Research Article

Bearing Fault Dominant Symptom Parameters Selection Based on Canonical Discriminant Analysis and False Nearest Neighbor Using GA Filtering Signal

Shilun Zuo¹ and Zhiqiang Liao²

¹School of Mechanical Engineering, Chongqing University of Science and Technology, Chongqing 401331, China

²Graduate School of Bio-resources, Mie University, 1577 Kurimamachiya-cho, Tsu, Mie 5148507, Japan

Correspondence should be addressed to Zhiqiang Liao; zhiqiangliao@126.com

Received 1 November 2019; Revised 31 January 2020; Accepted 4 February 2020; Published 14 April 2020

Academic Editor: Roman Lewandowski

Copyright © 2020 Shilun Zuo and Zhiqiang Liao. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Symptom parameter is a popular method for bearing fault diagnosis, and it plays a crucial role in the process of building a diagnosis model. Many symptom parameters have been performed to extract signal fault features in time and frequency domains, and the improper selection of parameter will significantly influence the diagnosis result. For dealing with the problem, this paper proposes a novel dominant symptom parameters selection scheme for bearing fault diagnosis based on canonical discriminant analysis and false nearest neighbor using GA filtered signal. The original signal was filtered by a genetic algorithm (GA) at first and then mapped to the new characteristic subspace through the canonical discriminant analysis (CDA) algorithm. The map distance in the new characteristic subspace is calculated by the false nearest neighbor (FNN) method to interpret the dominance of symptom parameters. The dominant symptom parameters brought to the bearing diagnosis system can improve the diagnosis result. The effectiveness of the proposed method has been demonstrated by the diagnosis model and by comparison with other methods.

1. Introduction

Bearing is a common component in rotating machinery. Bearing failures can cause severe accidents and result in property loss and even personal injury [1, 2]. For that reason, it is significant to monitor and diagnose bearing faults. Previous studies indicate that bearing diagnosis methods mainly focus on data-driven methods [3, 4]. It employs the fault features to build a diagnosis model for diagnosing. The primary process of diagnosis model-based method is as follows: collect the vibration signal in a fault state and build a useful diagnosis model for fault diagnosis, utilizing signal filtering and feature extraction. It has been widely applied in bearing fault diagnosis [5]. In this way, fault symptom parameter selection is an essential operation, which will significantly influence the diagnosis result. There are many symptom parameters which are available to extract fault

features in time and frequency domains. Among these symptom parameters, some symptom parameters are sensitive and dominant, and the others are insensitive. Using all parameters in the model directly may increase processing time and decrease the diagnosis accuracy. It is essential to choose the most representative symptom parameters before diagnosis.

Dominant symptom parameters selection involves much work. It aims to select the optimal feature subset to decrease both the time and complexity of the diagnosis process. Generally, based on the techniques in evaluating the cost function of feature selection approaches, the relevant selection algorithm can broadly be categorized into three groups [6, 7]: the filter, the wrapper, and the hybrid methods. Filter-based symptom parameter selection discovers the general characteristics of a dataset. It is independent of any classification model or clustering model

[8, 9]. The filter-based method is less time consuming and of less efficiency. The wrapper-based method works on a specific model and selects the parameter subset aiming to optimize the model performance according to evaluation criteria [10, 11]. It is more time consuming and more effective than the filter models. The hybrid method tends to take advantage of filter methods for their computational cheapness and wrapper methods for their high accuracy [12, 13].

In the field of bearing fault diagnosis, many techniques were applied in symptom parameter selection under filter-based, wrapper-based, and hybrid methods. Yu et al. [14] proposed features selection by adjusting the rand index and standard deviation ratio (FSASR) with the K-means method and standard deviation (STD) to select the sensitive statistical characteristics of bearing fault signal. Meng et al. [15] applied the binary value of the gravitational search algorithm (BGSA) to find the essential features from the feature set. Sugumaran et al. [16] illustrated a method based on a decision tree to select bearing fault features. Li et al. [17] used a multiscale morphological filter signal processing method to select a train axle bearing fault feature. As the mentioned methods have some limitations in different working conditions, they need more feature methods to adapt to various harsh working conditions. This work focuses on a novel method for selecting dominant symptom parameters. Different from the previous methods, the proposed method imposes the FNN distance in CDA characteristic subspace to represent the performance of symptom parameters which can pick out the most representative symptom parameters in time and frequency domains for fault diagnosis. For avoiding the influences of signal noise, before the symptom parameters calculation, the measured signals go through automatic signal filtering by the GA method. According to the variance analysis theory, canonical discriminant analysis (CDA) can project data to the new characteristic subspace with the maximum between-group difference and the minimum intergroup difference. Canonical discriminant function can get the maximum canonical score to represent data information. False nearest neighbor (FNN) works in the new characteristic subspace, and with the embedded dimension increasing, the system may recover the high-dimensional space state gradually, while the false neighbor point with high similarity will be successively swept. An evaluation criterion is designed according to the similarity (FNN-based space distance) to select the dominant symptom parameters. In this paper, FNN can be used to choose the symptom parameters which show super performance in representing the fault information of roller bearing. The effectiveness of the proposed method was demonstrated by the support vector machine (SVM) diagnosis model and by comparison with other symptom parameters selection methods.

The rest of the paper is organized as follows: Section 2 will introduce GA filtering, symptom parameters of fault diagnosis, CDA and FNN dominant symptom parameters selection, and evaluation criterion, respectively. In Section 3, the experimental validation is carried to examine the

performance of the proposed method. The conclusion is drawn in Section 4.

2. Dominant Symptom Parameters Selection Scheme of Bearing Fault Diagnosis

2.1. GA Filtering. Before the symptom parameters calculation, signal filtering should be performed to decrease the noise influences. The GA filtering has been published in our previous works [1]. As shown in Figure 1(a), in the signal frequency domain, the fault signal and noise are other existence, and the high-pass, low-pass, and bandpass filtering do not work in this situation. If it makes a binary string, each frequency is expressed by 0 or 1. When we find the optimal binary string, in which all of the noise-corresponding frequency positions set the binary string to 0, then the signal frequency spectrum multiplied by optimal binary string can obtain only fault components to be retained. The transform of the multiplied frequency domain signal to the time domain is a filtered process.

As an example, Figure 1 presents the principle of GA filtering. Figure 1(a) shows the original signal frequency domain with noise and fault signal. N represents noise, and S represents fault signal. Figure 1(b) shows the original signals of the time domain. When we perform the GA filtering to the signal, we can get the optimal binary string, as shown in Figure 1(c). The noise-corresponding binary string is set to 0. After the optimal binary string is multiplied with the original signal frequency spectrum, the filtered time-domain signal will be obtained as shown in Figure 1(d). Figure 1(d) shows the filtered signal and can be applied to the next operation.

2.2. Fault Diagnosis Symptom Parameters. There are many symptom parameters used for bearing fault diagnosis. They can carry information about faults. The sensitivity and dominance of parameters are different in different work conditions. They cover from the time domain to the frequency domain. Equations (1)~(18) show different symptom parameters [18].

Time domain is as follows:

$$p_1 = \frac{1}{N} \sum_{t=1}^N x(t), \quad (1)$$

$$p_2 = \sqrt{\frac{1}{N} \sum_{t=1}^N x^2(t)}, \quad (2)$$

$$p_3 = \left(\frac{1}{N} \sum_{t=1}^N \sqrt{|x(t)|} \right)^2, \quad (3)$$

$$p_4 = \frac{1}{N} \sum_{t=1}^N |x(t)|, \quad (4)$$

$$p_5 = \frac{1}{2} (\max(x(t)) - \min(x(t))), \quad (5)$$

FIGURE 1: (a) Original signal frequency domain; (b) original time signal; (c) GA-filtered frequency domain; and (d) the filtered signal.

$$p_6 = \sqrt{\frac{1}{N-1} \sum_{t=1}^N (x(t) - p_1)^2}, \quad (6)$$

$$p_7 = \frac{1}{N-1} \sum_{t=1}^N \left(\frac{x(t) - p_1}{p_6} \right)^3, \quad (7)$$

$$p_8 = \frac{1}{N-1} \sum_{t=1}^N \left(\frac{x(t) - p_1}{p_6} \right)^4, \quad (8)$$

$$p_9 = \frac{p_5}{p_2}, \quad (9)$$

$$p_{10} = \frac{p_5}{p_3}, \quad (10)$$

$$p_{11} = \frac{p_2}{p_4}, \quad (11)$$

$$p_{12} = \frac{p_5}{p_4}, \quad (12)$$

$$p_{13} = \frac{1}{M} \sum_{n=1}^M s(k), \quad (13)$$

$$p_{14} = \sqrt{\frac{1}{M-1} \sum_{n=1}^M (s(k) - p_{13})^2}, \quad (14)$$

$$p_{15} = \frac{\sum_{k=1}^M ((f(k) - \text{SGF})/\text{STDF})^3 s(k)}{\sum_{k=1}^M s(k)}, \quad (15)$$

$$p_{16} = \frac{\sum_{k=1}^M ((f(k) - \text{SGF})/\text{STDF})^4 s(k)}{\sum_{k=1}^M s(k)}, \quad (16)$$

$$p_{17} = \frac{1}{M-1} \sum_{n=1}^M \left(\frac{s(k) - p_{13}}{p_{14}} \right)^3, \quad (17)$$

$$p_{18} = \frac{1}{M-1} \sum_{n=1}^M \left(\frac{s(k) - p_{13}}{p_{14}} \right)^4, \quad (18)$$

$$\text{SGF} = \frac{\sum_{k=1}^M f(k)s(k)}{\sum_{k=1}^M s(k)}, \quad (19)$$

$$\text{STDF} = \sqrt{\frac{\sum_{k=1}^M (f(k) - \text{SGF})^2 s(k)}{\sum_{k=1}^M s(k)}}. \quad (20)$$

Here, $x(t)$ is a signal series for $t = 1, 2, \dots, N$, where N is the length of the signal, and $s(k)$ is the spectrum of $x(t)$ for $k = 1, 2, \dots, M$, where M is the spectrum line. $f(k)$ is the frequency amplitude.

The amplitude of $s(k)$ has a significant influence on the value of such symptom parameters. As a result, each of the symptom parameters has a different order of magnitude. So, before calculating the symptom parameters, both of the spectrum and symptom parameters should be normalized:

$$s(k) = \frac{s'(k)}{\sum_{k=1}^M s'(k)}, \quad (21)$$

where $s(k)$ is the spectrum of normalization. Also,

$$p'_i = \frac{p_i - \bar{p}}{\text{std}_{p_i}}, \quad (22)$$

where p_i is the calculated fault symptom parameter and p'_i is the fault symptom parameter of normalization. $\bar{p}$ is the mean value of such symptom parameters under the normal state, and std_{p_i} is the standard deviation symptom parameter under the normal state.

2.3. CDA and FNN Dominant Symptom Parameters Selection. Canonical discriminant analysis (CDA) is a dimension reduction method developed from the principal component analysis (PCA) method [19] and canonical correlation analysis (CCA) method [20]. Given the label values and quantitative variables, it is possible to extract the canonical components that are linear combinations of the quantitative variables. These canonical components can be applied to evaluate the unlabelled observations. At the same time, the resulting scores are applied for discrimination [21].

Assume symptom parameters S as follows:

$$S = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1j} \\ s_{21} & s_{22} & \cdots & s_{2j} \\ \vdots & \vdots & \vdots & \vdots \\ s_{i1} & s_{i2} & \cdots & s_{ij} \end{bmatrix}. \quad (23)$$

The individual parameter mean vectors are $\bar{S}$, where

$$\bar{s}_j = \frac{1}{h} \sum_{i=1}^h s_{ij}, \quad i = 1, 2, \dots, h, j = 1, 2, \dots, l, \quad (24)$$

and overall mean centres of all symptom parameters $\bar{s}$ are

$$\bar{s} = \frac{1}{h \times l} \sum_{i=1}^h \sum_{j=1}^l s_{ij}, \quad i = 1, 2, \dots, h, j = 1, 2, \dots, l. \quad (25)$$

The parameter variance-covariance matrix is

$$Q_w = \frac{E}{\sum_{i=1}^l h_i - l}, \quad (26)$$

where E is illustrated by

$$E = \sum_{i=1}^l \sum_{j=1}^h (s_{ij} - \bar{s}_j)(s_{ij} - \bar{s}_j)^T. \quad (27)$$

Q_b is expressed by

$$Q_b = \sum_{j=1}^l h_j (\bar{s}_j - \bar{s})(\bar{s}_j - \bar{s})^T. \quad (28)$$

It aims to make the between-class variance as large as possible and, at the same time, the within-class variance as small as possible; it chooses β to maximize the objective function λ and is expressed by

$$\lambda(\beta) = \frac{\beta^T Q_b \beta}{\beta^T Q_w \beta}. \quad (29)$$

For this reason, differentiating λ on β and letting it be 0, we calculate

$$Q_w^{-1} Q_b \beta = \lambda \beta, \quad (30)$$

where β are the eigenvectors relating to the r eigenvalues of $Q_w^{-1} Q_b$ and maximum λ . If the eigenvalues rank by descending order, such as $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r$, then $\beta_1 \geq \beta_2 \geq \dots \geq \beta_r$ may be treated as the corresponding coefficients of the canonical variables, respectively. The first canonical score can be shown by $\text{Can}_1 = \beta_1^T s$, and it interprets that the single best linear discriminator of these g data. The second canonical score can be expressed by $\text{Can}_2 = \beta_2^T s$ and is selected in the same way but subject to additional restriction; it is uncorrelated with $\text{Can}_1 = \beta_1^T s$. It will be interpreted as it is the second-best linear discriminator of these g data.

False nearest neighbor (FNN) algorithm is employed to select dominant symptom parameters. It can pick out the most representative symptom parameters in time and frequency domains. The FNN is an algorithm proposed by Kennel et al. in 1992 [22] for calculating the embedding dimension in the chaos theory. Here, the FNN is applied to calculate the distance of different fault symptom parameters vectors. When the symptom parameters set is selected, the space dimension m will be changed and the high similarity false neighbors will be removed in the phase space reconstruction process.

In m space reconstruction, each of the space vectors is

$$s(i) = \{s(i), s(i + \tau), \dots, s(i + (m - 1)\tau)\}, \quad (31)$$

where $s(i)$ is the symptom parameter, m is the dimension, and τ is the delay time. $s(i)$ has a nearest neighbor $s^{NN}(i)$ ($s^{NN}(i)$ is the point other than $s(i)$) at a certain distance. The distance between $s(i)$ and $s^{NN}(i)$ is

$$R_m^2(i) = \|s(i) - s^{NN}(i)\|. \quad (32)$$

When the dimension changes from m to $m + 1$, the distance between two vectors will change from $R_m(i)$ to $R_{m+1}(i)$:

$$R_{m+1}(i) = \sqrt{R_m^2(i) + \|s(i + \tau m) - s^{NN}(i + \tau s)\|}. \quad (33)$$

When $R_{m+1}(i)$ is more significant than $R_m(i)$, it can be regarded that it is the false nearest neighbor in m and $m + 1$ dimensions. Otherwise, it is the nearest neighbor. The nearest neighbor means that the fault symptom parameter is

a strong correlation, so it should remove the nearest neighbor symptom parameters. Usually, the distance is used to measure the false nearest neighbor.

With the FNN algorithm, assume that a $S_{h \times l}$ high symptom parameters matrix (h is the sample number and l is the number of symptom parameters) is obtained. The FNN-CDA dominant symptom parameters method employs the CDA algorithm to obtain the canonical coefficient from $S_{h \times l}$ characteristic subspace so that it can build the canonical discriminant functions. Based on these functions, it can calculate the canonical score of symptom parameters Can_{s_i} . Then, one symptom parameter can be successively set from the original parameter set $S_{h \times l} = (s_1, s_2, \dots, s_i, \dots, s_l)$ to zero, thus obtaining a new parameter set $\hat{S}_{h \times l} = (s_1, s_2, \dots, 0, \dots, s_l)$. After the CDA operation, $\hat{S}$ can get the canonical score $\text{Can}_{\hat{s}_i}$. The false nearest neighbor algorithm is used to calculate the distance R_i between Can_{s_i} and $\text{Can}_{\hat{s}_i}$.

According to the distances in the descending order, the small distance means less contribution to fault diagnosis, and therefore, it is a nondominant symptom parameter.

2.4. Evaluation Criterion. For determining the dominant symptom parameters, an evaluation criterion is proposed in this paper. It is defined as follows.

Assume that R_i ($1 \leq i \leq l$) is the value of FNN distance and l is the number of symptom parameters. R_i is ranked as $R_1 \geq R_2 \geq \dots \geq R_l \geq 0$. Individual percentage pr_i is defined as

$$\text{pr}_i = \frac{R_i}{\sum_{i=1}^m R_i}. \quad (34)$$

The cumulative percent of the first t types in the sequence is defined as

$$\text{SR}_t = \frac{\sum_{i=1}^t R_i}{\sum_{i=1}^l R_i}, \quad t = 1, 2, \dots, l. \quad (35)$$

When the threshold SR_t is set to N_t , the ideal result will be available. If the threshold is less than N_t or more substantial than N_t , the diagnosis accuracy will decrease. If $\text{SR}_t > N_t$, it is deemed that the dominant symptom parameters include the first t symptom parameter. The threshold SR_t refers to principal component analysis theory: the percentage of cumulative sum over 80% can represent the primary information of the signal.

2.5. Framework of the Proposed Method. According to the previous introductions of the method, the schematic of the dominant symptom parameters selection bearing fault diagnosis method is shown in Figure 2:

The process of the proposed method is summarized as follows:

- (1) The signal is acquired from the condition diagnosis test experimental bench under various fault conditions (normal condition and fault conditions).
- (2) The acquired original vibration signals were filtered by the GA filtering method.

- (3) The filtered signals were employed to calculate the symptom value of all time- and frequency-domain symptom parameters. They cover the fault information in time and frequency domains.
- (4) CDA is performed to calculate the canonical discriminant function and obtain the canonical score.
- (5) Successively set the single symptom parameter from all parameters to zero and recalculate the canonical score.
- (6) Combining the canonical scores (vector) and the canonical scores of all parameters to calculate the space distance by FNN in characteristic subspace. It can evaluate the performance of symptom parameters.
- (7) Based on the distances in descending order, according to the evaluation criteria, select the dominant symptom parameters.
- (8) The SVM algorithm is performed to build the discriminative model for verifying the effectiveness.

The schematic of the proposed method is shown in Figure 2.

3. Experiments and Discussion

In this section, the results of the performed experiments are presented. The experiment in each section represents a further step to achieve the objectives described in the previous section.

3.1. Experimental Conditions. The effectiveness of the presented strategy will be validated by the analysis of three rotating machinery fault signals of the outer race fault (O), inner race fault (I), roller element fault (E), and normal signal (N).

Figure 3 shows the experimental bench for the condition diagnosis test. It consists of a loading, motor, and rotor system. The ball bearing (type: NSK 205) is applied for bearing condition diagnosis. As shown in Figure 4, the fault depth is 0.05 mm and width is 0.3 mm.

In this work, each state's original vibration signals were measured by using an accelerometer with a 50,000 Hz sampling frequency. The accelerometer is fixed in the vertical direction of the bearing. The speed of the servo motor was 1500 rpm.

3.2. Bearing Fault Diagnosis Case

3.2.1. GA-Based Filtering. The GA-based filtering requires designing a fitness function to evaluate the similarity between extracted diagnosis state noise signals and normal noise signals.

For this reason, a "statistical information evaluation function" is created in this work. The probability density distribution function (PDDF) is illustrated as follows:

FIGURE 2: Schematic of the proposed method.

FIGURE 3: Bearing fault diagnosis experimental system.

FIGURE 4: Bearing fault of outer race, inner race, and roller. (a) Outer race flaw. (b) Inner race flaw. (c) Roller flaw.

$$q_j = \frac{1}{\pi} \sqrt{\frac{\int_0^\infty (2\pi t)^2 P_x(t) dt}{\int_0^\infty P_x(t) dt e^{(-j^2/2\sigma^2)}}} \quad (36)$$

And the frequency density distribution function (FDDF) can be shown as follows:

$$p_j = \frac{q_j \sqrt{2\pi}}{2 \int_0^\infty (2\pi t)^2 P_x(t) dt} \quad (37)$$

The statistical information evaluation function can combine the PDDF with FDDF, and it is expressed as

$$\begin{aligned} I_{pq} &= \sum_{j=1}^n \frac{|\log\{(p_j^*/p_j) \times (q_j^*/q_j)\}|}{n} \\ &= \sum_{j=1}^n \frac{|\log\{(p_j^*/p_j)\}|}{n} + \sum_{j=1}^n \frac{|\log\{(q_j^*/q_j)\}|}{n} \quad (38) \\ &= I_p + I_q, \end{aligned}$$

where $p_x(t)$ is the probability density value at t ; p_j^* is the PDDF of the extracted noise signals in a diagnosis signal; p_j is the PDDF of noise signals at a normal state; q_j^* is the FDDF of the extracted noise signals at diagnosis state; q_j is the FDDF of noise signal at a normal state; and j is from 1 to n . The detailed settings can be seen in our past research achievement [1].

Parts of GA filtered signal under different bearing fault signal can be seen in Figure 5.

From the raw signal, the features are submerged in the strong background noises. After the GA-based filtering, the fault-related information can be observed in the GA-filtered signal clearly; it works as condition surveillance, and the precise fault diagnosis is shown in the later sections.

3.2.2. Symptom Parameters Extraction. The data length is 360448, and the data segments into 88. According to the sampling frequency and RPM, the subsegments contain more than one rotation signal information. When the data length of each subsegment is less than one rotation signal information, the classification accuracy will be lower. The chosen data employ the symptom parameters in equations (1)~(18) to calculate the symptom parameters values. As the data characteristic, symptom parameters values normalize the data in the range from -1 to 1 by the normalization method. For illustration, the result shown in the paper is the mean of the absolute values of symptom parameters. The values are shown in Figure 6.

As observed in Figure 6, under different fault conditions (normal and fault conditions), the symptom parameters values are different. Every symptom parameter of four conditions is shown in subplots SP1~SP18. Every subplot is different. In Subplots 17 and 18, the symptom parameter value of normal signal compared with other conditions is the biggest. In Subplots 1, 5, 8, 9, 15, and 16, inner race fault is

compared with normal, outer race fault, and roller fault, the symptom parameter being the biggest. In Subplots 10 and 11, the symptom parameter value of the outer race fault compared with other conditions is bigger. In Subplots 2, 3, 6, and 14, the symptom parameter of the roller fault compared with other conditions is the biggest. Some histograms of symptom parameters look the same (e.g., Subplots 10 and 12), but the analytical ability will be quite different.

3.2.3. Canonical Discriminant Analysis. Normal condition symptom parameters were combined with the inner race fault, outer race fault, and roller fault symptom parameters to build a data set and further to be used for CDA operation. The normal condition symptom parameters are labelled as one, inner race fault symptom parameters are labelled as two, outer race fault symptom parameters are labelled as three, and roller fault symptom parameters are labelled as four. The CDA gets the canonical function, which allows the symptom parameters and the type to establish the primary and secondary canonical (e.g., two canonical functions) correlation. Here, each bearing condition with 88 symptom parameters obtained from equations (1)~(18) was used for the CDA operation to calculate the canonical function. The primary component proportion is 71.64%, and the secondary component proportion is 24.26%.

Based on the canonical function, the symptom parameter canonical score can be calculated. The scatter diagram plotted with the canonical score can be calculated by the canonical discriminant function, as shown in Figure 7.

The result shown in Figure 7 is the canonical score of four bearing conditions. From the result, it can be seen that all of the fault conditions (inner race, outer race, and roller) are separable from the normal condition. At the same time, the canonical scores under the fault condition can also be separated from each other. The symptom parameter with canonical score distribution can have a maximum between-group difference and the minimum intergroup difference. According to the canonical score, it can get the fault canonical score vector Can_s .

Successively, set one symptom parameter from all parameters to zero and calculate the new canonical score Can_{s_0} .

The FNN is employed to select the dominant symptom parameters based on the canonical score vector Can_s , and Can_{s_0} is shown in the later section.

3.2.4. False Nearest Neighbor Analysis. The false nearest neighbor analysis was employed to calculate the distance from all symptom parameters sets and new symptom parameters (set one symptom parameter to zero). The Can_s and Can_{s_0} are brought to the algorithm to get the FNN space distance. The FNN space distance values can be seen in Table 1.

The result shown from Table 1 can be found when one of the symptom parameters is set to zero, and the distance will be sharply changed. Mapped to the dominance, it suggests that each symptom parameter may have different impacts. For Symptom Parameter 14, the distance is maximum, which indicates that it is the most sensitive. Similarly, Symptom

FIGURE 5: GA-based filtered signal. (a) Raw signal. (b) Fitness signal. (c) Filtered signal.

FIGURE 6: SP absolute values under different states (SP: symptom parameter).

Parameter 15 has a minimum correlation, being more insensitive compared to other symptom parameters. From the symptom parameters from one to eighteen, some of the time-

domain symptom parameters have a big distance, such as Symptom Parameter 4. Some frequency-domain symptom parameters have a big distance such as Symptom Parameter 13.

FIGURE 7: Canonical score of all symptom parameters (the primary component proportion is 71.64%, and the secondary component proportion is 24.26%).

3.2.5. Evaluation Criterion to Select the Dominant Symptom Parameters. For selecting the dominant symptom parameter, the evaluation criterion requires to calculate the percentage of each symptom parameter. At the same time, all the percentages are ranked in descending order to calculate the cumulative percentage. The raw percentages are shown in Figure 8, and the ranking table and cumulative value are shown in Table 2.

Figure 8 and Table 2 show the maximum and minimum distances intuitively. Table 2 shows the ranking of percentages of symptom parameters, wherein the cumulative percentage is shown in the last column. According to the evaluation criterion, when the threshold is set to 80%, it can get the optimal result. The influence of the threshold will be described in the later section. So according to the table, Symptom Parameters 14, 13, 4, 11, 1, 7, 8, 12, 18, 9, 10, and 17 are dominant symptom parameters. Other Symptom Parameters 6, 3, 5, 16, 2, and 15 are labelled as insensitive parameters and should be removed from fault diagnosis data sets.

3.2.6. SVM Fault Classification Analysis. For analyzing and verifying the performance of the selected parameters, this work used the support vector machine to build a classification model for fault diagnosis. The model used RBF as the kernel function, and the kernel parameter γ_{svm} and the penalty parameter C_{svm} of the error classification are referred in [23]. 70 of each kind of symptom parameters from the normal signal, inner race fault, outer race fault, and roller fault were used to train the model. The other 18 symptom parameters of the same kind were used for diagnosis and verification.

The results of all symptom parameters and dominant symptom parameter set from the trained SVM diagnosis classification model are shown in Figures 9 and 10.

Compared with all symptom parameters settings and the dominant symptom parameters set, it can be found that the dominant symptom parameters are brought in the model; the accuracy of the SVM model output is superior to all symptom parameters. It indicates that the

dominant symptom parameters selection can reduce the data complication and lead to better fault diagnosis results. It verifies the validity and feasibility of the algorithm.

In order to prove that the dominant parameter has the best fault diagnosis performance, here, it presents some different sets of new symptom parameters subset to test the accuracy. The result is shown in Table 3.

According to the data in Table 3, the selected dominant parameters have the best preference for diagnostic classification: the accuracy is 97.23%. When the model uses all symptom parameters, the accuracy will decline to 93.06%. The nondominant parameters have the poor ability of fault diagnosis: the accuracy is 51.39%. The first several dominant parameters result (such as numbers 6 and 7) is worse than dominant parameters. It is indicated that the selected dominant symptom parameters have super characteristics for fault diagnosis and the threshold value is appropriate.

3.3. Comparison with High-Pass Filtering. For verifying the advantage of GA filtering, this section employs high-pass filtering to test the proposed method. According to the artificial experience, the pass frequency of high-pass filtering is larger than 1500 Hz. The other operations are the same as in the case.

As the same operation given in the cases, the classified result of the different symptom parameters is shown in Table 4.

From Table 4, it can be found that the selected dominant symptom parameters have super characteristics for fault diagnosis than all symptom parameters and all classification results of high-pass filtering are worse than the GA-based filtering results. It shows the effectiveness of GA-based dominant symptom parameters.

3.4. Influence of Evaluation Criterion Threshold. For discussing the influence of the evaluation criterion threshold SR_t , this section uses the result of bearing fault diagnosis

TABLE 1: Distance calculated by FNN.

No.	$P1 = 0$	$P2 = 0$	$P3 = 0$	$P4 = 0$	$P5 = 0$	$P6 = 0$	$P7 = 0$	$P8 = 0$	$P9 = 0$	$P10 = 0$	$P11 = 0$	$P12 = 0$	$P13 = 0$	$P14 = 0$	$P15 = 0$	$P16 = 0$	$P17 = 0$	$P18 = 0$
Distance	4.350	1.905	3.406	6.274	2.721	3.439	4.287	4.164	3.974	3.688	5.968	4.116	6.708	18.069	1.280	1.929	3.466	4.082

FIGURE 8: Raw percentages of all symptom parameters.

TABLE 2: Symptom parameter cumulative percent.

No.	FNN distance	Symptom parameter	Percent (%)	Cumulative percent (%)
1	18.07	14	21.56	21.56
2	6.71	13	8.00	29.56
3	6.27	4	7.48	37.04
4	5.97	11	7.12	44.16
5	4.35	1	5.19	49.35
6	4.29	7	5.11	54.46
7	4.16	8	4.97	59.43
8	4.12	12	4.91	64.34
9	4.08	18	4.87	69.21
10	3.97	9	4.74	73.95
11	3.69	10	4.40	78.35
12	3.47	17	4.13	82.49
13	3.44	6	4.10	86.59
14	3.41	3	4.06	90.65
15	2.72	5	3.25	93.90
16	1.93	16	2.30	96.20
17	1.90	2	2.27	98.47
18	1.28	15	1.53	100.00

cases to analyze the influence. Table 5 displays the SVM classification result with dominant parameters under the different thresholds.

In Table 5, when the threshold SR_t is set to 80 or 85, the classification result is optimal. When the threshold value is larger than or less than that, the classification result will less than 97.23. So, when the threshold value set to 80% can get the optimal result, the setting of SR_t in experiment is 80.

3.5. Compariosn with Other Features Selection Methods. To further test the effectiveness of the presented strategy, the experiments have been performed to compare with non-negative matrix factorization (NMF) [13], PCA approach [24], and backward stepwise selection (BSS) and forward

FIGURE 9: All symptom parameters result (accuracy: 93.06%).

FIGURE 10: Dominant symptom parameter result (accuracy: 97.23%).

TABLE 3: Diagnosis classification result.

Number	Symptom parameter subset	Accuracy (%)
1	All symptom parameter	93.06
2	Dominant parameter	97.23
3	P6 + dominant parameter	93.06
4	P6, P3 + dominant parameter	93.06
5	P6, P3, P5, P16, P2, P15	51.39
6	P14, P13, P4, P11, P1	83.33
7	P14, P13, P4	75.00
8	P14	38.89

TABLE 4: Diagnosis classification result.

Number	Symptom parameter subset	Accuracy (%)
1	All symptom parameters	88.89
2	Dominant parameter	93.06
3	Nondominant parameter	51.39

TABLE 5: SVM classification result under different thresholds.

Threshold SR_t	SVM classification result (%)
60	87.50
65	90.28
70	90.28
75	90.28
80	97.23
85	97.23
90	93.06
95	93.06

TABLE 6: Comparison of different selection methods.

Method	Symptom parameter subset	Accuracy (%)
NMF	14, 4, 11, 2	94.44
PCA	14, 11, 9, 18, 16, 2	88.89
BSS	1, 8, 9, 10, 12, 14, 15, 17, 18	88.89
FSS	1, 8, 9, 10, 12, 14, 15, 17, 18	88.89
Our method	14, 13, 4, 11, 1, 7, 8, 12, 18, 9, 10, 17	97.23

TABLE 7: Comparison of different classification methods.

Method	Symptom parameter subset	Accuracy (%)
SVM	14, 13, 4, 11, 1, 7, 8, 12, 18, 9, 10, 17	97.23
BPNN	14, 13, 4, 11, 1, 7, 8, 12, 18, 9, 10, 17	97.23
SVM	1, 8, 9, 10, 12, 14, 15, 17, 18	88.89
BPNN	1, 8, 9, 10, 12, 14, 15, 17, 18	88.89

stepwise selection (FSS). The experiment data are the data of bearing fault diagnosis case given in Section 3.2 of the experiment. The SVM model will verify the subset of the selected parameter. The results are reported in Table 6.

The results reported in Table 6 shows that the proposed method achieves a better result than NMF and PCA methods. It demonstrated the effectiveness of the dominant symptom parameters scheme.

3.6. Comparison with Other Classification Methods. To further verify the effectiveness of the presented strategy, the experiments have been performed to compare with back-propagation neural network (BPNN). The experiment data are the data of bearing fault diagnosis case given in Section 3.2 of the experiment. The SVM and BPNN models will verify the subset of the selected parameter. The results are reported in Table 7.

The results reported in Table 7 shows that the selected dominant parameters have an influence on classification results.

4. Conclusions

This work proposes a novel dominant symptom parameters selection scheme for bearing fault diagnosis. It is composed of GA-based filtering, time- and frequency- domain symptom parameters calculation, CDA, FNN, and SVM model. The raw signal is filtered through the GA-based

method at first. Then, the canonical components are extracted by CDA. FNN was performed to calculate the distance in the CDA characteristic subspace. According to the evaluation criterion, the dominant symptom parameters can be selected. The proposed method is helpful to omit the in-depth knowledge of object properties and is also effective in eliminating parameter multicollinearity. FNN selects the most representative symptom parameters for various types of fault.

The method has been verified through roller bearing fault signals. The results show that the accuracy of dominant symptom parameters is 97.23%; it is better than 93.06% of all symptom parameters. It is an effective method for bearing fault diagnosis.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the Scientific and Technological Research Program of Chongqing Municipal Education Commission, KJZD-K201801502.

References

- [1] Z. Liao, L. Song, P. Chen, and S. Zuo, "An automatic filtering method based on an improved genetic algorithm-with application to rolling bearing fault signal extraction," *IEEE Sensors Journal*, vol. 17, no. 19, pp. 6340–6349, 2017.
- [2] T. Gong, Y. Yuan, X. Yuan, X. Wang, X. Wu, and Y. Li, "Iterative asymmetric multiscale morphology and its application to fault detection for rolling element bearing," *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, vol. 232, no. 2, pp. 316–330, 2018.
- [3] X. H. He, D. Wang, Y. F. Li, and C. H. Zhou, "A novel bearing fault diagnosis method based on Gaussian restricted Boltzmann machine," *Mathematical Problems In Engineering*, vol. 2016, Article ID 2957083, 8 pages, 2016.
- [4] M. Van, P. Franciosa, and D. Ceglarek, "Rolling element bearing fault diagnosis using integrated nonlocal means denoising with modified morphology filter operators," *Mathematical Problems In Engineering*, vol. 2016, Article ID 9657285, 14 pages, 2016.
- [5] L. Ren, W. Lv, S. Jiang, and Y. Xiao, "fault diagnosis using a joint model based on sparse representation and SVM," *IEEE Transactions on Instrumentation and Measurement*, vol. 65, no. 10, pp. 2313–2320, 2016.
- [6] Y. Yang, Y. Liao, G. Meng, and J. Lee, "A hybrid feature selection scheme for unsupervised learning and its application in bearing fault diagnosis," *Expert Systems with Applications*, vol. 38, no. 9, pp. 11311–11320, 2011.
- [7] H. Xia, J. Zhuang, and D. Yu, "Multi-objective unsupervised feature selection algorithm utilizing redundancy measure and

- negative epsilon-dominance for fault diagnosis,” *Neurocomputing*, vol. 146, pp. 113–124, 2014.
- [8] Z. Xu, Y. Li, Z. Wang, and J. Xuan, “A selective fuzzy ARTMAP ensemble and its application to the fault diagnosis of rolling element bearing,” *Neurocomputing*, vol. 182, pp. 25–35, 2016.
 - [9] J. Zheng, “Rolling bearing fault diagnosis based on partially ensemble empirical mode decomposition and variable predictive model-based class discrimination,” *Archives of Civil and Mechanical Engineering*, vol. 16, no. 4, pp. 784–794, 2016.
 - [10] L. Lu, J. Yan, and C. W. de Silva, “Dominant feature selection for the fault diagnosis of rotary machines using modified genetic algorithm and empirical mode decomposition,” *Journal of Sound and Vibration*, vol. 344, pp. 464–483, 2015.
 - [11] X. Zhang, W. Chen, B. Wang, and X. Chen, “Intelligent fault diagnosis of rotating machinery using support vector machine with ant colony algorithm for synchronous feature selection and parameter optimization,” *Neurocomputing*, vol. 167, pp. 260–279, 2015.
 - [12] B. Li, P.-l. Zhang, H. Tian, S.-s. Mi, D.-s. Liu, and G.-q. Ren, “A new feature extraction and selection scheme for hybrid fault diagnosis of gearbox,” *Expert Systems with Applications*, vol. 38, no. 8, pp. 10000–10009, 2011.
 - [13] L. Liang, F. Liu, M. Li, K. He, and G. Xu, “Feature selection for machine fault diagnosis using clustering of non-negation matrix factorization,” *Measurement*, vol. 94, pp. 295–305, 2016.
 - [14] X. Yu, F. Dong, E. Ding, S. Wu, and C. Fan, “Rolling bearing fault diagnosis using modified LFDA and EMD with sensitive feature selection,” *Ieee Access*, vol. 6, pp. 3715–3730, 2018.
 - [15] M. Luo, C. Li, X. Zhang, R. Li, and X. An, “Compound feature selection and parameter optimization of ELM for fault diagnosis of rolling element bearings,” *ISA Transactions*, vol. 65, pp. 556–566, 2016.
 - [16] V. Sugumaran, V. Muralidharan, and K. I. Ramachandran, “Feature selection using decision tree and classification through proximal support vector machine for fault diagnostics of roller bearing,” *Mechanical Systems and Signal Processing*, vol. 21, no. 2, pp. 930–942, 2007.
 - [17] Y. Li, X. Liang, J. Lin, Y. Chen, and J. Liu, “Train axle bearing fault detection using a feature selection scheme based multi-scale morphological filter,” *Mechanical Systems And Signal Processing*, vol. 101, pp. 435–448, 2018.
 - [18] M. Zeng, Y. Yang, J. Zheng, and J. Cheng, “Maximum margin classification based on flexible convex hulls for fault diagnosis of roller bearings,” *Mechanical Systems and Signal Processing*, vol. 66–67, pp. 533–545, 2016.
 - [19] S. J. Dong, D. H. Sun, B. P. Tang, Z. Y. Gao, W. T. Yu, and M. Xia, “A Fault diagnosis method for rotating machinery based on PCA and morlet kernel SVM,” *Mathematical Problems In Engineering*, vol. 2014, Article ID 293878, 8 pages, 2014.
 - [20] R. Pu and D. Liu, “Segmented canonical discriminant analysis of in situ hyperspectral data for identifying 13 urban tree species,” *International Journal of Remote Sensing*, vol. 32, no. 8, pp. 2207–2226, 2011.
 - [21] N. Strachan and L. Kell, “A potential method for the differentiation between haddock fish stocks by computer vision using canonical discriminant analysis,” *ICES Journal Of Marine Science*, vol. 52, no. 1, pp. 145–149, 1995.
 - [22] M. B. Kennel, R. Brown, and H. D. I. Abarbanel, “Determining embedding dimension for phase-space reconstruction using a geometrical construction,” *Physical Review A*, vol. 45, no. 6, pp. 3403–3411, 1992.
 - [23] K. Zhu, X. Song, and D. Xue, “A roller bearing fault diagnosis method based on hierarchical entropy and support vector machine with particle swarm optimization algorithm,” *Measurement*, vol. 47, pp. 669–675, 2014.
 - [24] A. Malhi and R. X. Gao, “PCA-based feature selection scheme for machine defect classification,” *IEEE Transactions on Instrumentation and Measurement*, vol. 53, no. 6, pp. 1517–1525, 2004.