IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 28 June 2023, accepted 31 July 2023, date of publication 14 August 2023, date of current version 17 August 2023.
Digital Object Identifier 10.1109/ACCESS.2023.3304700

RESEARCH ARTICLE

Rolling Bearing Fault Diagnosis Under Different Severity Based on Statistics Detection Index and Canonical Discriminant Analysis

XUEWEI SONG[1,2,3], ZHIQIANG LIAO[1,2,3], BAOZHU JIA[1,2,3], DEFENG KONG[1,2,3], AND JINZHANG NIU[1,2]
[1]Naval Architecture and Shipping College, Guangdong Ocean University, Zhanjiang 524088, China
[2]Technical Research Center for Ship Intelligence and Safety Engineering of Guangdong Province, Zhanjiang 524088, China
[3]Guangdong Provincial Key Laboratory of Intelligent Equipment for South China Sea Marine Ranching, Zhanjiang 524088, China
Corresponding authors: Zhiqiang Liao (zhiqiangliao@126.com) and Baozhu Jia (jiabzh@gdou.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 52201355 and Grant 52071090, and in part by the Program for Scientific Research Start-Up Funds of Guangdong Ocean University.

ABSTRACT Bearing failures are the most frequent causes of breakdowns in rotating machinery. Different levels of severity in these failures exhibit distinct fault characteristics in the vibration signal. This paper presents a bearing fault diagnosis method that considers different severity levels, involving the selection of statistics detection index symptom parameter and the application of canonical discriminant analysis (CDA). Initially, kurtosis is employed to detect abnormalities in the bearing. Subsequently, statistical analysis theory is utilized to extract efficient symptom parameters from the time domain and frequency domain vibration signals. As a statistical analysis method, CDA can discriminate between different signals by maximizing the between-group difference and minimizing the inter-group difference. By analyzing the distribution of CDA canonical scores, bearing faults can be intuitively diagnosed. The proposed method is validated using vibration signals obtained from an experimental bench with three different bearing conditions (normal, inner race fault, outer race fault) exhibiting varying severity levels. The results demonstrate the effectiveness and feasibility of diagnosing faults under different severity levels.

INDEX TERMS Bearing fault diagnosis, canonical discriminant analysis, statistics detection index, symptom parameter selection.

I. INTRODUCTION

Rolling bearings are vital components of rotating machines, and their condition has a significant impact on production. Therefore, monitoring and diagnosing the status of bearings play a crucial role in ensuring equipment safety and minimizing accidents and losses [1], [2]. A bearing fault signal carries information not only about the overall health condition of the machine but also the severity of the fault. Accurate assessment of fault severity is essential for obtaining reliable diagnosis results [3].

In existing research on bearing fault diagnosis, the established diagnosis models are often based on a specific level of bearing fault severity. However, the accuracy of bearing fault diagnosis is often impacted when the severity of the bearing fault varies [4], [5]. Therefore, it is imperative to explore methods for diagnosing bearing faults under varying levels of severity. One of the most commonly used approaches for severity diagnosis is analyzing the energy values in the frequency or time-frequency domains of vibration signals [3], [6]. There are some papers focused on the fault severity assessment. Reference [2] analyzed the Lempel-Ziv complexity used to evaluate fault deterioration in [7], and proposed a new version of the Lempel-Ziv complexity for bearing fault severity assessment. The new method jointly applied with the continuous wavelet transform to identify the best scale for diagnosis based on Lempel-Ziv complexity. Aditya Sharma etc. in [8] employs machine

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

learning techniques support vector machine (SVM) and artificial neural network (ANN) for estimating the severity of the fault. The proposed methodology used the most appropriate features to train and diagnose the SVM and ANN model. According the output result, the severity of faults can be analyzed. In the work by Pacheco [9], Feature selection by attribute clustering using rough set theory was successfully applied to the fault severity classification. In [10], a fault identification and severity assessment of rolling element bearings was proposed based on EMD and fast kurtogram. The fault severity assessment is according to the energy percentage of the most fault-related IMF.

Reviewing related studies reveals that many methods for diagnosing faults under different severity levels rely on comparing different datasets [11], such as vibration signals from the healthy condition, slight fault, medium fault, and severe fault. However, several difficulties arise in the process. Firstly, a fault of the same severity may exhibit varying signal strengths or magnitudes in different operating conditions or bearing sizes. Secondly, different gain settings used during data measurement can lead to significant differences in signal magnitudes. Therefore, relying solely on signal "strength" for fault assessment without knowledge of the gain values can yield misleading diagnosis results. Thirdly, in the case of new machines and new working conditions, the availability of measured signal sets may be limited. Considering the aforementioned challenges, this paper proposes a bearing fault diagnosis method under different severity levels.

To determine the normality of the bearing, a commonly used noise cancellation technique, such as high-pass filtering, is employed. Following this, kurtosis is utilized to detect any abnormalities in the bearing. Subsequently, the filtered signal undergoes parameter extraction in the time and frequency domains, utilizing statistics detection index. The extracted symptom parameters are then subjected to the canonical discriminant analysis (CDA) algorithm to obtain canonical scores. These scores enable the presentation of fault condition distributions. By analyzing the distribution of CDA canonical scores, the diagnosis of bearing faults under different severity levels is performed.

The innovation and merits of the proposed method can be summarized as follows:

(1) The proposed method eliminates the need for manually selecting object properties and instead automatically selects dominant parameters using statistics detection index.

(2) The developed diagnostic model can effectively diagnose faults of different severity levels.

(3) The proposed method utilizes CDA to fuse the performance of feature parameters for intelligent fault diagnosis of rolling bearings.

The structure of this paper is as follows: Section II introduces the method and theory of bearing fault diagnosis under different severity proposed in this paper, including the introduction of statistical detection indices and CDA. In Section III, experimental data from various fault severities (slight, medium, and severe) of inner race, outer race, and roller faults under different rotational speeds (500 RPM, 1000 RPM, and 1500 RPM) are utilized to verify the effectiveness of the proposed method. Finally, Section IV provides a summary and conclusions of this paper.

II. STATISTICS DETECTION INDEX SYMPTOM PARAMETER SELECTION

A. SYNTHETIC DETECTION INDEX (SDI)

According to our previous research achievement [12], supposing that bearing symptom parameter (SP) x_1 and x_2 are calculated from vibration signal in condition 1 and condition 2 respectively. Both of bearing symptom parameter is conforms the normal distribution. It is expressed by $N(\mu_1, \sigma_1)$ and $N(\mu_2, \sigma_2)$ respectively. Here, μ means the SP mean and σ means the standard deviation of SP. The larger the value of $|x_1 - x_2|$, the superior the sensitivity of discrimination between the SP in condition 1 and the SP in condition 2. Let $z = x_2 - x_1$, it also follows to the normal distribution $N(\mu_2 - \mu_1, \sigma_2 - \sigma_1)$. The density function of z can be seen in Eq. 1:

$$f(z) = \frac{1}{\sqrt{2\pi\left(\sigma_1^2 + \sigma_2^2\right)}} \exp\{\frac{\{z - \left(\mu_1^2 + \mu_2^2\right)\}^2}{2\left(\sigma_1^2 + \sigma_2^2\right)}\} \quad (1)$$

here, $\mu_2 \geq \mu_1$ (when $\mu_1 \geq \mu_2$, it can draw the same conclusion). The probability can be showed in Eq. 2:

$$P_0 = \int_{-\infty}^{0} f(z) d_z \quad (2)$$

where, $1 - P_0$ is mean the "Discrimination Rate (DR)". With the substitution:

$$\mu = \frac{z - (\mu_2 - \mu_1)}{\sqrt{\sigma_1^2 + \sigma_1^2}} \quad (3)$$

into Eq. (1) and (2), the P_0 can be calculated by:

$$P_0 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-DI} \exp(-\frac{\mu^2}{2}) d_\mu \quad (4)$$

where, the detection index (DI) is obtained by:

$$DI = \frac{|\mu_2 - \mu_1|}{\sqrt{\sigma_1^2 + \sigma_2^2}} \quad (5)$$

Obviously, the larger the value of the DI, the superior the value of the "Discrimination Rate ($DR = 1 - P_0$)" will be, the better the SP will get. Therefore, DI is employed as the index of the performance to evaluate the discriminant sensitivity of the SP.

So, suppose symptom parameter for diagnosing is M and fault types is N, and the synthetic detection index (SDI) can be defined in Eq. 6:

$$SDI = \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \sum_{k=1}^{M} \frac{|\mu_{ik} - \mu_{jk}|}{\sqrt{\sigma_{ik}^2 + \sigma_{jk}^2}} \quad (6)$$

TABLE 1. Symptom parameters of time domain, frequency domain.

Time domain		Frequency domain
Mean $p_1=\frac{1}{N}\sum_{k=1}^{N}x(k)$	Skewness $p_7=\frac{1}{N-1}\sum_{n=1}^{N}\left(\frac{x(n)-p_1}{p_6}\right)^3$	Spectral amplitude mean $p_{13}=\frac{1}{M}\sum_{n=1}^{M}s(k)$
Root mean square $p_2=\sqrt{\frac{1}{N}\sum_{k=1}^{N}x^2(k)}$	Kurtosis $P_8=\frac{1}{N-1}\sum_{n=1}^{N}\left(\frac{x(n)-p_1}{p_6}\right)^4$	Spectral amplitude standard deviation $p_{14}=\sqrt{\frac{1}{M\text{-}1}\sum_{n=1}^{M}\left(s(k)\text{-}p_{13}\right)^2}$
Square root amplitude $p_3=\left(\frac{1}{N}\sum_{n=1}^{N}\sqrt{\lvert x(n)\rvert}\right)^2$	Crest factor $P_9=\frac{p_5}{p_2}$	Spectral frequency skewness $p_{15}=\sum_{k=1}^{M}\frac{\left(\frac{f(k)\text{-SGF}}{STDF}\right)^3 s(k)}{\sum_{j=1}^{M}s(j)}$
Mean amplitude $P_4=\frac{1}{N}\sum_{n=1}^{N}\lvert x(n)\rvert$	Clearance factor $P_{10}=\frac{p_5}{p_3}$	Spectral frequency kurtosis $p_{16}=\sum_{k=1}^{M}\frac{\left(\frac{f(k)\text{-SGF}}{STDF}\right)^4 s(k)}{\sum_{j=1}^{M}s(j)}$
Maximum peak $p_5=\frac{1}{2}\left(\max\left(x(n)\right)-\min\left(x(n)\right)\right)$	Shape factor $P_{11}=\frac{p_2}{p_4}$	Spectral amplitude skewness $p_{17}=\frac{1}{M\text{-}1}\sum_{n=1}^{M}\left(\frac{s(k)\text{-}p_{13}}{p_{14}}\right)^3$
Standard deviation $p_6=\sqrt{\frac{1}{N-1}\sum_{n=1}^{N}\left(x(n)-p_1\right)^2}$	Impulse factor $P_{12}=\frac{p_5}{p_4}$	Spectral amplitude kurtosis $p_{18}=\frac{1}{M\text{-}1}\sum_{n=1}^{M}\left(\frac{s(k)\text{-}p_{13}}{p_{14}}\right)^4$

$\text{SGF}=\sum_{k=1}^{M}f(k)s(k)\Big/\sum_{k=1}^{M}s(k)$; $STDF=\sqrt{\sum_{k=1}^{M}\left(f(k)\text{-SGF}\right)^2 s(k)\Big/\sum_{k=1}^{M}s(k)}$.

Notes: $x(n)$ is a signal series for $n=1,2,\cdots,N$, where N is the number of data point, $s(k)$ is the spectrum of $x(n)$ for $k=1,2,\cdots,M$, where M is the number of spectrum. $f(k)$ is the frequency value of the kth spectrum.

B. SYMPTOM PARAMETER SELECTION

There are many symptom parameters for fault diagnosis, it can represent different degrees of fault information and the sensitivity and dominant are different in certain situation. The symptom parameters cover from time domain, frequency domain. Table 1 shows that the different symptom parameters according to [13].

The amplitude of $f(k)$ have a deep influence to symptom parameter's value, it may loads that each of symptom parameters has a different order of magnitude. So, before calculating the symptom parameters, it needs normalization the spectrum and normalization the symptom parameters.

$$f'(k)=f(k)\Big/\sum_{i=1}^{N}f(k) \tag{7}$$

where, $f'(k)$ is spectrum of normalization.

$$p_i'=(p_i-\bar{p})/Sp_i \tag{8}$$

where p_i is the calculated fault symptom parameters, p_i' fault symptom parameters of normalization. $\bar{p}$ is the mean symptom parameter's value of normal condition, Sp_i is the standard deviation symptom parameters of normal condition.

When there is a fault in the rolling bearing, due to the presence of impact components in the collected vibration signal, kurtosis, Skewness, and other statistics will react to the impact signal. Usually, when there is no fault in the bearing signal, the vibration signal kurtosis of the bearing will be around 3 and the Skewness will be around 0. When the rolling bearing have fault, the more severe the fault, the farther the kurtosis will be from 3, and the more the Skewness will deviate from 0.

III. CDA FOR FAULT DIAGNOSIS

A. CANONICAL DISCRIMINANT ANALYSIS

CDA is a dimension reduction technique derived from principal component analysis and canonical correlation analysis. Given a classification variable and several quantitative variables it is possible to derive canonical variables (summarizing between-class variation) that are linear combinations of the quantitative variables. These canonical variables can then be evaluated for unclassified observations and the resulting scores used for discrimination [14].

Referring to [15], CDA assumes g independent samples $\mathbf{X}$

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n_1} \\ x_{21} & x_{22} & \cdots & x_{2n_2} \\ \vdots & \vdots & \vdots & \vdots \\ x_{g1} & x_{g2} & \cdots & x_{gn_g} \end{bmatrix} \tag{9}$$

The individual sample mean vectors is $\bar{\mathbf{x}}$, there

$$\bar{x}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} x_{ij}, i = 1, 2, \cdots, g \tag{10}$$

and over all mean centers of all samples $\bar{\mathbf{x}}$ is

$$\bar{x} = \frac{1}{\sum_{i=1}^{g} n_i}\sum_{i=1}^{g}\sum_{j=1}^{n_i} x_{ij} = \frac{1}{\sum_{i=1}^{g} n_i}\sum_{i=1}^{g} n_i\bar{x}_i \tag{11}$$

The sample variance covariance matrix is

$$Q_w = E\Big/\left(\sum_{i=1}^{g} n_i - g\right) \tag{12}$$

where E is the pooled within-class sums of squares and cross product matrix, expressed as:

$$E = \sum_{i=1}^{g}\sum_{j=1}^{n_i}\left(x_{ij} - \bar{x}_i\right)\left(x_{ij} - \bar{x}_i\right)^T \tag{13}$$

Q_b is the between-class sum of squares and cross-product matrix, which is given by:

$$Q_b = \sum_{i=1}^{g} n_i\left(x_i - \bar{x}_i\right)\left(x_i - \bar{x}_i\right)^T \tag{14}$$

Based on the objective of CDA that makes between-class variance as large as possible and within-class variance as small as possible, choosing β to maximize the objective function λ

$$\lambda(\beta) = \frac{\beta^T Q_b \beta}{\beta^T Q_w \beta} \tag{15}$$

To do so, differentiating λ with respect to β and letting it be 0 yields

$$Q_w^{-1} Q_b \beta = \lambda\beta \tag{16}$$

where β are eigenvectors corresponding to the r non-zero eigenvalues of $Q_w^{-1}Q_b$ and maximizing λ.

If the r non-zero eigenvalues are ranked from largest to least as $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r$, then $\beta_1 \geq \beta_2 \geq \cdots \geq \beta_r$ are corresponding coefficients of the respective canonical variables. The first canonical score $Can_1 = \beta_1^T x$, which may be interpreted as the single best linear discriminator of these g samples. The second canonical score $Can_2 = \beta_2^T x$, is chosen in the same way but subject to the additional restriction that it is uncorrelated with the first canonical variable. It may be interpreted as the second best linear discriminator of these g populations. Proceed in the similar way to derive the other component variables.

TABLE 2. Symptom parameters of time domain, frequency domain.

Discriminant function (Learning) (Canonical variable z_{ik})	Canonical variable in diagnosis z_{ix}
$\underset{k=O,I,R}{z_{1k}} = \sum_{j=1}^{M} a_{1j} p_{jk}$	$z_{1x} = \sum_{j=1}^{M} a_{1j} p_{jx}$
$\underset{k=O,I,R}{z_{2k}} = \sum_{j=1}^{M} a_{2j} p_{jk}$	$z_{2x} = \sum_{j=1}^{M} a_{2j} p_{jx}$

Because the basic interest of conducting CDA is in the reduction of dimensionality of the raw data while keeping the maximum separation of different classes, it is interested in evaluating the first several canonical variables corresponding to the first several eigenvectors. According to [16], it uses $p-$value to test whether a canonical variable is significantly correlated to a set of dummy variables. It can obtain a list of $\gamma = \min(p, g-1)$ canonical variables, uncorrelated with each other.

B. CDA FAULT DIAGNOSIS METHOD

For bearing fault diagnosing, the first canonical variable and second canonical variable were obtained through CDA. The diagnosis procedure has two steps: one is learning stage and the other is diagnosis stage. In learning stage, the single fault (inner race fault, outer race fault and roller fault) employs the Table 1 to calculate the symptom parameters p_{jk} $(j = 1\ M,\ k = O, I, R)$, then the discriminant function (canonical variable) can be calculated by Table 2. The discriminant function z_{1k}, z_{2k} (k =I (inner race fault), (outer race fault), (roller fault))can represent first canonical variable and second canonical variable. In diagnosis stage, the fault employs the Table 1 to calculate the symptom parameters p_{jx} $(j = 1\ M)$, and z_{1x}, z_{2x} represent the diagnosis stage canonical variable. The canonical variable can be seen in Table 2.

Depend on the distribution of CDA, the fault can be inferred. Compare with the normal data in the CDA distribution, the farther diagnosis data get away from normal data, the severer the fault severity could be. The closer diagnosis data get away from normal data, the less serious the fault severity could be.

IV. THE PROPOSED METHOD

The proposed method was used for diagnosing bearing fault under different fault severity. It involves statistics detection index and canonical discriminant analysis. The proposed method acquires the bearing signal of normal condition and fault condition. The specific steps of the proposed method are illuminated as follows:

Step 1: Data acquisition. The bearing normal condition and fault condition (e.g., slight(small) fault, medium(middle) fault, severe(big) fault) vibration signal was recorded. Each

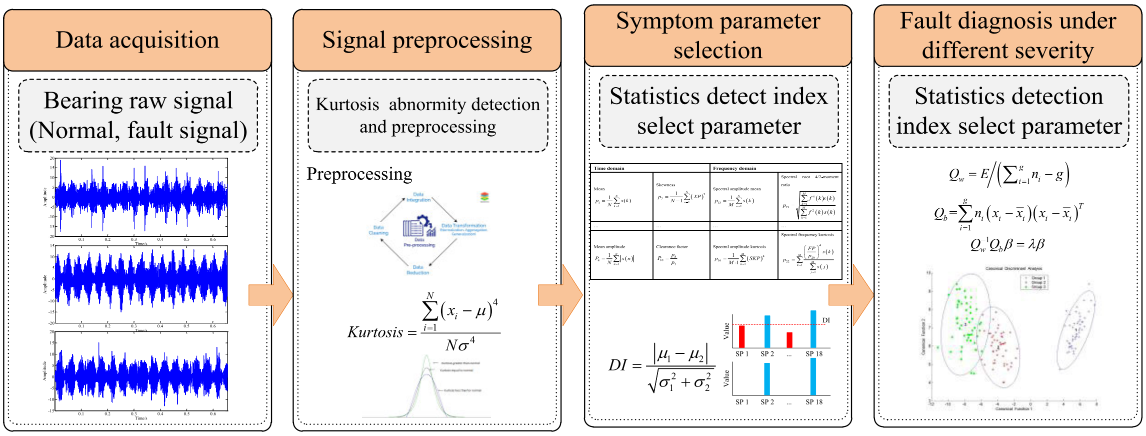

FIGURE 1. Framework of proposed method.

condition acquisition condition such as sampling frequency and sampling time should be the same.

Step 2: Signal preprocessing. Considering that the magnitude of different bearing conditions has a major difference, the method normalizes the raw data in the range (-1, 1).

Step 3: Kurtosis abnormity detection. Before symptom parameters calculation stage, the bearing must be decided if it is in normal or abnormal. If the bearing is abnormal, the measured vibration signal is used for further step by the proposed method. In order to determine whether the bearing is normal condition or not, after the usually noise cancelling (high pass filtering method) is employed, kurtosis is used to detect bearing abnormity.

Step 4: Symptom parameters calculation. The time domain and frequency domain of raw signal symptom parameters was calculated by Table 1 to represent the condition information of bearing.

Step 5: DI symptom parameters selection. Considering that some of the symptom parameter may insensitive or redundancy. The statistics detection index was employed to select the optimal subset to fault diagnosis.

Step 6: Canonical discriminant analysis. The selected symptom parameters were brought into the CDA algorithm to calculate the canonical score.

Step 7: Fault diagnosis under different severity. According to the CDA canonical score distribution. The bearing fault can be diagnosis under different fault severity.

The framework of the presented method was shown in Fig. 1.

V. EXPERIMENT ANALYSIS

In this part the results of the some performed experiments are presented. Each of experiment represents a further step to obtained the objectives described in the before section.

A. EXPERIMENT PLATFORM

In this section, the effectiveness of the proposed technique will be validated by the analysis of three rotating machinery fault signal of outer race defect, inner race defect and roller defect and normal signal without fault.

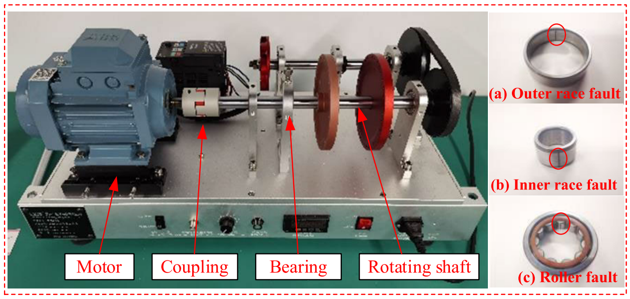

FIGURE 2. Bearing fault diagnosis experimental system and bearing fault of outer race, inner race and roller.

As shown in Fig. 2, the experimental platform is for bearing fault diagnosis test, it includes a servo motor, a coupling, a bearing and rotating shaft. The bearings have three different type fault which are outer race fault, inner race fault, and the roller fault, they were artificially made by using electro discharge machining.

B. EXPERIMENTAL VERIFICATION

1) RAW VIBRATION SIGNAL AND ITS PROBABILITY DISTRIBUTION

Acquiring the vibration signal of three fault signal and normal signal from experiment platform at 1000 RPM, and the sampling frequency is 100 kHz. The Raw vibration signal, normalized signal and its probability density function (PDF) can be seen in Fig. 3. As observed from Fig. 3, the PDF of normal condition is followed normal distribution. The fault condition PDF didn't obey normal distribution.

2) CONDITION MONITORING BASED ON KURTOSIS

Before symptom parameters calculation stage of the proposed method, the bearing must be decided if it is in normal or abnormal. If the bearing is abnormal, the measured vibration signal is used for further step by the proposed

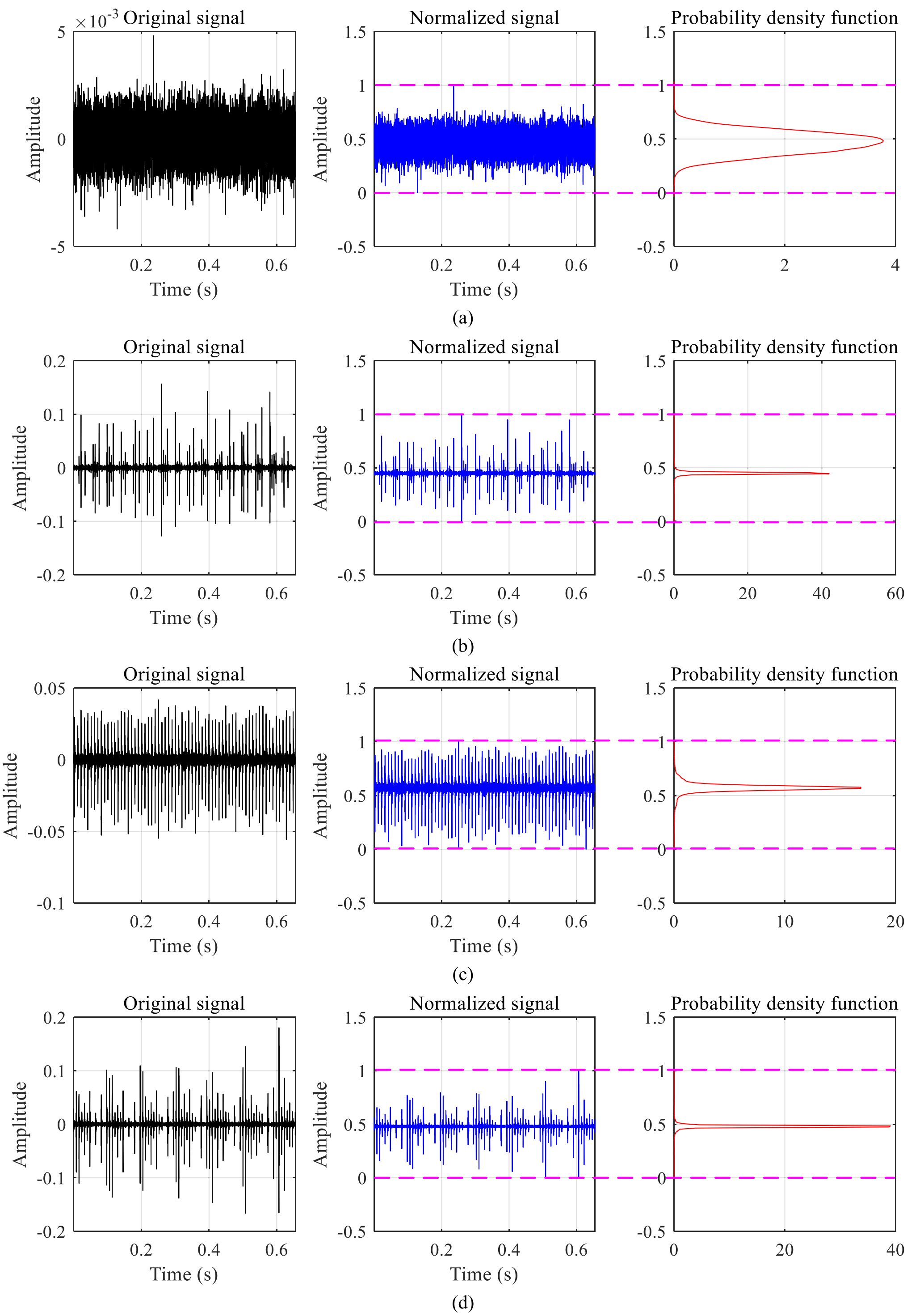

FIGURE 3. Raw vibration signal and probability density function (a: Normal signal, b: Inner race fault, c: Outer race fault, d: Roller fault.

method. Kurtosis is one of the most effective parameters to detect abnormity of bearing. In order to determine whether the bearing is normal condition or not, after the usually noise cancelling (high pass filtering) is employed, kurtosis can be used to detect bearing abnormity. Kurtosis has been successfully used to detect bearing faults in some occasions [17], [18]. It definition is:

$$Kurtosis = \sum_{i=1}^{N} (x_i - \mu)^4 \Big/ N\sigma^4 \tag{17}$$

where, μ is vibration signal mean value, σ is vibration signal standard deviation, and N is length of vibration signal x_i.

TABLE 3. Average kurtosis values of each condition.

Type	Kurtosis Value
Normal condition	2.73
Inner race fault condition	115.97
Outer race fault condition	17.71
Roller race fault condition	141.33

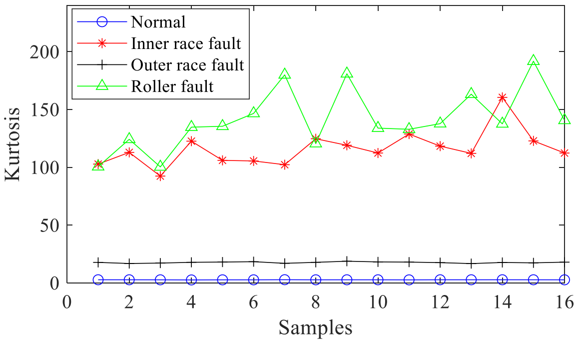

FIGURE 4. The value of kurtosis in bearing fault of outer race, inner race and roller.

According to literatures [17], [18], when vibration signal kurtosis value exceeds 5, there is fault. This paper adopts this analysis method as condition monitoring to judge abnormal condition vibration signals. Kurtosis is employed for condition monitoring, it only judge bearing health or not, the further diagnosis will employ the proposed method. Fig. 4 shows the kurtosis distribution values of the samples in different states, and Table 3 shows the average kurtosis values of each condition.

From the Table 3, the result shown that the kurtosis value of normal condition is approximate to 3, it verified the theory of method kurtosis analysis for fault diagnosis: at normal condition, the kurtosis value around 3. The kurtosis value of Inner race fault, outer race fault, roller race fault is 97.27, 54.18 and 103.42 respectively. It is farther away from 3. It indicated there are fault in measured signal. It can continue further operation.

3) SDI SYMPTOM PARAMETER SELECTION

The experiment employs the inner race fault with severe(big) fault size (d=0.25mm, w=0.7mm), medium(middle) fault size (d=0.15mm, w=0.5mm), slight(small) fault size (d=0.05mm, w=0.3mm) to assess the proposed method effectiveness, after the band filtering and normalization, the kurtosis of slight(small), medium(middle) and severe(big) fault are 57.23, 79.47 and 102.43 respectively. It manifests all of the signal is abnormal. Separating the severe fault, medium fault, and slight fault to 16 sub-segment respectively, and calculate the symptom parameters of each of sub-segment. Through Eq. 5 to calculate the DI value to select the parameters, the DI value can be seen in Table 4.

P=0 indicates that the corresponding feature parameter P is set to zero, with the aim of evaluating the impact of the P value on the DI value. According to the references [19] theory, experiments result and signal characteristics in this study, the relationship between the SDI value and the severity of fault is as follows. Table 5 exhibits the condition diagnosis sensitivity standard, choosing the DI value large than 1.3, the result can be seen in Fig. 5.

According to the Table 5 and Fig. 5, choose p2: root mean square, p4: mean amplitude, p6: standard deviation, p8: kurtosis, p9: crest factor, p10: clearance factor, p12:

TABLE 4. DI value of symptom parameters.

No.	P_1=0	P_2=0	P_3=0	P_4=0	P_5=0	P_6=0
DI	1.17	1.37	1.29	3.50	0.45	1.44
No.	P_7=0	P_8=0	P_9=0	P_{10}=0	P_{11}=0	P_{12}=0
DI	1.25	1.31	1.37	2.76	0.94	5.14
No.	P_{13}=0	P_{14}=0	P_{15}=0	P_{16}=0	P_{17}=0	P_{18}=0
DI	0.46	1.63	0.79	0.80	1.55	1.59

TABLE 5. Diagnosis sensitivity for condition diagnosis.

Detection index	Discrimination rate	Sensitivity
<0.85	<80%	Low
0.85～1.30	80%～90%	Slightly low
1.30～1.65	90%～95%	Medium
1.65～2.33	95%～99%	Severe
>2.33	>99%	Extremely Severe

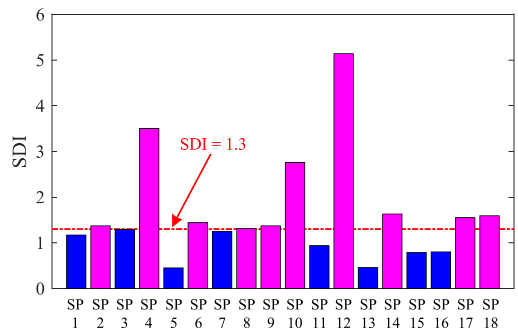

FIGURE 5. The symptom parameters DI value.

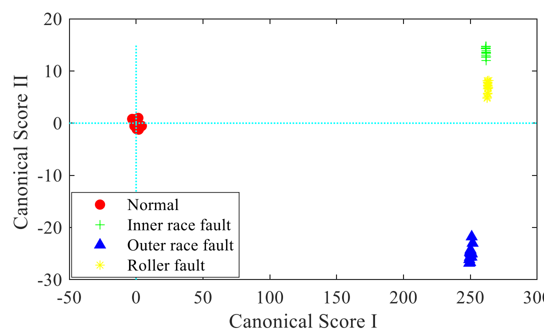

FIGURE 6. CDA canonical score of bearing fault diagnosis (1500RPM).

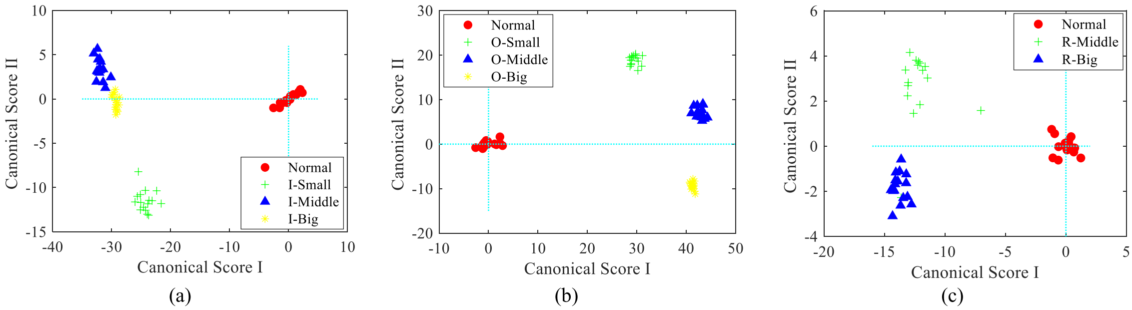

FIGURE 7. **CDA canonical score of small fault, middle fault and big fault at 500 RPM: (a) Inner race fault; (b) Outer race fault; (c) Roller fault.**

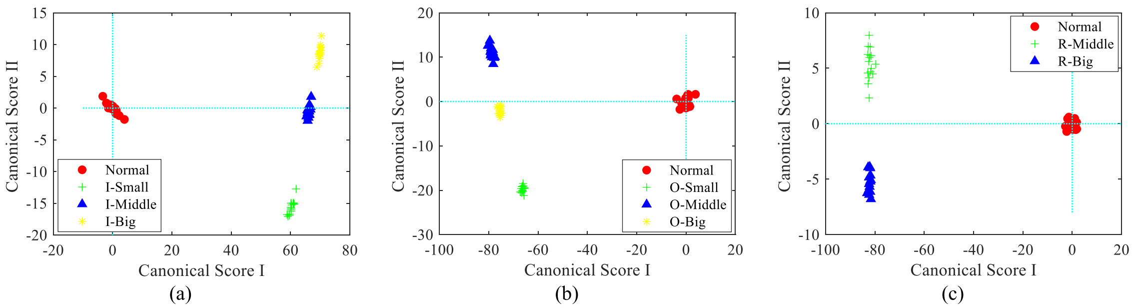

FIGURE 8. **CDA canonical score of small fault, middle fault and big fault at 1000 RPM: (a) Inner race fault; (b) Outer race fault; (c) Roller fault.**

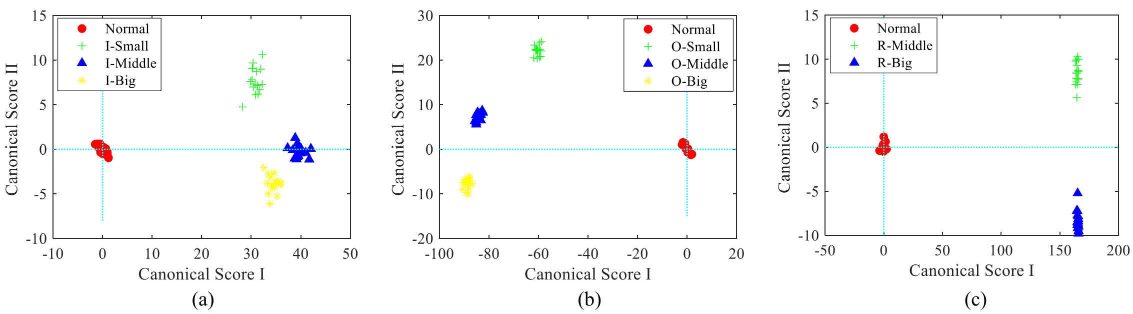

FIGURE 9. **CDA canonical score of small fault, middle fault and big fault at 1500 RPM: (a) Inner race fault; (b) Outer race fault; (c) Roller fault.**

impulse factor, p14: spectral amplitude standard deviation, p17: spectral amplitude skewness, p18: spectral amplitude kurtosis as a symptom parameters set.

4) CDA FAULT DIAGNOSIS

The symptom parameters brought into CDA algorithm can obtain the canonical function like

$$f_1 = 1.38x_1' - 2.27x_2' + \cdots - 12.11x_{10}' + 19.21 \quad (18)$$

$$f_2 = 0.77x_1' - 5.47x_2' + \cdots + 2.63x_{10}' - 7.03 \quad (19)$$

Based on the canonical function, it can calculate the canonical score of CDA. The first canonical score and second canonical score of CDA distribution figure as shown in Fig. 6. It can be seen from the Fig. 6 that the three fault types and the normal state are clearly distinguished, indicating that the method in this study can effectively identify bearing faults.

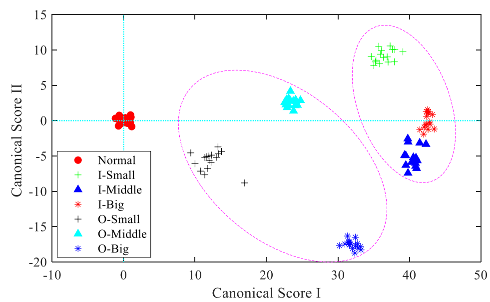

FIGURE 10. **Small, middle, big fault of inner and outer race fault (500 RPM).**

C. FAULT SEVERITY ANALYSIS

1) BEARING SINGLE FAULT DIAGNOSIS MODEL

Fig. 7~9 show the CDA canonical score of small fault, middle fault and big fault at 500 RPM, 1000 RPM, and 1500 RPM. Where, I-Small denotes inner race small fault, I-Middle

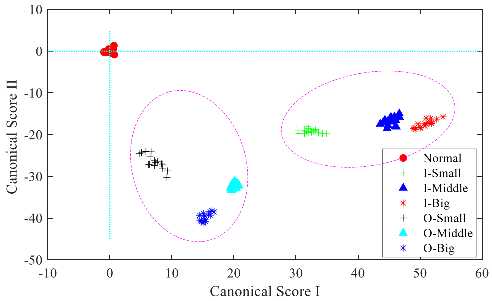

FIGURE 11. Small, middle, big fault of inner and outer race fault (1000 RPM).

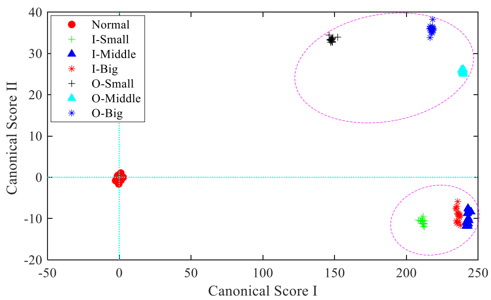

FIGURE 12. Small, middle, big fault of inner and outer race fault (1500 RPM).

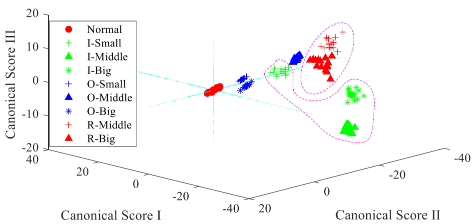

FIGURE 13. Small, middle, big fault of inner, outer race and roller (500 RPM).

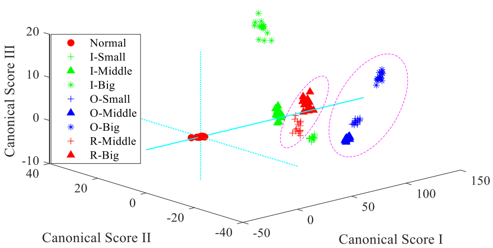

FIGURE 14. Small, middle, big fault of inner, outer race and roller (1000 RPM).

denotes inner race middle fault and I-Big denotes inner race big fault, so as to O-Big denotes outer race big fault and R-Big denotes roller race big fault. It can be seen that different fault severity are effectively identified.

2) SMALL, MIDDLE, BIG FAULT OF INNER AND OUTER RACE FAULT

As same to the operation of Section V-C1, the small fault, middle fault, big fault in inner and outer race fault at 500 RPM, 1000 RPM, 1500 RPM were applied to calculate

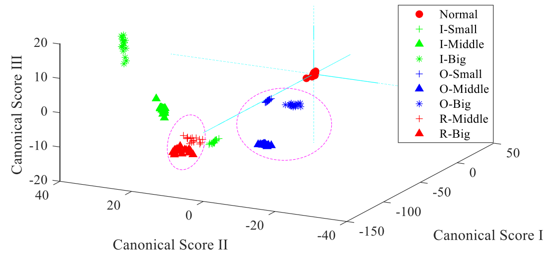

FIGURE 15. Small, middle, big fault of inner, outer race and roller (1500 RPM).

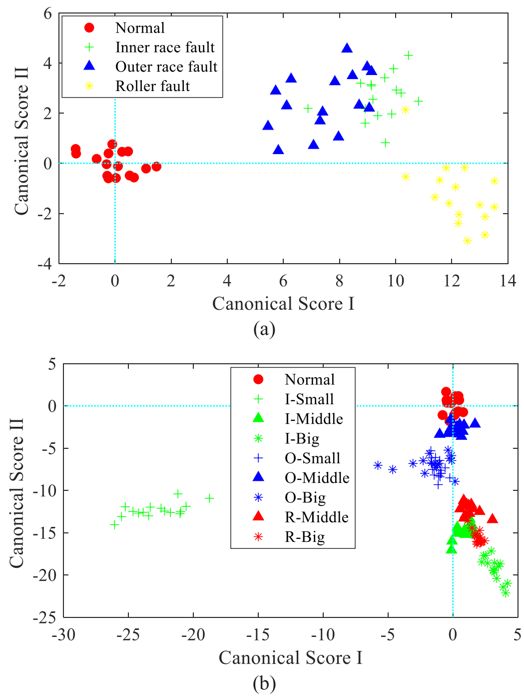

FIGURE 16. All symptom parameters CDA canonical score distribution.

the canonical score of CDA. After using CDA for analysis, the lighter the degree of rolling bearing failure, the closer the distance between canonical components and normal signals, and the higher the degree of failure, the farther the distance between canonical components and normal signals. Therefore, the degree of failure of rolling bearings can be analyzed by distance. The result can be seen from Fig. 10∼12.

As observed from Fig. 10, the inner race fault and outer race fault and normal are separate from each other. According to the distance farther from the normal signal, it can infer the fault severity. The big fault has the strong impulse amplitude and has a big distance farther from normal condition. Conversely, the small fault has a small distance relative to big fault. Based on that, it can infer the fault severity of each kind of fault.

3) SMALL, MIDDLE, BIG FAULT OF INNER, OUTER RACE FAULT AND ROLLER FAULT

Fig. 13∼15 show the canonical score of CDA of the small fault, middle fault, big fault in inner and outer race fault

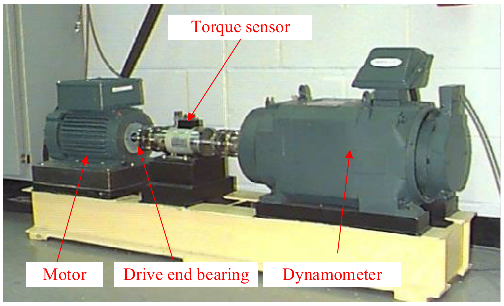

FIGURE 17. Test rig of bearings (CWRU Bearing Data).

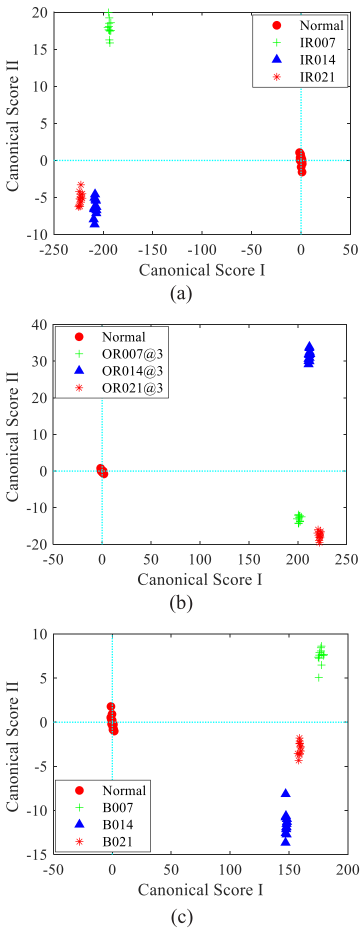

FIGURE 18. CDA canonical score of small fault, middle fault and big fault of CWRU: (a) Inner race fault; (b) Outer race fault; (c) Ball fault.

at 500 RPM, 1000 RPM, 1500 RPM, the results show that the method can also distinguish the different fault degrees of the three fault states.

D. A COMPARISON OF ALL SYMPTOM PARAMETERS AND SDI SELECTED SYMPTOM PARAMETERS

For further proved that the advantage of SDI selected symptom parameters, a comparison between all symptom parameters and SDI selected symptom parameters was presented in this section. Fig. 16 shows the CDA canonical score distribution result of all symptom parameters. The result of Fig. 16 shows that using all symptom parameter to fault diagnosis and analyze the fault severity is unsuccessful. Even, some of fault signal have certain overlap with normal data. It manifested that the SDI symptom parameter selection is indispensable. It proved the effectiveness of proposed method.

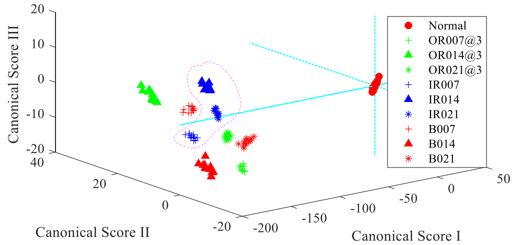

FIGURE 19. Small, middle, big fault of inner, outer race and ball (CWRU).

E. VALIDATION EXPERIMENT OF CASE WESTERN RESERVE UNIVERSITY

The bearing fault experimental signal of Case Western Reserve University (CWRU) are employed to verify the effectiveness of the proposed method. The experimental platform device is shown as Fig. 17.

As same to the operation of Section V-C1, the small fault, middle fault, big fault in inner, outer race and roller fault at 1797 RPM were applied to calculate the canonical score of CDA. The result can be seen from Fig. 18 ∼ 19. Where, IR007 denotes inner race small fault (fault width: 0.007 inch), IR014 denotes inner race middle fault (fault width: 0.014 inch) and IR021 denotes inner race big fault (fault width: 0.021 inch), so as to OR021 denotes outer race big fault and B021 denotes ball big fault. It can be seen that different fault degrees are effectively identified.

VI. CONCLUSION

Fault severity is a crucial factor in bearing fault diagnosis. This paper presents a method for diagnosing bearings under different fault severity levels, utilizing statistical detection index for symptom parameter selection and canonical discriminant analysis. After noise cancellation, kurtosis, one of the most effective parameters, is initially employed to detect any abnormalities in the bearing. The statistical detection index, based on statistical analysis theory, is then utilized to extract the most sensitive symptom parameters. These extracted parameters are used to calculate the symptom parameter values, which represent the fault features. By applying the sensitive symptom parameters to the CDA algorithm, canonical scores are obtained. Based on the distribution of canonical scores, fault diagnosis can be performed under different fault severity levels.

The proposed method can diagnose the condition of the bearing as healthy, slight fault, medium fault, and severe fault without the need to know specific characteristic parameters, such as fault characteristic frequency. Moreover, the proposed method can effectively diagnose the fault. Experimental analysis of inner race and outer race faults in the bearing at different rotational speeds (500 RPM, 1000 RPM, and 1500 RPM) verifies the effectiveness of the proposed method.

Although the proposed method demonstrates superiority in achieving fault diagnosis under different severity levels, there are certain aspects that may require improvement. Firstly, the threshold for the statistical detection index needs to be manually set. Secondly, the time-consuming nature of the method is not specifically considered in this paper. These factors will be taken into consideration in future research endeavors.

REFERENCES

[1] R. B. Randall and J. Antoni, "Rolling element bearing diagnostics—A tutorial," *Mech. Syst. Signal Process.*, vol. 25, no. 2, pp. 485–520, Feb. 2011.

[2] H. Zhong, J. Liu, L. Wang, Y. Ding, and Y. Qian, "Bearing fault diagnosis based on kernel independent component analysis and antlion optimization," *Trans. Inst. Meas. Control*, vol. 43, no. 16, pp. 3573–3587, Dec. 2021.

[3] H. Hong and M. Liang, "Fault severity assessment for rolling element bearings using the Lempel–Ziv complexity and continuous wavelet transform," *J. Sound Vibrat.*, vol. 320, nos. 1–2, pp. 452–468, Feb. 2009.

[4] C. R. Soto-Ocampo, J. D. Cano-Moreno, J. M. Mera, and J. Maroto, "Bearing severity fault evaluation using contour maps—Case study," *Appl. Sci.*, vol. 11, no. 14, p. 6452, Jul. 2021.

[5] D. H. C. de Sá Só Martins, D. P. Viana, A. A. de Lima, M. F. Pinto, L. Tarrataca, F. L. E. Silva, R. H. R. Gutiérrez, T. De Moura Prego, U. A. B. V. Monteiro, and D. B. Haddad, "Diagnostic and severity analysis of combined failures composed by imbalance and misalignment in rotating machines," *Int. J. Adv. Manuf. Technol.*, vol. 114, nos. 9–10, pp. 3077–3092, Jun. 2021.

[6] X. Yan, W. Zhang, and M. Jia, "A bearing fault feature extraction method based on optimized singular spectrum decomposition and linear predictor," *Meas. Sci. Technol.*, vol. 32, no. 11, Nov. 2021, Art. no. 115023.

[7] R. Yan and R. X. Gao, "Complexity as a measure for machine health evaluation," *IEEE Trans. Instrum. Meas.*, vol. 53, no. 4, pp. 1327–1334, Aug. 2004.

[8] A. Sharma, M. Amarnath, and P. Kankar, "Feature extraction and fault severity classification in ball bearings," *J. Vibrat. Control*, vol. 22, no. 1, pp. 176–192, Jan. 2016.

[9] F. Pacheco, M. Cerrada, R.-V. Sánchez, D. Cabrera, C. Li, and J. V. de Oliveira, "Attribute clustering using rough set theory for feature selection in fault severity classification of rotating machinery," *Exp. Syst. Appl.*, vol. 71, pp. 69–86, Apr. 2017.

[10] L. Cheng, S. Fu, H. Zheng, Y. Huang, and Y. Xu, "Fault identification and severity assessment of rolling element bearings based on EMD and fast kurtogram," *J. Vibroeng.*, vol. 18, no. 6, pp. 3668–3683, Sep. 2016.

[11] M. Xia, H. Shao, D. Williams, S. Lu, L. Shu, and C. W. de Silva, "Intelligent fault diagnosis of machinery using digital twin-assisted deep transfer learning," *Rel. Eng. Syst. Saf.*, vol. 215, Nov. 2021, Art. no. 107938.

[12] K. Li, P. Chen, and S. Wang, "An intelligent diagnosis method for rotating machinery using least squares mapping and a fuzzy neural network," *Sensors*, vol. 12, no. 5, pp. 5919–5939, May 2012.

[13] M. Zeng, Y. Yang, J. Zheng, and J. Cheng, "Maximum margin classification based on flexible convex hulls for fault diagnosis of roller bearings," *Mech. Syst. Signal Process.*, vols. 66–67, pp. 533–545, Jan. 2016.

[14] N. Strachan, "A potential method for the differentiation between haddock fish stocks by computer vision using canonical discriminant analysis," *ICES J. Mar. Sci.*, vol. 52, no. 1, pp. 145–149, Feb. 1995.

[15] R. Pu and D. Liu, "Segmented canonical discriminant analysis of in situ hyperspectral data for identifying 13 urban tree species," *Int. J. Remote Sens.*, vol. 32, no. 8, pp. 2207–2226, Mar. 2011.

[16] R. Khattree and D. N. Naik, *Multivariate Data Reduction and Discrimination with SAS Software*. Cary, NC, USA: SAS Inst., 2000.

[17] R. B. W. Heng and M. J. M. Nor, "Statistical analysis of sound and vibration signals for monitoring rolling element bearing condition," *Appl. Acoust.*, vol. 53, nos. 1–3, pp. 211–226, Jan. 1998.

[18] X. Niu, L. Zhu, and H. Ding, "New statistical moments for the detection of defects in rolling element bearings," *Int. J. Adv. Manuf. Technol.*, vol. 26, nos. 11–12, pp. 1268–1274, Nov. 2005.

[19] K. Li, P. Chen, and H. Wang, "Intelligent diagnosis method for rotating machinery using wavelet transform and ant colony optimization," *IEEE Sensors J.*, vol. 12, no. 7, pp. 2474–2484, Jul. 2012.

XUEWEI SONG received the M.S. degree from Anhui University, China. She is currently a Lecturer with the Naval Architecture and Shipping College, Guangdong Ocean University, China. Her current research interests include machinery maintenance technology, motor diagnosis technology, machinery condition monitoring and diagnosis technology, and intelligent control.

ZHIQIANG LIAO received the Ph.D. degree from Mie University, Japan. He is currently a Lecturer with the Naval Architecture and Shipping College, Guangdong Ocean University, China. His current research interests include machinery maintenance technology, machinery condition monitoring and diagnosis technology, and intelligent robot systems.

BAOZHU JIA received the Ph.D. degree in marine engineering from Dalian Maritime University, in 2006. Currently, he is a Professor with the Guangdong Ocean University and working with the Naval Architecture and Shipping College. His research interests include architecture and ocean engineering, automatic control, and intelligent fault diagnosis.

DEFENG KONG received the Ph.D. degree from Harbin Engineering University, in 2020. He is currently a Lecturer with the Naval Architecture and Shipping College, Guangdong Ocean University, China. His current research interests include the analysis of electromagnetic field and the vibration control of electrical machines.

JINZHANG NIU is currently an Associate Professor with the Naval Architecture and Shipping College, Guangdong Ocean University, China. His current research interests include the analysis of signal acquisition and signal processing.