

Article

# Rolling Bearing Diagnosis Based on Composite Multiscale Weighted Permutation Entropy

Xiong Gan <sup>1</sup>, Hong Lu <sup>1</sup>, Guangyou Yang <sup>2,\*</sup> and Jing Liu <sup>2</sup>

<sup>1</sup> School of Mechanical and Electronic Engineering, Wuhan University of Technology, Wuhan 430070, China; ganxiong001@126.com (X.G.); landzh@whut.edu.cn (H.L.)

<sup>2</sup> Institute of Agricultural Machinery, Hubei University of Technology, Wuhan 430068, China; liujing\_hbut@126.com

\* Correspondence: pekka@126.com; Tel.: +86-027-59750857

Received: 30 September 2018; Accepted: 22 October 2018; Published: 24 October 2018



**Abstract:** In this paper, composite multiscale weighted permutation entropy (CMWPE) is proposed to evaluate the complexity of nonlinear time series, and the advantage of the CMWPE method is verified through analyzing the simulated signal. Meanwhile, considering the complex nonlinear dynamic characteristics of fault rolling bearing signal, a rolling bearing fault diagnosis approach based on CMWPE, joint mutual information (JMI) feature selection, and k-nearest-neighbor (KNN) classifier (CMWPE-JMI-KNN) is proposed. For CMWPE-JMI-KNN, CMWPE is utilized to extract the fault rolling bearing features, JMI is applied for sensitive features selection, and KNN classifier is employed for identifying different rolling bearing conditions. Finally, the proposed CMWPE-JMI-KNN approach is used to analyze the experimental dataset, the analysis results indicate the proposed approach could effectively identify different fault rolling bearing conditions.

**Keywords:** CMWPE; rolling bearing; fault diagnosis

## 1. Introduction

Rolling bearings are one of the most vulnerable parts in mechanical equipment, the working condition of rolling bearing has great influence on the reliability of mechanical system. Therefore, it is valuable to monitor and diagnose the rolling bearing [1].

Recently, various rolling bearing fault diagnosis approaches have been proposed. Non-adaptive time-frequency signal analysis techniques including Wigner–Ville Distribution (WVD) [2], Short-Time Fourier Transformation (STFT) [3], and wavelet-transform (WT) [4], which are widely used to extract feature in fault bearing diagnosis. However, WVD has the problem of cross interference, STFT is merely a single-resolution signal analysis method by using a fixed short-term window function, and WT is difficult to choose suitable wavelet bases. Meanwhile, adaptive time-frequency signal analysis techniques such as Empirical Mode Decomposition (EMD) [5,6], Local Mean Decomposition (LMD) [7,8], and Intrinsic Time-Scale Decomposition (ITD) [9,10] are extensively employed in fault bearing feature extraction. However, these methods generally have the disadvantages of envelope error, mode mixing, and end effect. Furthermore, fault bearing diagnosis method based on image processing [11] has been proposed. Nevertheless, sound data are susceptible to noise from surrounding equipment and environment, the image processing method must convert the vibration data into image data, instead of analyzing the vibration signal directly.

Mechanical systems usually have nonlinear dynamic models due to the instantaneous change of friction, load conditions, friction, and stiffness. For a mechanical condition monitoring system, the vibration signal acquired by the sensor reflects the relevant characteristics of the mechanical system. Different rolling bearing failures cause different mechanical system responses, the vibration

signal exhibits complex nonlinear characteristics [12,13]. Lots of nonlinear dynamic approaches such as approximate entropy (APE) [14,15], sample entropy (SE) [16,17], and multiscale sample entropy (MSE) [18,19] have been applied to extract rolling bearing features and identify different rolling bearing conditions.

Bandt et al. [20] first proposed the time series complexity analysis method called permutation entropy (PE), which has higher calculation efficiency and stronger anti-noise ability, compared with APE and SE. PE was adopted for feature extraction of rolling bearing in [21], which indicated that PE could stand for the randomness and dynamic change of vibration signal. In order to evaluate the complexity of time series within different scale factors, Li et al. [22] proposed multiscale permutation entropy (MPE). MPE was also applied to extract fault features and identify different conditions for rolling bearing in [23]. However, PE and MPE ignore the amplitude difference between the same permutation patterns and do not include the amplitude information of the time series. To solve this problem, Fadlallah et al. [24] proposed weight permutation entropy (WPE), which demonstrated better performance in quantifying the information complexity. Yin et al. [25] combined the WPE and multiscale analysis, and proposed the multiscale weighted permutation entropy (MWPE). However, as the scale factor increases, the length of the time series becomes shorter, which results in a sudden change for MWPE [26,27]. In order to improve the statistical reliability, composite multiscale weighted permutation entropy (CMWPE) is put forward in this paper. CMWPE averages the weighted permutation entropy within multiple scale factors, which improves the reliability of entropy estimation. Then, we adopt CMWPE for extracting rolling bearing features. However, not all CMWPE values are closely related to fault information. Joint mutual information (JMI) is applied for selecting the sensitive CMWPE features. In order to intelligently identify different rolling bearing conditions,  $k$ -nearest-neighbor (KNN) classifier is employed. Therefore, a novel fault diagnosis approach based on CMWPE, JMI, and KNN (simplified into CMWPE-JMI-KNN) is proposed and used to analyze the standard bearing dataset. The analysis results validate the effectiveness of the proposed CMWPE-JMI-KNN method in identifying different fault rolling bearing conditions.

This paper is organized as follows. The MPE, MWPE, and CMWPE are introduced in Section 2, and the superiority of the CMWPE method is also verified in this part by analyzing the simulated signal. In Section 3, the JMI feature selection method is illustrated, and the CMWPE-JMI-KNN method is proposed. In Section 4, experimental validation is presented. In Section 5, conclusions are drawn.

## 2. MPE, MWPE, CMWPE

### 2.1. MPE

#### 2.1.1. PE

Input: Time series  $X = \{x(1), x(2), \dots, x(N)\}$ , embedding dimension  $m$ , time delay  $\tau$

Output:  $PE(X, m, \tau)$

Step 1. For embedding dimension  $m$ , time delay  $\tau$ , the time series  $X$  can be reconstructed in phase space as  $X^{m,\tau} = \{X^{m,\tau}(1), X^{m,\tau}(2), \dots, X^{m,\tau}(k), \dots, X^{m,\tau}(N - (m - 1)\tau)\}$ ,  $X^{m,\tau}(k)$  can be expressed as

$$X^{m,\tau}(k) = \{x(k), x(k + \tau), \dots, x(k + (m - 1)\tau)\} \quad (1)$$

where  $k = 1, 2, \dots, N - (m - 1)\tau$ .

Step 2. Rearrange  $\{x(k), x(k + \tau), \dots, x(k + (m - 1)\tau)\}$  in an increasing order as  $\{x(k + (v_1 - 1)\tau) \leq x(k + (v_2 - 1)\tau) \leq \dots \leq x(k + (v_m - 1)\tau)\}$ , and obtain the symbol index sequence  $\pi_l^{m,\tau} = [v_1, v_2, \dots, v_m]$ ,  $\pi_l^{m,\tau}$  is one of the  $m!$  distinct symbols  $\{\pi_l^{m,\tau}\}_{l=1}^{m!}$ .

Step 3. Define  $p(\pi_l^{m,\tau})$  as

$$p(\pi_l^{m,\tau}) = \frac{\|\{k : k \leq N - (m - 1)\tau, \text{type}(X^{m,\tau}(k)) = \pi_l^{m,\tau}\}\|}{N - (m - 1)\tau} \quad (2)$$

where  $\text{type}(\cdot)$  represents the map from pattern space to symbol space,  $\|\cdot\|$  represents the cardinality of a set.

$p(\pi_l^{m,\tau})$  can be also expressed as

$$p(\pi_l^{m,\tau}) = \frac{\sum_{k=1}^{N-(m-1)\tau} 1_{\text{u:type(u)=}\pi_l^{m,\tau}}(X^{m,\tau}(k))}{\sum_{k=1}^{N-(m-1)\tau} 1_{\text{u:type(u)=}\Pi}(X^{m,\tau}(k))} \tag{3}$$

where  $1_A(u) = \begin{cases} 1, & \text{if } u \in A \\ 0, & \text{if } u \notin A \end{cases}$ ,  $\Pi = \{\pi_l^{m,\tau}\}_{l=1}^{m!}$ .

Step 4.  $PE(X, m, \tau)$  can be expressed as

$$PE(X, m, \tau) = -\sum_{l:\pi_l^{m,\tau} \in \Pi} p(\pi_l^{m,\tau}) \ln(p(\pi_l^{m,\tau})) \tag{4}$$

### 2.1.2. MPE

Input: Time series  $X = \{x(1), x(2), \dots, x(N)\}$ , embedding dimension  $m$ , time delay  $\tau$ , largest scale factor  $s_{max}$

Output:  $MPE$

Initialization:  $MPE = \emptyset, s = 1$

Step 1. For  $X = \{x(1), x(2), \dots, x(N)\}$ , the coarse-grained time series  $y^s = \{y^s(j)\}_{j=1}^{\lfloor N/s \rfloor}$  can be expressed as

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

where  $j = 1, 2, \dots, \lfloor N/s \rfloor$ ,  $\lfloor N/s \rfloor$  is the largest positive integer no more than  $N/s$ .

Step 2.  $MPE(X, m, \tau, s)$  can be expressed as

$$MPE(X, m, \tau, s) = PE(y^s, m, \tau) \tag{6}$$

Step 3.  $MPE = MPE \cup MPE(X, m, \tau, s), s = s + 1$ . Repeat steps 1–3 until  $s$  is larger than  $s_{max}$ .

## 2.2. MWPE

### 2.2.1. WPE

Input: Time series  $X = \{x(1), x(2), \dots, x(N)\}$ , embedding dimension  $m$ , time delay  $\tau$

Output:  $WPE(X, m, \tau)$

Step 1. For embedding dimension  $m$ , time delay  $\tau$ , the time series  $X$  can be reconstructed in phase space as  $X^{m,\tau} = \{X^{m,\tau}(1), X^{m,\tau}(2), \dots, X^{m,\tau}(k), \dots, X^{m,\tau}(N - (m - 1)\tau)\}$ ,  $X^{m,\tau}(k)$  can be expressed as

$$X^{m,\tau}(k) = \{x(k), x(k + \tau), \dots, x(k + (m - 1)\tau)\} \tag{7}$$

where  $k = 1, 2, \dots, N - (m - 1)\tau$ .

Step 2. Rearrange  $\{x(k), x(k + \tau), \dots, x(k + (m - 1)\tau)\}$  in an increasing order as  $\{x(k + (v_1 - 1)\tau) \leq x(k + (v_2 - 1)\tau) \leq \dots \leq x(k + (v_m - 1)\tau)\}$ , and obtain the symbol index sequence  $\pi_l^{m,\tau} = [v_1, v_2, \dots, v_m], \pi_l^{m,\tau}$  is one of the  $m!$  distinct symbols  $\{\pi_l^{m,\tau}\}_{l=1}^{m!}$ .

Step 3. Define  $p_w(\pi_l^{m,\tau})$  as

$$p_w(\pi_l^{m,\tau}) = \frac{\sum_{k=1}^{N-(m-1)\tau} 1_{\text{u:type(u)=}\pi_l^{m,\tau}}(X^{m,\tau}(k))w_k}{\sum_{k=1}^{N-(m-1)\tau} 1_{\text{u:type(u)=}\Pi}(X^{m,\tau}(k))w_k} \tag{8}$$

where  $w_k = \frac{1}{m} \sum_{q=1}^m [x(k + (q - 1)\tau) - \bar{X}^{m,\tau}(k)]^2$ ,  $\bar{X}^{m,\tau}(k) = \frac{1}{m} \sum_{q=1}^m [x(k + (q - 1)\tau)]$ .

Step 4.  $WPE(X, m, \tau)$  can be expressed as

$$WPE(X, m, \tau) = - \sum_{l: \pi_l^{m,\tau} \in \Pi} p_w(\pi_l^{m,\tau}) \ln(p_w(\pi_l^{m,\tau})) \tag{9}$$

### 2.2.2. MWPE

Input: Time series  $X = \{x(1), x(2), \dots, x(N)\}$ , embedding dimension  $m$ , time delay  $\tau$ , largest scale factor  $s_{max}$

Output:  $MWPE$

Initialization:  $MWPE = \emptyset, s = 1$

Step 1. For  $X = \{x(1), x(2), \dots, x(N)\}$ , the coarse-grained time series  $y^s = \{y^s(j)\}_{j=1}^{\lfloor N/s \rfloor}$  can be expressed as

$$y^s(j) = \frac{1}{s} \sum_{i=(j-1)s+1}^{js} x(i) \tag{10}$$

where  $j = 1, 2, \dots, \lfloor N/s \rfloor$ .

Step 2.  $MWPE(X, m, \tau, s)$  can be expressed as

$$MWPE(X, m, \tau, s) = WPE(y^s, m, \tau) \tag{11}$$

Step 3.  $MWPE = MWPE \cup MWPE(X, m, \tau, s), s = s + 1$ . Repeat steps 1–3 until  $s$  is larger than  $s_{max}$ .

### 2.3. CMWPE

Input: Time series  $X = \{x(1), x(2), \dots, x(N)\}$ , embedding dimension  $m$ , time delay  $\tau$ , largest scale factor  $s_{max}$

Initialization:  $CMWPE = \emptyset, s = 1$

Step 1. For  $X = \{x(1), x(2), \dots, x(N)\}$ , the coarse-grained time series  $y^{s,q} = \{y^{s,q}(j)\}_{j=1}^{\lfloor (N+1)/s \rfloor - 1}$  can be expressed as

$$y^{s,q}(j) = \frac{1}{s} \sum_{i=(j-1)s+q}^{js+q-1} x(i) \tag{12}$$

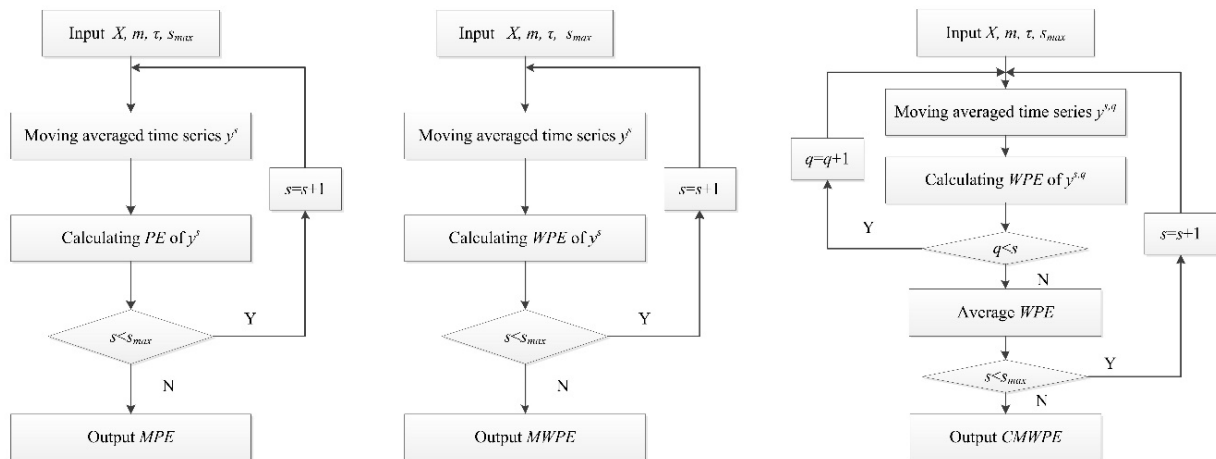
where  $j = 1, 2, \dots, \lfloor (N + 1)/s \rfloor - 1, q = 1, 2, \dots, s$ .

Step 2.  $CMWPE$  averages the  $WPE$  values,  $CMWPE(X, m, \tau, s)$  can be expressed as

$$CMWPE(X, m, \tau, s) = \frac{1}{s} \sum_{q=1}^s WPE(y^{s,q}, m, \tau) \tag{13}$$

Step 3.  $CMWPE = CMWPE \cup CMWPE(X, m, \tau, s), s = s + 1$ . Repeat steps 1–3 until  $s$  is larger than  $s_{max}$ .

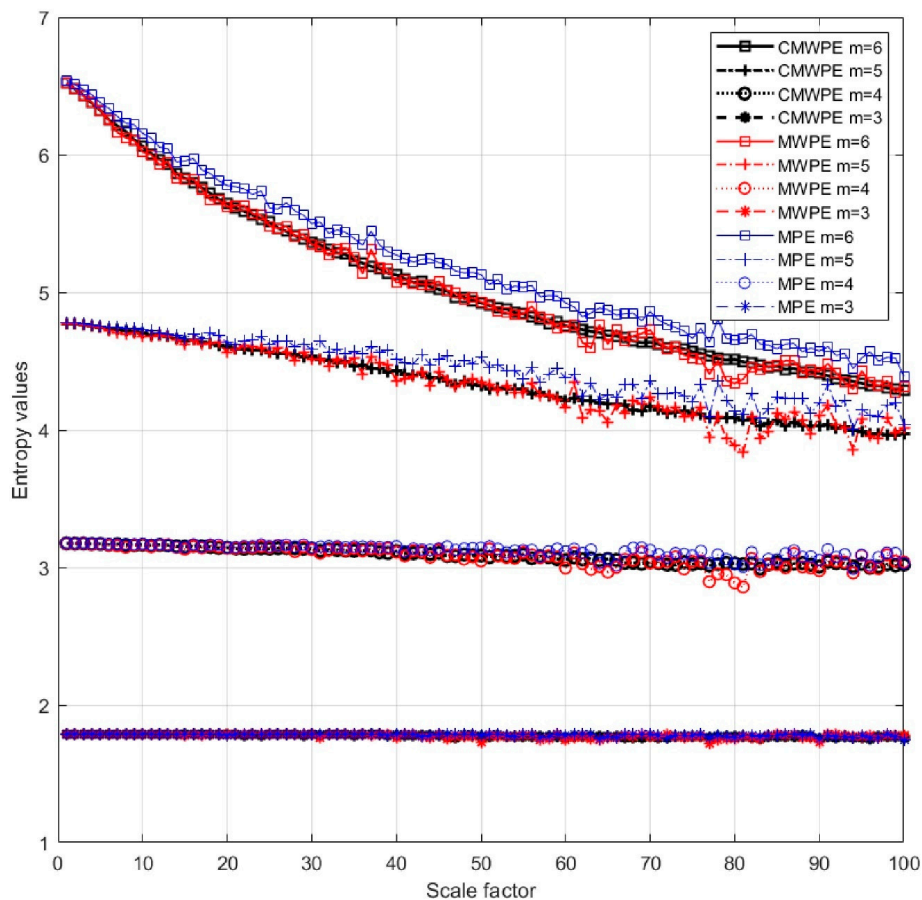
Figure 1 shows the flowchart of the MPE, MWPE, and CMWPE.



**Figure 1.** The flowchart of the multiscale permutation entropy (MPE), multiscale weighted permutation entropy (MWPE), and composite multiscale weighted permutation entropy (CMWPE).

2.4. Comparisons between MPE, MWPE, CMWPE

To prove the advantage of the CMWPE method, the white noise with 10,000 data points are generated, and we analyze the MPE, MWPE, and CMWPE results on white noise. According to the previous reports [25,28], we select embedding dimension  $m = 3, 4, 5, 6$ , time delay  $\tau = 1$ , the largest scale factor  $s_{max} = 100$ . The results of MPE, MWPE, and CMWPE on white noise are shown in Figure 2.



**Figure 2.** MPE, MWPE, and CMWPE values of the white noise.

As shown, all entropy values increase with the increasing embedding dimension  $m$ . It is noticeable that, when  $m = 3, 4$ , with the increase of scale factor, the declines of MPE, MWPE, and CMWPE are not evident, and the superiority of multiscale analysis cannot be displayed effectively. When  $m = 5, 6$ , with the increasing scale factor, MPE, MWPE and CMWPE all show a significant decreasing trend. In addition, the values of CMWPE and WMPE decrease faster than that of MPE, indicating CMWPE and MWPE can be more sensitive to extract the time series complexity including amplitude information. Furthermore, compared with MWPE, CMWPE can reduce the fluctuation and standard deviation, which demonstrates its superiority. In general, CMWMP not only measures the complexity of time series incorporating amplitude information within multiple scales, but also improves the reliability of entropy evaluation.

### 3. Fault Diagnosis Approach Based on CMWPE, JMI, and KNN

Due to friction, load, impact, and signal transmission interference in mechanical system, not all the CMWPE features are closely related to fault information. If all the CMWPE features are input into the classifier as fault features for identifying bearing condition, the recognition accuracy and efficiency may be decrease. Therefore, sensitive features need to be selected, and CMWPE-JMI-KNN method is proposed in this paper. For CMWPE-JMI-KNN method, JMI algorithm [29,30] is used to select sensitive features, and KNN classifier is applied to identify different fault rolling bearing conditions.

#### 3.1. JMI Feature Selection

Give original features set  $X = \{x_1, x_2, \dots, x_s\}$ , label information  $C$ , the size of sensitive features subset  $p$ , JMI approach is employed to optimize sensitive features subset  $Y = \{y_1, y_2, \dots, y_p\}$ . JMI algorithm selects the first sensitive feature  $y_1$  based on the maximum mutual information according to the following expression

$$y_1 = \operatorname{argmax}_{i=1,2,\dots,s} \{I(x_i; C)\} \quad (14)$$

where  $I(x_i; C)$  represents the mutual information of  $x_i$  and  $C$ .

Then, the sequential forward search algorithm is adopted to obtain the sensitive features subset  $Y$  until the subset size  $|Y|$  is equal to  $p$ . Suppose that  $k$  features have been selected, choose the  $k + 1$  feature  $y_{k+1}$  by the following expression

$$y_{k+1} = \operatorname{argmax}_{x_i \in X-Y} \sum_{j=1}^k I(x_i; C|y_j) \quad (15)$$

where  $I(x_i; C|y_j)$  represents the condition mutual information of  $x_i$  and  $C$  under given  $y_j$ .

Input: Original features set  $X = \{x_1, x_2, \dots, x_s\}$ , label information  $C$ , the size of sensitive features subset  $p$

Output: Sensitive features subset  $Y = \{y_1, y_2, \dots, y_p\}$

Initialization:  $Y = \emptyset, k = 0$

Step 1. Select the first sensitive feature  $y_1$  based on the maximum mutual information,  $y_1 = \operatorname{argmax}_{i=1,2,\dots,s} \{I(x_i; C)\}, Y = \{y_1\}, k = 1$ .

Step 2. Based on  $y_{k+1} = \operatorname{argmax}_{x_i \in X-Y} \sum_{j=1}^k I(x_i; C|y_j)$  to select the  $k + 1$  sensitive feature  $y_{k+1}$ ,  $Y = Y \cup \{y_{k+1}\}, k = k + 1$ . Repeat step 2 until the size of sensitive features subset  $|Y|$  is equal to  $p$ .

#### 3.2. CMWPE-JMI-KNN

Based on the advantages of CMWPE, JMI, and KNN, the proposed CMWPE-JMI-KNN approach can be described as follows.

Input: Training samples, training label  $L_{train}$ , testing samples, embedding dimension  $m$ , time delay  $\tau$ , maximum scale factor  $s_{max}$ , the number of sensitive features  $p$

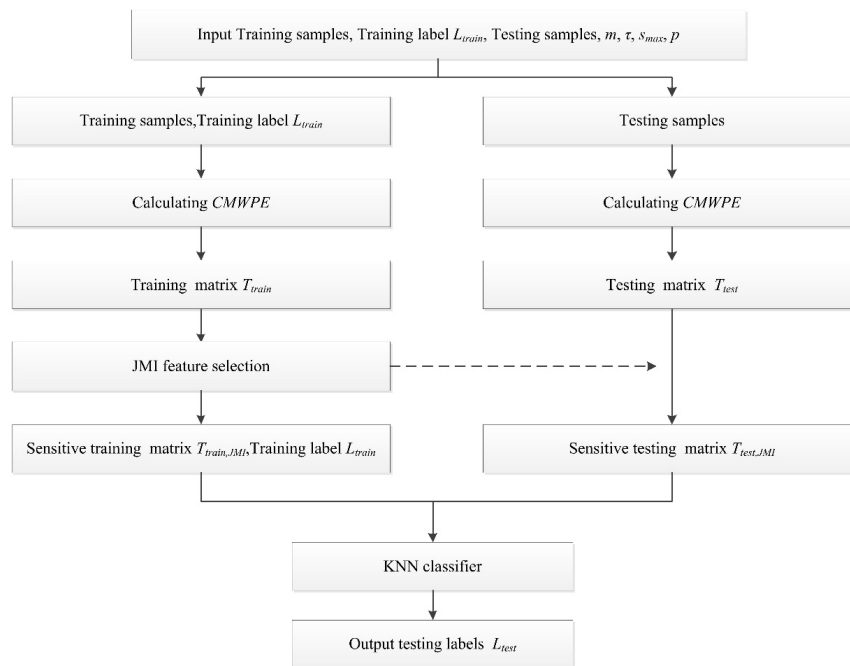
Output: Testing label  $L_{test}$

Step 1. Calculating CMWPE of the training samples to form training matrix  $T_{train}$ , adopting JMI method to select the first  $p$  features to construct into sensitive low-dimension matrix  $T_{train,JMI}$ .

Step 2. Calculating CMWPE of the testing samples to form testing matrix  $T_{test}$ , forming the sensitive low-dimension matrix  $T_{test,JMI}$  based on the selected results in training samples.

Step 3. Input  $T_{train,JMI}$ ,  $T_{test,JMI}$ ,  $L_{train}$  to the KNN classifier for identifying the rolling bearing conditions of testing samples, output the testing label  $L_{test}$ .

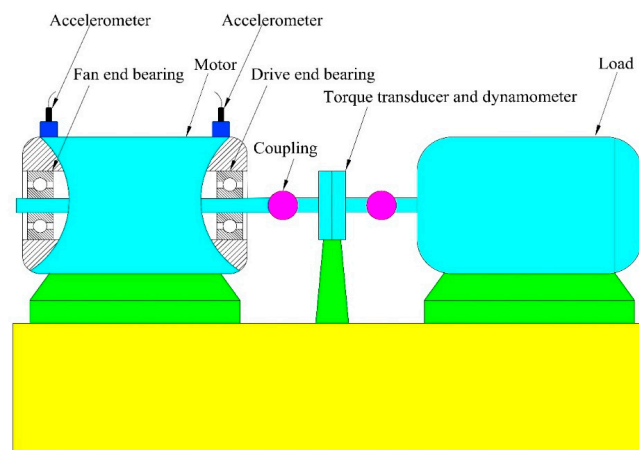
Figure 3 shows the flowchart of the proposed CMWPE-JMI-KNN method.



**Figure 3.** Flowchart of the proposed CMWPE-JMI-KNN method. JMI, joint mutual information; KNN,  $k$ -nearest neighbor.

#### 4. Experimental Validation

Standard experiment dataset is provided by Rolling Bearing Data Center of Case Western Reserve University [31]. The proposed CMWPE-JMI-KNN method is applied to analyze the dataset for verifying its effectiveness. Figure 4 shows the schematic diagram of bearing test bench, the experiment uses 6205-2RS JEM SKF bearing and adopts accelerometer to collect the vibration signal.



**Figure 4.** Schematic diagram of bearing test bench.



The standard dataset contains vibration data of the drive end bearing under ten conditions, including normal condition (noted into Normal), inner race fault conditions with fault size 0.1778 mm, 0.3556 mm, 0.5334 mm (noted into IRF1, IRF2, IRF3), ball fault conditions with fault size 0.1778 mm, 0.3556 mm, 0.5334 mm (noted into BF1, BF2, BF3), outer race fault conditions with fault size 0.1778 mm, 0.3556 mm, 0.5334 mm (noted into ORF1, ORF2, ORF3). Figure 5 shows the vibration data under ten conditions. The sampling frequency is 12 kHz, the motor speed is 1772 rpm, and the load is 0 HP. Due to friction, load, shock, and noise interference, it is hard to identify concrete bearing condition from the time domain waveform.

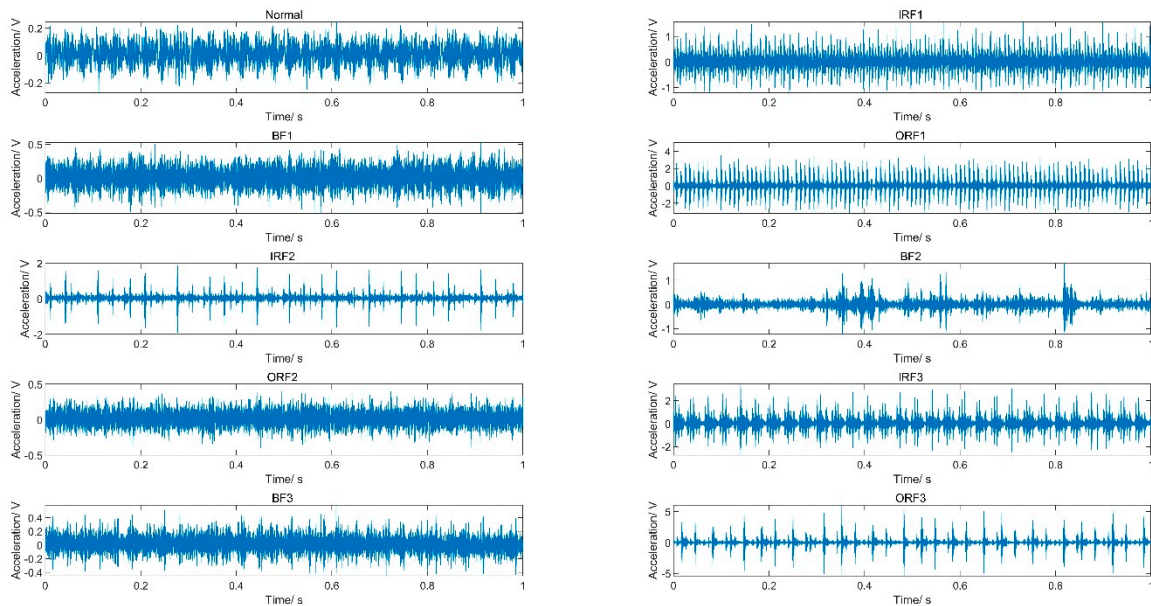


Figure 5. Vibration data under 10 conditions.

The vibration data are divided into multiple data samples without overlap. Each condition has 110 samples, and every sample contains  $N = 1024$  points. Twenty-two randomly selected samples in each condition were used to form the training samples, and the remaining 88 samples were used as testing samples. The detailed information of the training samples and testing samples are shown in Table 1. Here, the embedding dimensions  $m = 3, 4, 5, 6$ , time delay  $\tau = 1$ , and maximum scale factor  $s_{max} = 100$ , sensitive features size  $p = 10$ . The proposed CMWPE-JMI-KNN method is applied to analysis the data, which can be described as follows.

Step 1. Calculate CMWPE of the training samples to form training matrix  $T_{train} \in \mathbb{R}^{22 \times 1000}$ . Sort 100 features using JMI, and select the first 10 features to construct into sensitive low-dimension matrix  $T_{train,JMI} \in \mathbb{R}^{22 \times 100}$ .

Step 2. Calculate CMWPE of the testing samples to form testing matrix  $T_{test} \in \mathbb{R}^{88 \times 1000}$ , and obtain the sensitive low-dimension matrix  $T_{test,JMI} \in \mathbb{R}^{88 \times 100}$  based on the selected results in the training samples.

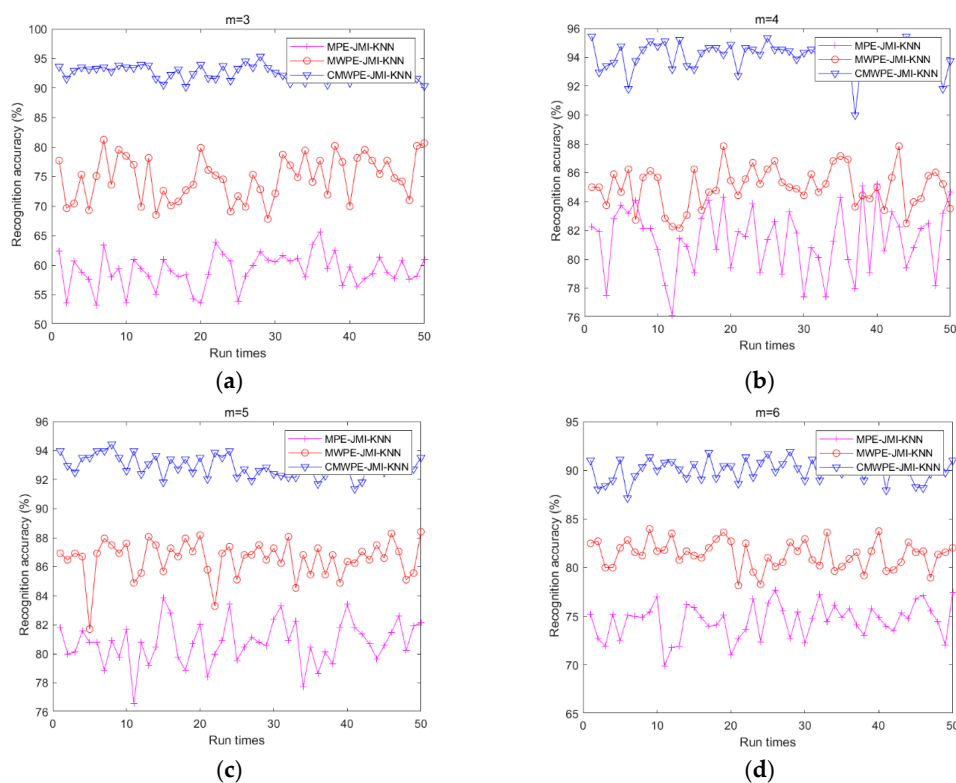
Step 3. Input  $T_{train,JMI}$ ,  $T_{test,JMI}$  and  $L_{train}$  to the KNN classifier ( $k = 1$ ) for identifying different conditions of the testing samples, output the testing label  $L_{test}$ .



**Table 1.** The detailed information of the training samples and testing samples.

| Bearing Condition | Fault Diameter (mm) | Motor Load (HP) | Motor Speed (rpm) | Label | Number of Training Samples | Number of Testing Samples |
|-------------------|---------------------|-----------------|-------------------|-------|----------------------------|---------------------------|
| Normal            | 0                   | 0               | 1772              | 1     | 22                         | 88                        |
| IRF1              | 0.1778              | 0               | 1772              | 2     | 22                         | 88                        |
| BF1               | 0.1778              | 0               | 1772              | 3     | 22                         | 88                        |
| ORF1              | 0.1778              | 0               | 1772              | 4     | 22                         | 88                        |
| IRF2              | 0.3556              | 0               | 1772              | 5     | 22                         | 88                        |
| BF2               | 0.3556              | 0               | 1772              | 6     | 22                         | 88                        |
| ORF2              | 0.3556              | 0               | 1772              | 7     | 22                         | 88                        |
| IRF3              | 0.5334              | 0               | 1772              | 8     | 22                         | 88                        |
| BF3               | 0.5334              | 0               | 1772              | 9     | 22                         | 88                        |
| ORF3              | 0.5334              | 0               | 1772              | 10    | 22                         | 88                        |

To prove the effectiveness of the proposed CMWPE-JMI-KNN method and the advantage of CMWPE, CMWPE is substituted with MPE and MWPE in CMWPE-JMI-KNN to extract the fault bearing features. Similarly, JMI is employed to select sensitive features and KNN classifier is applied for recognizing the different bearing conditions. Under the same experimental settings, all the three methods were conducted with 50 run times for comparison. The recognition results are shown in Figure 6 and Table 2. As demonstrated, for  $m = 3, 4, 5, 6$ , CMWPE-JMI-KNN is superior to MWPE-JMI-KNN and MPE-JMI-KNN in recognition ability. Especially, the maximum recognition accuracy of CMWPE-JMI-KNN reaches to 95.45% with  $m = 4$ . In addition, the recognition accuracy of MWPE-JMI-KNN method is less than that of CMWPE-JMI-KNN, which obtains the moderate recognition accuracy. Furthermore, the MPE-JMI-KNN method has the worst recognition accuracy. The results validate the effectiveness of the proposed CMWPE-JMI-KNN method in recognizing different fault conditions, and the advantage of CMWPE in extracting fault bearing features.



**Figure 6.** Recognition accuracy of the three methods. (a)  $m = 3$ ; (b)  $m = 4$ ; (c)  $m = 5$ ; (d)  $m = 6$ .

Table 2. Recognition accuracy of the three methods.

| Experiments | CMWPE-JMI-KNN |       |       | MWPE-JMI-KNN |       |       | MPE-JMI-KNN  |       |       |
|-------------|---------------|-------|-------|--------------|-------|-------|--------------|-------|-------|
|             | Accuracy (%)  |       |       | Accuracy (%) |       |       | Accuracy (%) |       |       |
|             | Max           | Min   | Mean  | Max          | Min   | Mean  | Max          | Min   | Mean  |
| $m = 3$     | 95.34         | 90.22 | 92.70 | 81.25        | 67.84 | 74.79 | 65.68        | 53.18 | 59.13 |
| $m = 4$     | 95.45         | 90.00 | 94.11 | 87.84        | 82.15 | 85.00 | 85.22        | 76.13 | 81.35 |
| $m = 5$     | 94.43         | 91.36 | 93.01 | 88.40        | 81.70 | 86.56 | 83.86        | 76.59 | 80.79 |
| $m = 6$     | 93.06         | 87.15 | 90.17 | 83.97        | 78.18 | 81.38 | 77.72        | 69.88 | 74.55 |

To further display the necessity of JMI feature selection, random feature selection method is utilized to replace JMI in CMWPE-JMI-KNN, MWPE-JMI-KNN, and MPE-JMI-KNN methods to randomly select 10 features. Similarly, CMWPE, MWPE, and MPE are used for extracting the bearing features, and KNN classifier is employed to recognize the different bearing conditions. For comparison, all the three methods combined with random selected features were conducted with 50 run times under the same experimental settings. The recognition results are shown in Figure 7 and Table 3. It was found that all the recognition accuracies of three methods with random feature selection are lower than those of integrated with JMI, which reveals the necessity of JMI in selecting sensitive features. Furthermore, among these three methods, CMWPE-RANDOM-KNN method achieved the highest recognition accuracy with  $m = 3, 4, 5, 6$ , which further demonstrates the advantage of the CMWPE for feature extraction.

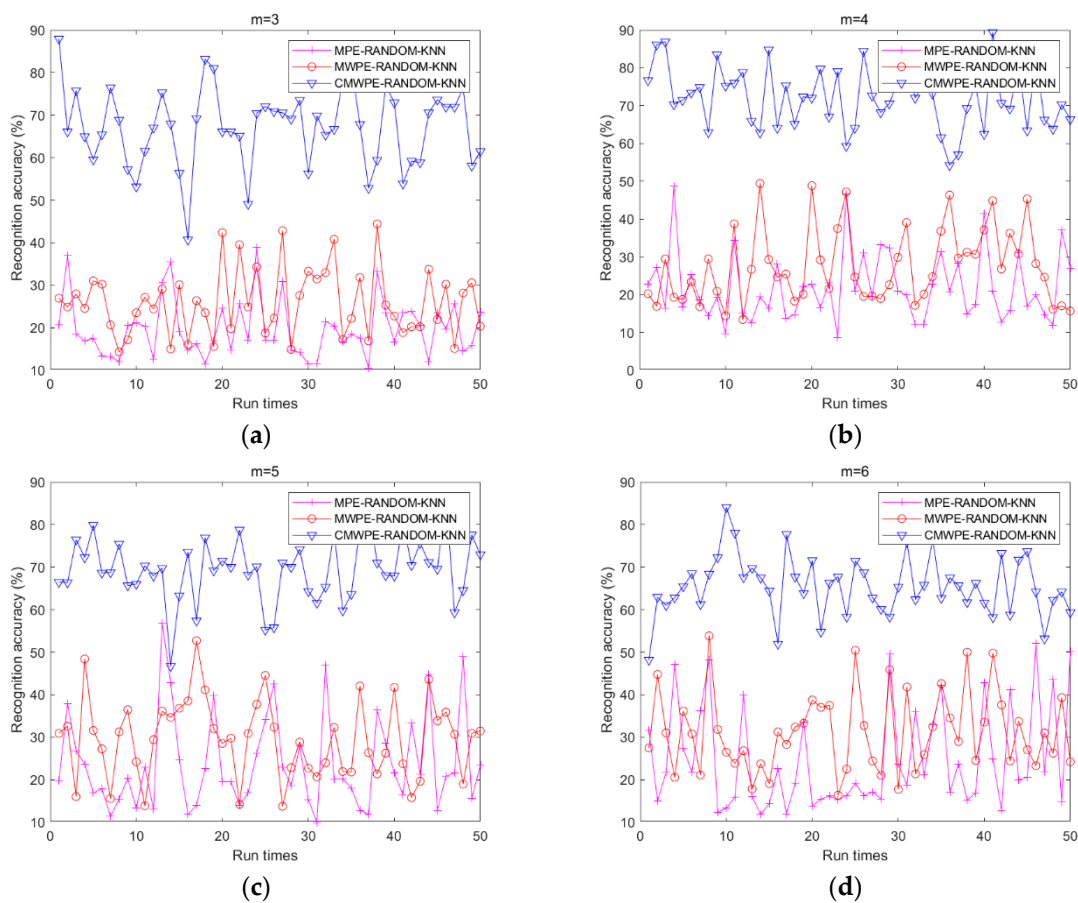
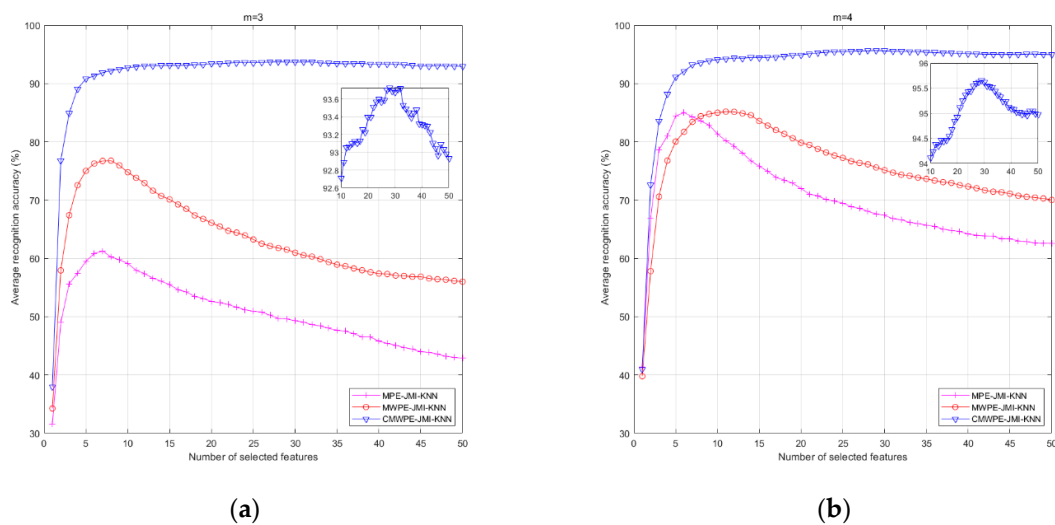


Figure 7. Recognition accuracy of the three methods with random selected features. (a)  $m = 3$ ; (b)  $m = 4$ ; (c)  $m = 5$ ; (d)  $m = 6$ .

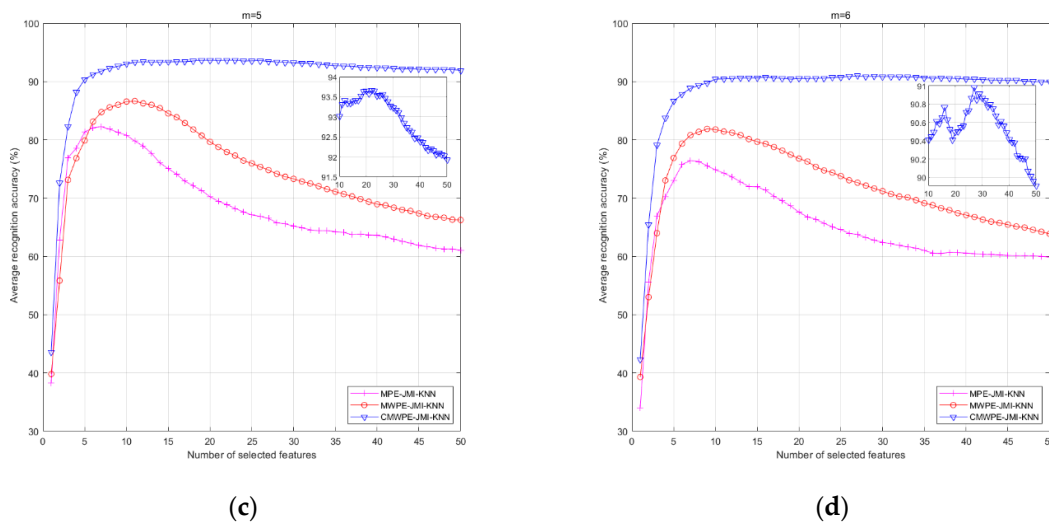
**Table 3.** Recognition accuracy of the three methods with random selected features.

| Experiments | CMWPE-RANDOM-KNN |       |       | MWPE-RANDOM-KNN |       |       | MPE-RANDOM-KNN |       |       |
|-------------|------------------|-------|-------|-----------------|-------|-------|----------------|-------|-------|
|             | Accuracy (%)     |       |       | Accuracy (%)    |       |       | Accuracy (%)   |       |       |
|             | Max              | Min   | Mean  | Max             | Min   | Mean  | Max            | Min   | Mean  |
| $m = 3$     | 87.95            | 40.68 | 67.10 | 44.43           | 14.31 | 25.88 | 38.97          | 10.45 | 19.74 |
| $m = 4$     | 89.43            | 54.20 | 72.18 | 49.43           | 13.40 | 27.48 | 48.63          | 8.63  | 22.02 |
| $m = 5$     | 86.47            | 46.70 | 69.56 | 52.72           | 13.75 | 29.57 | 56.93          | 10.00 | 23.91 |
| $m = 6$     | 84.09            | 48.18 | 65.48 | 53.86           | 16.36 | 31.15 | 52.15          | 11.81 | 24.94 |

In order to analyze the relationship between the number of selected features and recognition accuracy, all the three methods were conducted with 50 run times to obtain average recognition accuracy for each number of selected features. Figure 8 shows the corresponding average recognition accuracy for various numbers of selected features with  $m = 3, 4, 5, 6$ . It was found that, compared with MWPE-JMI-KNN and MPE-JMI-KNN, the CMWPE-JMI-KNN method achieved much higher recognition accuracy. Especially when  $m = 4$ , the proposed CMWPE-JMI-KNN method achieved the highest recognition accuracy as 95.66% when the first 29 sensitive features were selected by JMI. The analysis results verify the advantage of the proposed CMWPE-JMI-KNN method. In addition, too large or too small selected features will lead to the decline in the recognition accuracy. For the CMWPE-JMI-KNN method, when  $m = 3, 4, 5, 6$ , the optimal number of sensitive features are 28, 29, 22, and 27, respectively. The reason is that, it contains less fault information if too small selected sensitive features are selected. On the contrary, if the number of selected sensitive features is too large, it will lead to redundancy of fault information and reduce the recognition accuracy.

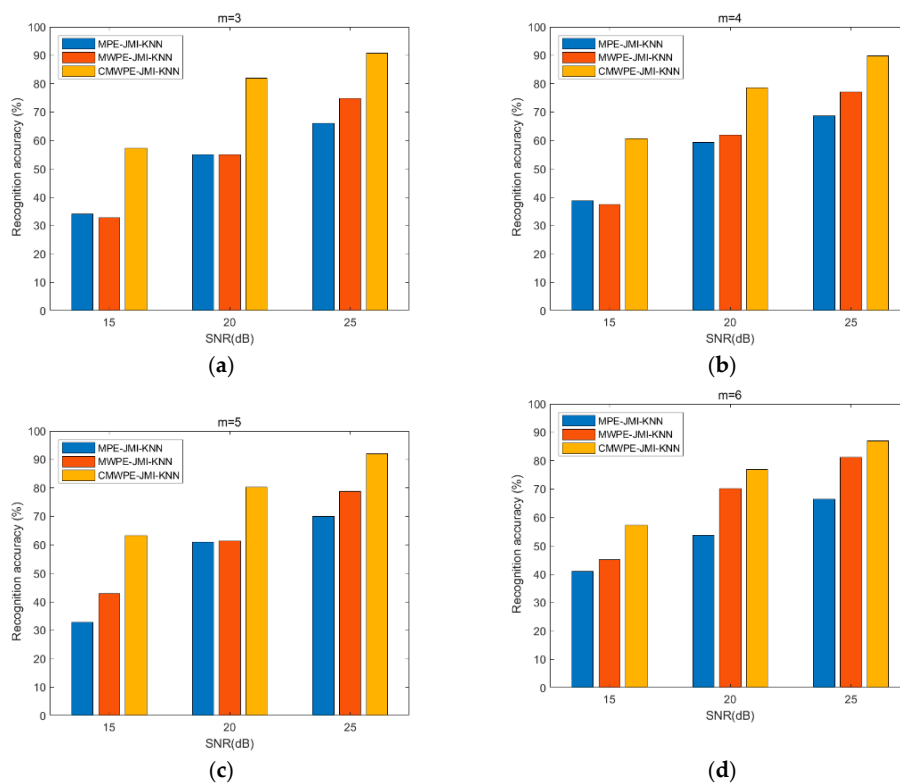


**Figure 8.** Cont.



**Figure 8.** The average recognition accuracy with various number of selected features. (a)  $m = 3$ ; (b)  $m = 4$ ; (c)  $m = 5$ ; (d)  $m = 6$ .

To test the noise robustness of the CMWPE-JMI-KNN method, random noise is added to the vibration signal of the rolling bearing. Figure 9 shows the identification results for the CMWPE-JMI-KNN, MWPE-JMI-KNN, and MPE-JMI-KNN methods under different SNR (signal-to-noise ratio) conditions. It can be seen from Figure 9 that under the same SNR condition, the fault recognition performance of the CMWPE-JMI-KNN method is better than the other two fault diagnosis methods. Moreover, the larger the SNR, the higher the identification accuracy for the three methods. Especially, when SNR = 25 dB, the proposed CMWPE-JMI-KNN method achieves the highest recognition accuracy as 90.68% with  $m = 3$ .



**Figure 9.** Recognition accuracy of the three methods under different SNR (signal-to-noise ratio) conditions. (a)  $m = 3$ ; (b)  $m = 4$ ; (c)  $m = 5$ ; (d)  $m = 6$ .

## 5. Conclusions

In this paper, a novel nonlinear dynamic approach named CMWPE is put forward, the comparison results verify the superiority of CMWPE method by analyzing the simulated signal. Based on the virtues of CMWPE, JMI, and KNN, the CMWPE-JMI-KNN approach for recognizing the fault bearing conditions is presented and employed to experimental dataset analysis. The analysis results validate the effectiveness of the proposed CMWPE-JMI-KNN approach, the advantage of CMWPE for extracting fault bearing information, and the necessity of JMI feature selection. The subsequent research will be focused on as follows:

1. Combining entropy theories and advanced signal processing techniques to further improve the recognition accuracy and anti-noise ability.
2. Applying the proposed diagnosis method to more types of mechanical fault diagnosis in real world industrial applications.

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

**Funding:** This research was funded by National Natural Science Foundation of China (No. 51275372, No. 61540027) and the National Key Research and Development Program of China (No. 2017YFD07006, No. 2018YFB0105300). The APC was funded by the National Key Research and Development Program of China (No. 2017YFD07006).

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

## References

1. Zheng, J.; Cheng, J.; Yang, Y. A rolling bearing fault diagnosis approach based on LCD and fuzzy entropy. *Mech. Mach. Theory* **2013**, *70*, 441–453. [[CrossRef](#)]
2. Gong, H.; Huang, W.; Zhao, K.; Li, S.; Zhu, Z. Time-frequency feature extraction based on fusion of Wigner-Ville distribution and wavelet scalogram. *J. Vib. Eng.* **2011**, *30*, 35–38.
3. Gao, H.; Liang, L.; Chen, X.; Xu, G. Feature extraction and recognition for rolling element bearing fault utilizing short-time Fourier transform and non-negative matrix factorization. *Chin. J. Mech. Eng.* **2015**, *28*, 96–105. [[CrossRef](#)]
4. Rodriguez, N.; Cabrera, G.; Lagos, C.; Cabrera, E. Stationary Wavelet Singular Entropy and Kernel Extreme Learning for Bearing Multi-Fault Diagnosis. *Entropy* **2017**, *19*, 541. [[CrossRef](#)]
5. 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](#)]
6. Li, Y.; Xu, M.; Wei, Y.; Huang, W. An improvement EMD method based on the optimized rational Hermite interpolation approach and its application to gear fault diagnosis. *Measurement* **2015**, *63*, 330–345. [[CrossRef](#)]
7. Zhang, Y.; Qin, Y.; Xing, Z.; Jia, L.; Cheng, X. Roller bearing safety region estimation and state identification based on LMD-PCA-LSSVM. *Measurement* **2013**, *46*, 1315–1324. [[CrossRef](#)]
8. Liu, H.; Han, M. A fault diagnosis method based on local mean decomposition and multi-scale entropy for roller bearings. *Mech. Mach. Theory* **2014**, *75*, 67–78. [[CrossRef](#)]
9. Zhang, L.; Li, P.; Li, M.; Zhang, S.; Zhang, Z. Fault diagnosis of rolling bearing based on ITD fuzzy entropy and GG clustering. *Chin. J. Sci. Instrum.* **2014**, *35*, 2624–2632.
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. Lu, C.; Wang, Y.; Ragulskis, M.; Cheng, Y. Fault Diagnosis for Rotating Machinery: A Method based on Image Processing. *PLoS ONE* **2016**, *11*, e0164111. [[CrossRef](#)] [[PubMed](#)]
12. Zhao, L.Y.; Wang, L.; Yan, R.Q. Rolling Bearing Fault Diagnosis Based on Wavelet Packet Decomposition and Multi-Scale Permutation Entropy. *Entropy* **2015**, *17*, 6447–6461. [[CrossRef](#)]
13. He, W.; Zi, Y.; Chen, B.; Wu, F.; He, Z. Automatic fault feature extraction of mechanical anomaly on induction motor bearing using ensemble super-wavelet transform. *Mech. Syst. Signal Process.* **2015**, *54–55*, 457–480. [[CrossRef](#)]

14. Pincus, S. Approximate entropy (ApEn) as a complexity measure. *Chaos* **1995**, *5*, 110–117. [[CrossRef](#)] [[PubMed](#)]
15. Li, Q.; Wang, T.Y.; Xu, Y.G.; He, H.L.; Zhang, Y. Fault diagnosis of rolling bearings based on chaos and two-dimensional approximate Entropy. *J. Vib. Eng.* **2007**, *20*, 268–273.
16. Richman, J.S.; Moorman, J.R. Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol.-Heart C* **2000**, *278*, H2039–H2049. [[CrossRef](#)] [[PubMed](#)]
17. Zhao, Z.H.; Yang, S.P. Sample entropy-based roller bearing fault diagnosis method. *J. Vib. Shock* **2012**, *31*, 136–154.
18. Costa, M.; Goldberger, A.L.; Peng, C.K. Multiscale Entropy Analysis of Complex Physiologic Time Series. *Phys. Rev. Lett.* **2007**, *89*, 705–708. [[CrossRef](#)] [[PubMed](#)]
19. Jinde, Z.; Junsheng, C.; Yang, Y. A rolling bearing fault diagnosis approach based on multiscale entropy. *J. Hum. Univ.* **2012**, *39*, 38–41.
20. Bandt, C.; Pompe, B. Permutation Entropy: A Natural Complexity Measure for Time Series. *Phys. Rev. Lett.* **2002**, *88*, 174102. [[CrossRef](#)] [[PubMed](#)]
21. Rao, G.Q.; Feng, F.Z.; Si, A.W.; Xie, J.L. Method for optimal determination of parameters in permutation entropy algorithm. *J. Vib. Shock* **2014**, *33*, 188–193.
22. Duan, L.; Xiaoli, L.; Zhenhu, L.; Logan, J.V.; Jamie, W.S. Multiscale permutation entropy analysis of EEG recordings during sevoflurane anesthesia. *J. Neural Eng.* **2010**, *7*, 046010. [[CrossRef](#)]
23. Li, Y.; Xu, M.; Wei, Y.; Huang, W. A new rolling bearing fault diagnosis method based on multiscale permutation entropy and improved support vector machine based binary tree. *Measurement* **2016**, *77*, 80–94. [[CrossRef](#)]
24. Fadlallah, B.; Chen, B.; Keil, A.; Príncipe, J. Weighted-permutation entropy: A complexity measure for time series incorporating amplitude information. *Phys. Rev. E* **2013**, *87*, 022911. [[CrossRef](#)] [[PubMed](#)]
25. Yin, Y.; Shang, P. Weighted multiscale permutation entropy of financial time series. *Nonlinear Dyn.* **2014**, *78*, 2921–2939. [[CrossRef](#)]
26. Wu, S.D.; Wu, C.W.; Lin, S.G.; Wang, C.C.; Lee, K.Y. Time Series Analysis Using Composite Multiscale Entropy. *Entropy* **2013**, *15*, 1069–1084. [[CrossRef](#)]
27. Zheng, J.; Pan, H.; Cheng, J. Rolling bearing fault detection and diagnosis based on composite multiscale fuzzy entropy and ensemble support vector machines. *Mech. Syst. Signal Process.* **2017**, *85*, 746–759. [[CrossRef](#)]
28. Yin, Y.; Shang, P. Multivariate weighted multiscale permutation entropy for complex time series. *Nonlinear Dyn.* **2017**, *88*, 1707–1722. [[CrossRef](#)]
29. Yang, H.; Moody, J. Data Visualization and Feature Selection: New Algorithms for Nongaussian Data. *Adv. Neural Inf. Process. Syst.* **2000**, *2*, 687–693.
30. Sheng, S.; Yang, H.; Wang, Y.; Pan, Y.; Tang, J. Joint mutual information feature selection for underwater acoustic targets. *J. Northwest. Polytech. Univ.* **2015**, *33*, 639–643.
31. Bearing Data Center, Case Western Reserve University. Available online: <http://csegroups.case.edu/bearingdatacenter/pages/download-data-file> (accessed on 23 October 2018).

