids. *Int. J. Damage Mech.* **2004**, *13*, 205–223. [CrossRef]
12. Amiri, M.; Khonsari, M.M. On the Role of Entropy Generation in Processes Involving Fatigue. *Entropy* **2012**, *14*, 24. [CrossRef]
13. Sosnovskiy, L.A.; Sherbakov, S.S. Mechanothermodynamic Entropy and Analysis of Damage State of Complex Systems. *Entropy* **2016**, *18*, 268. [CrossRef]
14. Sosnovskiy, L.A.; Sherbakov, S.S. A Model of Mechanothermodynamic Entropy in Tribology. *Entropy* **2017**, *19*, 115. [CrossRef]
15. Pincus, S. Approximate entropy (ApEn) as a complexity measure. *Chaos* **1995**, *5*, 110–117. [CrossRef]
16. Yan, R.; Gao, R.X. Approximate Entropy as a diagnostic tool for machine health monitoring. *Mech. Syst. Signal Process.* **2007**, *21*, 824–839. [CrossRef]
17. Richman, J.S.; Moorman, J.R. Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol.-Heart Circ. Physiol.* **2000**, *278*, H2039–H2049. [CrossRef]

18. Han, M.; Pan, J. A fault diagnosis method combined with LMD, sample entropy and energy ratio for roller bearings. *Measurement* **2015**, *76*, 7–19. [CrossRef]
19. Bandt, C.; Pompe, B. Permutation Entropy: A Natural Complexity Measure for Time Series. *Phys. Rev. Lett.* **2002**, *88*, 174102. [CrossRef]
20. Cheng, J.; Ma, X.; Yang, Y. Rolling bearing fault diagnosis method based on permutation entropy and VPMCD. *J. Vib. Shock* **2014**, *33*, 119–123.
21. Costa, M.; Goldberger, A.L.; Peng, C.K. Multiscale Entropy Analysis of Complex Physiologic Time Series. *Phys. Rev. Lett.* **2002**, *89*, 068102. [CrossRef]
22. Wu, S.-D.; Wu, P.-H.; Wu, C.-W.; Ding, J.-J.; Wang, C.-C. Bearing Fault Diagnosis Based on Multiscale Permutation Entropy and Support Vector Machine. *Entropy* **2012**, *14*, 1343. [CrossRef]
23. Wang, X.; Si, S.; Wei, Y.; Li, Y. The Optimized Multi-Scale Permutation Entropy and Its Application in Compound Fault Diagnosis of Rotating Machinery. *Entropy* **2019**, *21*, 170. [CrossRef]
24. Li, Y.; Wang, X.; Liu, Z.; Liang, X.; Si, S. The Entropy Algorithm and Its Variants in the Fault Diagnosis of Rotating Machinery: A Review. *IEEE Access* **2018**, *6*, 66723–66741. [CrossRef]
25. Zhang, L.; Xiong, G.; Liu, H.; Zou, H.; Guo, W. Bearing fault diagnosis using multi-scale entropy and adaptive neuro-fuzzy inference. *Expert Syst. Appl.* **2010**, *37*, 6077–6085. [CrossRef]
26. Ren, J.; Sun, G.; Chen, B.; Luo, M. Multi-scale Permutation Entropy Based On-line Milling Chatter Detection Method. *J. Mech. Eng.* **2015**, *51*, 206–212. [CrossRef]
27. Rostaghi, M.; Azami, H. Dispersion Entropy: A Measure for Time-Series Analysis. *IEEE Signal Process. Lett.* **2016**, *23*, 610–614. [CrossRef]
28. Rostaghi, M.; Ashory, M.R.; Azami, H. Application of dispersion entropy to status characterization of rotary machines. *J. Sound Vib.* **2019**, *438*, 291–308. [CrossRef]
29. Azami, H.; Escudero, J. Amplitude- and Fluctuation-Based Dispersion Entropy. *Entropy* **2018**, *20*, 210. [CrossRef]
30. Peng, H.; Long, F.; Ding, C. Feature Selection Based on Mutual Information: Criteria of Max-Dependency, Max-Relevance, and Min-Redundancy. *IEEE Trans. Pattern Anal. Mach. Intell.* **2005**, *27*, 1226–1238. [CrossRef]
31. Bearing Data Center, Case Western Reserve University. Available online: http://csegroups.case.edu/bearingdatacenter/pages/download-data-file (accessed on 13 March 2019).



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).