

## Research Article

# Feature Extraction Based on Adaptive Multiwavelets and LTSA for Rotating Machinery Fault Diagnosis

Na Lu <sup>1</sup>, Guangtao Zhang,<sup>2</sup> Zhihuai Xiao <sup>3</sup>, and Om Parkash Malik <sup>4</sup>

<sup>1</sup>School of Water Conservancy & Environment, Zhengzhou University, Zhengzhou, China

<sup>2</sup>Rundian Energy Science and Technology Co. Ltd., Zhengzhou 450052, China

<sup>3</sup>Department of Power and Mechanical Engineering, Wuhan University, Wuhan 430072, China

<sup>4</sup>Department of Electrical and Computer Engineering, University of Calgary, Calgary, Canada

Correspondence should be addressed to Na Lu; [pyzhaocool@126.com](mailto:pyzhaocool@126.com)

Received 23 May 2018; Accepted 18 November 2018; Published 22 January 2019

Academic Editor: Vadim V. Silberschmidt

Copyright © 2019 Na Lu et al. 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.

Feature extraction is a key procedure in the fault diagnosis of rotating machinery. To obtain fault features with lower dimensionality and higher sensitivity, a feature extraction method based on adaptive multiwavelets transform (AMWT) and local tangent space alignment (LTSA) is proposed in this paper. AMWT is first used to obtain multiple features from the vibration signals of the machine under test to form a high-dimensional feature set. Then, in order to avoid the adverse effect of the irrelevant features in this high-dimensional feature set on the fault diagnosis result, a detection index (DI) is investigated to evaluate the sensitivity of the features and those with lower sensitivity are removed. After that, LTSA is applied for feature fusion to reduce the redundant features in the high-dimensional feature set. To validate the proposed method, performance of four feature extraction schemes based on (i) wavelet and LTSA, (ii) Geronimo, Hardin, and Massopust (GHM) multiwavelets and LTSA, (iii) AMWT and principal component analysis (PCA), and (iv) AMWT and multidimensional scaling (MDS) is compared with the proposed method. The feature extraction results by these methods are then fed into K-medoids classifier to discriminate the faults. The results show that the proposed method can improve the sensitivity of the extracted features and obtain higher fault recognition rate.

## 1. Introduction

Rotating machinery is widely used in the modern industry [1, 2]. Faults occurring in rotating machinery may lead to fatal breakdowns. Therefore, it is important to diagnose the existence of the rotating machinery faults accurately at an early stage to avoid huge maintenance cost and catastrophic accidents [3, 4].

Feature extraction is a pivotal procedure in the process of fault diagnosis [5]. The application of an effective feature extraction method may improve the accuracy of fault diagnosis results significantly. Being rich in information on machinery health conditions, vibration signals are often processed by advanced signal processing techniques to extract features from the rotating machinery [6, 7]. Up to now, multiple methods have been extensively applied for

vibration signal processing, such as fast Fourier transform (FFT), short-time Fourier transform (STFT) [8], wavelet transform (WT) [9], Wigner-Ville distribution (WVD) [10], empirical mode decomposition (EMD) [11], multiwavelets transform (MWT) [12], etc. However, in spite of all the achievements these methods have made, they are fixed and independent of the measured signals. Hence, the sensitivity of the extracted features could be hardly enhanced further.

As an extension of WT and MWT, adaptive multiwavelets transform (AMWT) not only has perfect local property in time and frequency domain but also possesses properties of orthogonality, symmetry, short support, and higher order of vanishing moments, simultaneously. More importantly, it can change adaptively according to the characteristics of the vibration signals and enhance the

sensitivity of the extracted fault features [13]. Therefore, AMWT constructed based on two-scale similarity transform (TST) [14] has been investigated in this paper and applied to feature extraction for fault diagnosis of rotating machinery.

In order to know the machinery condition most comprehensively, many features are generally extracted from vibration signal beforehand to form a high-dimensional feature set. However, irrelevant or redundant features, that have adverse effect on the accuracy of fault diagnosis result, are often contained in this feature set. Additionally, applying such a high dimensional feature set directly for rotating machinery fault diagnosis is computationally demanding [15]. Therefore, investigating both feature selection and feature fusion methods to reduce irrelevant and redundant features, respectively, and obtain fault features with lower dimensionality and higher sensitivity is significant in increasing the accuracy of fault diagnosis results and reducing the complexity of the fault diagnosis algorithm.

Generally, feature selection can be divided into two categories: filters and wrappers [16]. Filters conduct feature selection depending on information of the feature subset [17], whereas wrappers generally employ the accuracy of the classifier result to evaluate the feature subset [18]. Compared with wrappers, filters are independent of classifiers, thus need less computational time and possess good generalizability [16]. Therefore, filters are considered in this paper. In this process, an effective criterion for feature sensitivity evaluation is essential to obtain more sensitive features from the high features set. Detection index (DI) [14, 19, 20] is employed in this paper to evaluate the rotating machinery vibration signal features for feature selection.

There are many classical feature fusion methods for redundant features reduction such as principal component analysis (PCA) [21] and multidimensional scaling (MDS) [22]. These methods are easy to be implemented. However, as linear methods, they fail to capture nonlinear characteristic of the dataset and thus are likely to exhibit poor performance for nonlinear feature fusion [23–26]. In recent years, many manifold learning techniques have been investigated, such as locally linear embedding (LLE) [27], isometric feature mapping (IsoMap) [28], local tangent space alignment (LTSA) [29], etc. Among them, LTSA aims to preserve local structures in small neighborhoods and discovers the intrinsic features of nonlinear manifolds [30]. It has been proven to be superior to LLE and IsoMap in obtaining intrinsic data structure [31, 32] and applied in vibration signals feature extraction [23, 33, 34].

In this paper, a new feature extraction method based on AMWT and LTSA is proposed. With this method, efforts have been made in three aspects in order to extract fault features with lower dimensionality and higher sensitivity from rotating machinery signals. Firstly, considering AMWT has the ability to change adaptively with signal characteristics, it is employed to improve the synthetic sensitivity of the candidate high-dimensional feature set. Secondly, with the evaluation index DI, feature selection is conducted to eliminate the irrelevant features from the

original high-dimensional feature set. Thirdly, feature fusion is implemented with LTSA to remove the redundant features from the obtained feature subset by the feature selection step. Results of the studies show that, with the proposed method, fault features with lower dimensionality and higher sensitivity can be obtained, and the accuracy of the rotating machinery fault diagnosis can be improved.

The rest of this paper is arranged as follows. In Section 2, theories of AMWT and LTSA are introduced briefly. The specific process of the proposed method is illustrated in Section 3. After that, effectiveness of the proposed method in feature extraction by application in rotating machinery fault diagnosis is demonstrated in Section 4. Finally, conclusions are described in Section 5.

## 2. Adaptive Multiwavelets and LTSA

*2.1. Adaptive Multiwavelets.* Adaptive multiwavelets have been applied in rotating machinery fault diagnosis and obtained good results [13, 14]. The TST-based adaptive multiwavelets are investigated in this paper for rotating machinery vibration fault feature extraction. They are constructed on the basis of multiwavelets and TST theory.

*2.1.1. Multiwavelets.* Different from scalar wavelet, multiwavelets [35] contain vector multiscaling functions  $\Phi(t)$  and multiwavelet functions  $\Psi(t)$  formed by several scaling functions  $\{\phi_r, r \in N\}$  and wavelet functions  $\{\psi_r, r \in N\}$ , respectively:

$$\begin{aligned}\Phi(t) &= [\phi_1(t), \phi_2(t), \dots, \phi_r(t)]^T, \\ \Psi(t) &= [\psi_1(t), \psi_2(t), \dots, \psi_r(t)]^T.\end{aligned}\quad (1)$$

Similar to the two-scale relations of scalar wavelet, multiwavelet functions and multiscaling functions of multiwavelets satisfy two-scale matrix equations:

$$\Phi(t) = \sqrt{2} \sum_{k=0}^M \mathbf{H}_k \Phi(2t-k) \quad k \in \mathbb{Z}, \quad (2)$$

$$\Psi(t) = \sqrt{2} \sum_{k=0}^M \mathbf{G}_k \Phi(2t-k), \quad (3)$$

where  $\{\mathbf{H}_k\}$  and  $\{\mathbf{G}_k\}$  represent  $r \times r$  two-scale matrix sequences.

In frequency domain, the two-scale matrix equations can be expressed as

$$\begin{aligned}\widehat{\Phi}(\omega) &= \mathbf{H}\left(\frac{\omega}{2}\right)\widehat{\Phi}\left(\frac{\omega}{2}\right), \\ \widehat{\Psi}(\omega) &= \mathbf{G}\left(\frac{\omega}{2}\right)\widehat{\Phi}\left(\frac{\omega}{2}\right),\end{aligned}\quad (4)$$

where  $\widehat{\Phi}(\omega)$  and  $\widehat{\Psi}(\omega)$  represent Fourier transformations of the multiscaling functions and multiwavelet functions, respectively;  $\mathbf{H}(\omega)$  and  $\mathbf{G}(\omega)$  represent the two-scale matrix symbols corresponding to the multiscaling and multiwavelet functions, respectively:

$$\mathbf{H}(\omega) = \frac{1}{2} \sum_{k=0}^M \mathbf{H}_k e^{-i\omega k}, \quad (5)$$

$$\mathbf{G}(\omega) = \frac{1}{2} \sum_{k=0}^M \mathbf{G}_k e^{-i\omega k}.$$

With the two-scale matrix equations (2) and (3), the decomposition and reconstruction of multiwavelets can be obtained as (6) and (7), respectively:

$$\mathbf{v}_{j,k} = \sum_n \mathbf{H}_{n-2k} \mathbf{v}_{j-1,n}, \quad (6)$$

$$\mathbf{w}_{j,k} = \sum_n \mathbf{G}_{n-2k} \mathbf{v}_{j-1,n}$$

$$\mathbf{v}_{j-1,k} = \sum_n \mathbf{H}_{k-2n}^* \mathbf{v}_{j,n} + \sum_n \mathbf{G}_{k-2n}^* \mathbf{w}_{j,n}, \quad (7)$$

where the superscript \* represents the complex conjugate transpose.

Since the two-scale matrix sequences are  $r \times r$  matrices, two or more input streams are required in the process of multiwavelet decomposition. Hence, preprocessing is needed to mapping a sequence of univariate data to  $r$  vectors. Among the proposed preprocessing methods, the oversampling scheme has been proven to be more effective for feature extraction than others [36]. Therefore, it is adopted in this paper.

**2.1.2. TST.** Two-scale similarity transform theory was first proposed in [37]. In this, TST was divided into singular TST and regular TST. Singular TST can generate multiwavelets with approximation order one higher than the original multiwavelets. Therefore, it attracts more attention. However, the obtained multiwavelets do not satisfy orthogonality. To make up this drawback, the singular TST was extended to biorthogonal multiwavelets [38].

Suppose  $\Phi$ ,  $\Psi$ ,  $\tilde{\Phi}$ , and  $\tilde{\Psi}$  are the multiscaling functions, multiwavelet functions, dual multiscaling functions, and dual multiwavelet functions, respectively,  $\mathbf{H}(\omega)$ ,  $\mathbf{G}(\omega)$ ,  $\tilde{\mathbf{H}}(\omega)$ , and  $\tilde{\mathbf{G}}(\omega)$  are their corresponding two-scale matrix symbols,  $\tilde{\mathbf{H}}(\omega)$  has approximation order of at least one, and  $\mathbf{M}(\omega)$  is a TST matrix. Then, the biorthogonal multiwavelets singular TST is

$$\mathbf{H}_{\text{new}}(\omega) = \frac{1}{2} \mathbf{M}(2\omega) \mathbf{H}(\omega) \mathbf{M}^{-1}(\omega), \quad (8)$$

$$\mathbf{G}_{\text{new}}(\omega) = \mathbf{G}(\omega) \mathbf{M}^{-1}(\omega).$$

Correspondingly, the biorthogonal multiwavelets singular inverse TST is

$$\begin{aligned} \tilde{\mathbf{H}}_{\text{new}}(\omega) &= 2\mathbf{M}^*(2\omega) \tilde{\mathbf{H}}(\omega) \mathbf{M}^*(\omega), \\ \tilde{\mathbf{G}}_{\text{new}}(\omega) &= \tilde{\mathbf{G}}(\omega) \mathbf{M}^*(\omega). \end{aligned} \quad (9)$$

**2.1.3. Adaptive Multiwavelets.** Adaptive multiwavelets are constructed by adding some free parameters to the TST matrix  $\mathbf{M}(\omega)$  in TST process of GHM multiwavelets [39].

The GHM multiwavelets have the properties of orthogonality, short support, symmetry, and approximation order of 2. The two-scale matrix sequences of GHM multiwavelets are

$$\mathbf{H}_0 = \begin{bmatrix} \frac{3}{5} & \frac{4\sqrt{2}}{5} \\ -\frac{1}{10\sqrt{2}} & -\frac{3}{10} \end{bmatrix},$$

$$\mathbf{H}_1 = \begin{bmatrix} \frac{3}{5} & 0 \\ \frac{9}{10\sqrt{2}} & 1 \end{bmatrix},$$

$$\mathbf{H}_2 = \begin{bmatrix} 0 & 0 \\ \frac{9}{10\sqrt{2}} & -\frac{3}{10} \end{bmatrix},$$

$$\mathbf{H}_3 = \begin{bmatrix} 0 & 0 \\ -\frac{1}{10\sqrt{2}} & 0 \end{bmatrix}, \quad (10)$$

$$\mathbf{G}_0 = \frac{1}{10} \begin{bmatrix} -\frac{1}{\sqrt{2}} & -3 \\ 1 & 3\sqrt{2} \end{bmatrix},$$

$$\mathbf{G}_1 = \frac{1}{10} \begin{bmatrix} \frac{9}{\sqrt{2}} & -10 \\ -9 & 0 \end{bmatrix},$$

$$\mathbf{G}_2 = \frac{1}{10} \begin{bmatrix} \frac{9}{\sqrt{2}} & -3 \\ 9 & -3\sqrt{2} \end{bmatrix},$$

$$\mathbf{G}_3 = \frac{1}{10} \begin{bmatrix} -\frac{1}{\sqrt{2}} & 0 \\ -1 & 0 \end{bmatrix}.$$

The two multiscaling functions and two multiwavelet functions of GHM are shown in Figure 1.

In order to construct adaptive multiwavelets, two TST matrices with free parameters are needed. They are

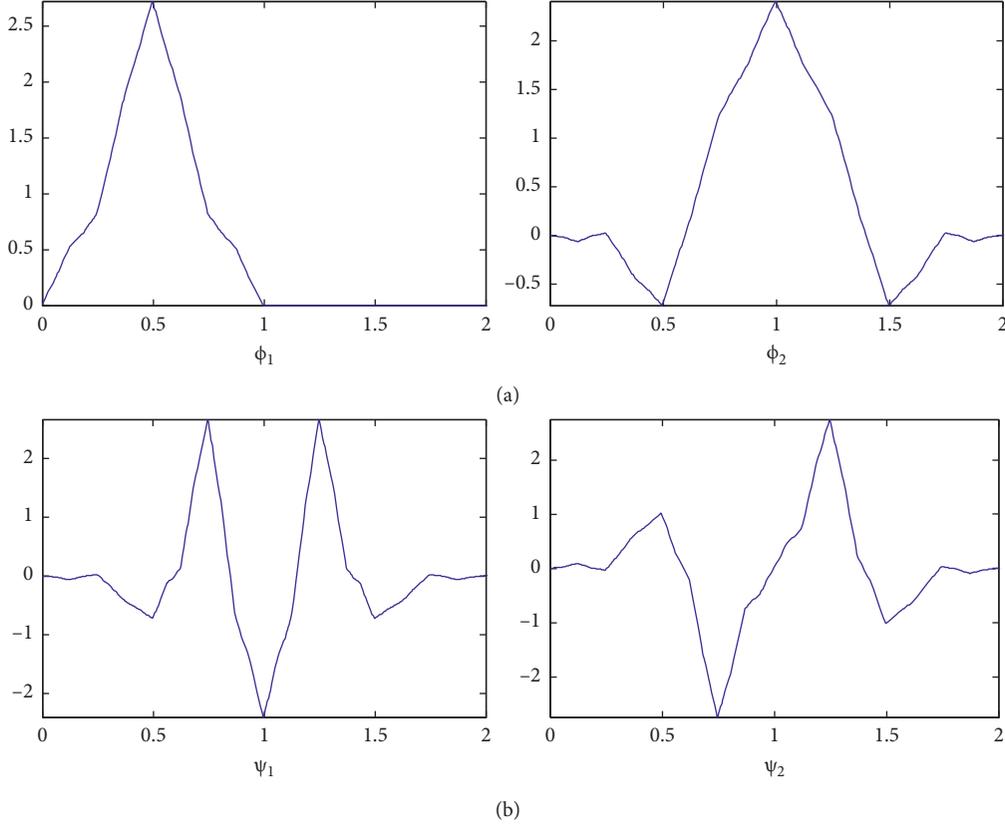


FIGURE 1: GHM multiwavelets: (a) multiscaling functions; (b) multiwavelet functions [14].

$$\mathbf{M}_1(\omega) = \begin{bmatrix} a(1 + e^{-i\omega}) & -2\sqrt{2}a \\ b(1 - e^{-i\omega}) & 0 \end{bmatrix}, \quad (11)$$

$$\mathbf{M}_2(\omega) = \begin{bmatrix} c & 0 \\ d(1 + e^{-i\omega}) & f(1 - e^{-i\omega}) \end{bmatrix}. \quad (12)$$

Substituting (11) into (8) and (9), new biorthogonal multiwavelets with two-scale matrix symbols of  $\mathbf{H}_{\text{new}}(\omega)$ ,  $\mathbf{G}_{\text{new}}(\omega)$ ,  $\tilde{H}_{\text{new}}(\omega)$ , and  $\tilde{G}_{\text{new}}(\omega)$  can be obtained. Then, taking  $\mathbf{H}_{\text{new}}(\omega)$ ,  $\mathbf{G}_{\text{new}}(\omega)$ ,  $\tilde{H}_{\text{new}}(\omega)$ , and  $\tilde{G}_{\text{new}}(\omega)$  as the original two-scale matrix symbols  $\tilde{H}(\omega)$ ,  $\tilde{G}(\omega)$ ,  $\mathbf{H}(\omega)$ , and  $\mathbf{G}(\omega)$  in (8) and (9), respectively, and substituting (12) into (8) and (9), the final biorthogonal multiwavelets with two-scale matrix symbols of  $\mathbf{H}_{\text{final}}(\omega)$ ,  $\mathbf{G}_{\text{final}}(\omega)$ ,  $\tilde{H}_{\text{final}}(\omega)$ , and  $\tilde{G}_{\text{final}}(\omega)$  can be obtained. Since the TST matrices shown as (11) and (12) are applied in this process, the two-scale matrix symbols  $\mathbf{H}_{\text{final}}(\omega)$ ,  $\mathbf{G}_{\text{final}}(\omega)$ ,  $\tilde{H}_{\text{final}}(\omega)$ , and  $\tilde{G}_{\text{final}}(\omega)$  are formed by free parameters  $a$ ,  $b$ ,  $c$ ,  $d$ , and  $f$ . Therefore, the final obtained biorthogonal multiwavelets can adaptively change their basis functions with the characteristics of signals.

**2.2. LTSA.** As a fundamental type of manifold learning method, LTSA employs locally linear transformation to extract low-dimensional manifold structure embedded in high-dimensional data and is capable of digging out valuable information and realizing feature dimension reduction. The

algorithm of LTSA generally includes the following three steps [29]:

- (1) Neighborhood identification: for each data point  $x_i \in R^D$  in sample set  $\mathbf{X} = \{x_1, x_2, \dots, x_N\}$ , identify its  $k$  neighborhood samples with K-nearest neighbor algorithm to construct neighborhood  $\mathbf{X}_{Ni} = \{x_{i1}, x_{i2}, \dots, x_{ik}\}$ .
- (2) Local information extraction: project the  $k$  points in  $\mathbf{X}_{Ni}$  into a tangent space of the manifold at  $x_i$ :

$$\theta_{ij} = \mathbf{Q}_i^T(x_{ij} - \bar{x}_{Ni}), \quad j = 1, 2, \dots, k, \quad (13)$$

where  $\bar{x}_{Ni} = (1/k)\mathbf{X}_{Ni}e_k$  is the neighborhood mean vector,  $e_k$  is a column vector with all ones;  $\theta_{ij}$  represents the local coordinate;  $\mathbf{Q}_i$  denotes the orthogonal basis of the tangent space which can be estimated by performing the singular decomposition of  $\mathbf{X}_{Ni} - \bar{x}_{Ni}e_k^T$ ; it consists of  $d$  left singular vectors of  $\mathbf{X}_{Ni} - \bar{x}_{Ni}e_k^T$  corresponding to the first  $d$  largest singular values.

- (3) Global coordinate determination: the global coordinate system can be obtained by linearly aligning the local coordinates:

$$T_{ij} = \bar{T}_{ij} + \mathbf{L}_i\theta_{ij} + \varepsilon_{ij} \quad j = 1, 2, \dots, k, \quad (14)$$

where  $\mathbf{L}_i \in R^{d \times d}$  is an affine transformation matrix,  $\varepsilon_{ij}$  is the reconstruction error,  $T_{ij}$  is the global coordinate, and  $\bar{T}_{ij}$  is the centralized vector of  $T_{ij}$ . Expressing (14) as matrix form, we get

$$\mathbf{T}_i = \frac{1}{k} \mathbf{T}_i e_k e_k^T + \mathbf{L}_i \Theta_i + \mathbf{E}_i, \quad (15)$$

where  $\mathbf{T}_i = [T_{i1}, T_{i2}, \dots, T_{ik}]$ ,  $\Theta_i = [\theta_{i1}, \theta_{i2}, \dots, \theta_{ik}]$ , and  $\mathbf{E}_i = [\varepsilon_{i1}, \varepsilon_{i2}, \dots, \varepsilon_{ik}]$ .

To preserve as much of the local geometry in the  $d$ -dimensional space, the reconstruction error needs to be minimized:

$$\min \sum_i \|\mathbf{E}_i\|_F^2 = \min \sum_i \left\| \mathbf{T}_i - \frac{1}{k} \mathbf{T}_i e_k e_k^T - \mathbf{L}_i \Theta_i \right\|_F^2, \quad (16)$$

where  $\|\bullet\|_F$  represents the matrix Frobenius norm. Obviously, the affine transformation matrix should be given by  $\mathbf{L}_i = \mathbf{T}_i (\mathbf{I} - (1/k) e_k e_k^T) \Theta_i^+$  to minimize the reconstruction error, where  $\Theta_i^+$  represents the Moore–Penrose generalized inverse of  $\Theta_i$ .

Let  $\mathbf{S}_i$  be the 0-1 selection matrix and  $\mathbf{W}_i = (\mathbf{I} - (1/k) e_k e_k^T) (\mathbf{I} - \Theta_i^+ \Theta_i)$  be the weight matrix of  $x_i$ , then

$$\min \sum_i \|\mathbf{E}_i\|_F^2 = \min \text{tr}(\mathbf{T} \mathbf{S} \mathbf{W} \mathbf{W}^T \mathbf{S}^T \mathbf{T}^T) = \min \text{tr}(\mathbf{T} \mathbf{B} \mathbf{T}^T), \quad (17)$$

where  $\mathbf{S} = [\mathbf{S}_1, \mathbf{S}_2, \dots, \mathbf{S}_n]$ ,  $\mathbf{W} = \text{diag}(\mathbf{W}_1, \mathbf{W}_2, \dots, \mathbf{W}_i)$ , and  $\mathbf{B} = \mathbf{S} \mathbf{W} \mathbf{W}^T \mathbf{S}^T$ .

In order to determine the global coordinate, let  $\mathbf{T} \mathbf{T}^T = \mathbf{I}$ . Then, the optimal global  $d$ -dimensional coordinate  $\mathbf{T}$  is obtained by the  $d$  eigenvectors of  $\mathbf{B}$ , corresponding to the second to  $d + 1$  smallest eigenvalues of  $\mathbf{B}$ .

### 3. Feature Extraction Based on Adaptive Multiwavelets and LTSA

The feature extraction method for rotating machinery vibration fault diagnosis mainly includes three steps. First, a high-dimensional candidate feature set is constructed via AMWT. Second, DI is used as the feature sensitivity evaluation index to remove irrelevant features from the candidate feature set and form a subset of vibration features. Third, LTSA is applied to the fusion of the redundant features in the feature subset to obtain favorable vibration features with lower dimensionality and higher sensitivity.

**3.1. High-Dimensional Feature Set Construction Based on Adaptive Multiwavelets.** To improve the sensitivity of the features in the original high-dimensional feature set, AMWT, rather than conventional fixed signal processing method, is applied to process the vibration signals in this paper. The maximum value of SDI is used as the optimal object, and genetic algorithm (GA) is used to find the optimal multiwavelets from the adaptive multiwavelets library. Then, the optimal multiwavelets are used to extract the values of some nondimensional symptom parameters

(NSPs) from the vibration signal to form the original high-dimensional features set.

**3.1.1. Summary of SDI.** SDI is an index formed by DI for feature sensitivity evaluation [19]. Suppose that  $s_1$  and  $s_2$  are values of a fault feature calculated from the signals measured in machinery state 1 and state 2, respectively. They obey the normal distributions  $N(\mu_1, \sigma_1)$  and  $N(\mu_2, \sigma_2)$ ,  $\mu_2 \geq \mu_1$ . Then, DI corresponding to these two states is

$$\text{DI} = \frac{\mu_2 - \mu_1}{\sqrt{\sigma_1^2 + \sigma_2^2}} \quad (18)$$

Suppose the number of features used for diagnosis and machinery states are  $M$  and  $N$ , respectively. Then, the SDI can be expressed as

$$\text{SDI} = \sum_{i=1}^{N-1} \sum_{j=i+1}^N \sum_{k=1}^M \frac{|\mu_{ik} - \mu_{jk}|}{\sqrt{\sigma_{ik}^2 + \sigma_{jk}^2}} \quad (19)$$

According to [19], the larger the value of DI, the more sensitive is the fault feature. In addition, it is obvious from (18) and (19) that SDI is the sum of the DI over all of the fault features and fault types. Therefore, increasing the value of SDI can improve the sensitivity of the fault feature set comprehensively.

**3.1.2. Types of NSPs.** To construct the high-dimensional candidate feature set, the following ten NSPs are taken into consideration:

$$P_1 = \frac{\sigma}{f}, \quad (20)$$

$$P_2 = \frac{\sum_{i=1}^N f_i^2}{\sigma^2}, \quad (21)$$

$$P_3 = \frac{\sum_{i=1}^N (f_i - \bar{f})^3}{N \sigma^3}, \quad (22)$$

$$P_4 = \frac{\sum_{i=1}^N (f_i - \bar{f})^4}{N \sigma^4}, \quad (23)$$

$$P_5 = \frac{\bar{f}_p}{f}, \quad (24)$$

$$P_6 = \frac{\bar{f}_p}{\sigma}, \quad (25)$$

$$P_7 = \frac{\left| \sum_{i=1}^{N_p} (f_{pi} - \bar{f}_p)^3 \right|}{N_p \sigma_p^3}, \quad (26)$$

$$P_8 = \frac{\left| \sum_{i=1}^{N_p} (f_{pi} - \bar{f}_p)^4 \right|}{N_p \sigma_p^4}, \quad (27)$$

$$P_9 = \frac{\left| \sum_{i=1}^{N_v} (f_{vi} - \bar{f}_v)^3 \right|}{N_v \sigma_v^3}, \quad (28)$$

$$P_{10} = \frac{\left| \sum_{i=1}^{N_v} (f_{vi} - \bar{f}_v)^4 \right|}{N_v \sigma_v^4}, \quad (29)$$

where  $f_i$  denotes the value of sampling point,  $f_{pi}$  and  $f_{vi}$  represent the peak and valley of  $f_i$ , respectively;  $\bar{f}$ ,  $\bar{f}_p$ , and  $\bar{f}_v$  represent the means of  $f_i$ ,  $f_{pi}$ , and  $f_{vi}$ , respectively;  $\sigma$ ,  $\sigma_p$ , and  $\sigma_v$  represent the standard deviations of  $f_i$ ,  $f_{pi}$ , and  $f_{vi}$ , respectively.

**3.1.3. High-Dimensional Features Set Construction.** In order to avoid the adverse effect of the noise component in signals on the extracted features, the original signals are first processed by GHM neighborhood coefficients denoising method [36] before input into the algorithm for the construction of high-dimensional features set. The flowchart of this algorithm is shown in Figure 2. Its specific steps are also given as follows:

- (1) Input the data denoised by GHM neighborhood coefficients denoising method.
- (2) Determine the decomposition level of the adaptive multiwavelets and assign it to variable  $n$ .
- (3) Choose the coefficients scale from the  $2(n+1)$  scales of coefficients of the adaptive multiwavelets for optimization.
- (4) Assign some values to the adaptive parameters  $a$ ,  $b$ ,  $c$ ,  $d$ , and  $f$ , and substitute them into (11) and (12) to get the TST matrices.
- (5) Perform two steps of biorthogonal multiwavelets singular TST and biorthogonal multiwavelets singular inverse TST with (8) and (9) to get new biorthogonal multiwavelets.
- (6) Preprocess the input data by oversampling preprocessing scheme [40].
- (7) Decompose the preprocessed data with the newly constructed biorthogonal multiwavelets in step (5).
- (8) Calculate the value of SDI corresponding to the chosen scale of coefficients with (19).
- (9) Compare the iteration number with the preassigned maximum step of the GA algorithm. If the former is smaller, turn to (4), otherwise, save the free parameters  $a$ ,  $b$ ,  $c$ ,  $d$ , and  $f$  and proceed to the next step.
- (10) Determine if all coefficients scales have been chosen. If so, proceed to the next step, otherwise, turn to step (3).
- (11) Use the obtained  $2(n+1)$  optimized multiwavelets to process the input data.
- (12) Extract the 10 NSPs in (20) to (29) from each of the chosen scale of coefficients corresponding to the  $2(n+1)$  optimized multiwavelets, respectively.
- (13) Output the  $10 * 2(n+1)$  features obtained in step (12) to form a high-dimensional candidate feature set.

**3.2. Feature Selection with DI.** By adaptive multiwavelets, an original feature set can be constructed. However, it contains multiple irrelevant and redundant features, which will have adverse effect on fault diagnosis.

According to Section 3.1.1, DI can be used as an index to evaluate the sensitivity of the fault features. Therefore, feature selection with DI is conducted to remove the irrelevant features from the constructed original feature set. Generally, different scales of coefficients contain different frequency components of the signal. In order to consider fault features in each frequency band, the 10 NSPs corresponding to each scale of multiwavelets coefficients are evaluated by DI, and then the 2 better NSPs are selected. The feature selection scheme is illustrated in Figure 3. After this process,  $4(n+1)$  NSPs can be obtained.

**3.3. Feature Fusion Based on LTSA.** With feature selection method, irrelevant features can be excluded from the high dimensional set. However, there still exist some redundant features in the obtained feature subset, which may have negative influence on the fault diagnosis process. Therefore, feature fusion should be implemented on the feature subset to obtain fault features with lower dimensionality and higher sensitivity.

Compared with the traditional PCA and MDS feature fusion methods, LTSA has illustrated better performance on processing nonlinear signals which are the common form of signals measured from rotating machinery. Hence, it is exploited in this paper to process the feature subset for feature fusion. In this process, GA is taken as the optimizing algorithm to select proper parameters of the neighborhood sample number  $k$  and the lower dimensionality  $d$  for LTSA algorithm, which have remarkable influence on the feature extraction result. As mentioned in Section 3.1.1, the value of DI can be used to evaluate the sensitivity of the features on discriminating two machinery states. Suppose that there are  $M$  machinery states to be recognized, then  $M(M-1)/2$  DI are related to each feature. Among these DI, the smallest one determines the comprehensive sensitivity of the feature. Therefore, the maximum value of the smallest DI is taken as the optimal object to search for the optimal parameters  $k$  and  $d$ . The flowchart of the feature fusion algorithm is illustrated in Figure 4. The specific steps of this algorithm are also illustrated as follows:

- (1) Input the features subset with the  $4(n+1)$  NSPs obtained in the feature selection process.
- (2) Assign some values to the parameters  $k$  and  $d$  for LTSA algorithm.
- (3) Use the LTSA algorithm to process the  $4(n+1)$  NSPs.
- (4) Calculate the DI corresponding to each obtained feature in the lower-dimensional space.
- (5) Determine if the maximum step reached. If so, continue to the next step. Otherwise, turn to step (2).
- (6) Output the  $d$  dimension features extracted by the optimal LTSA algorithm.

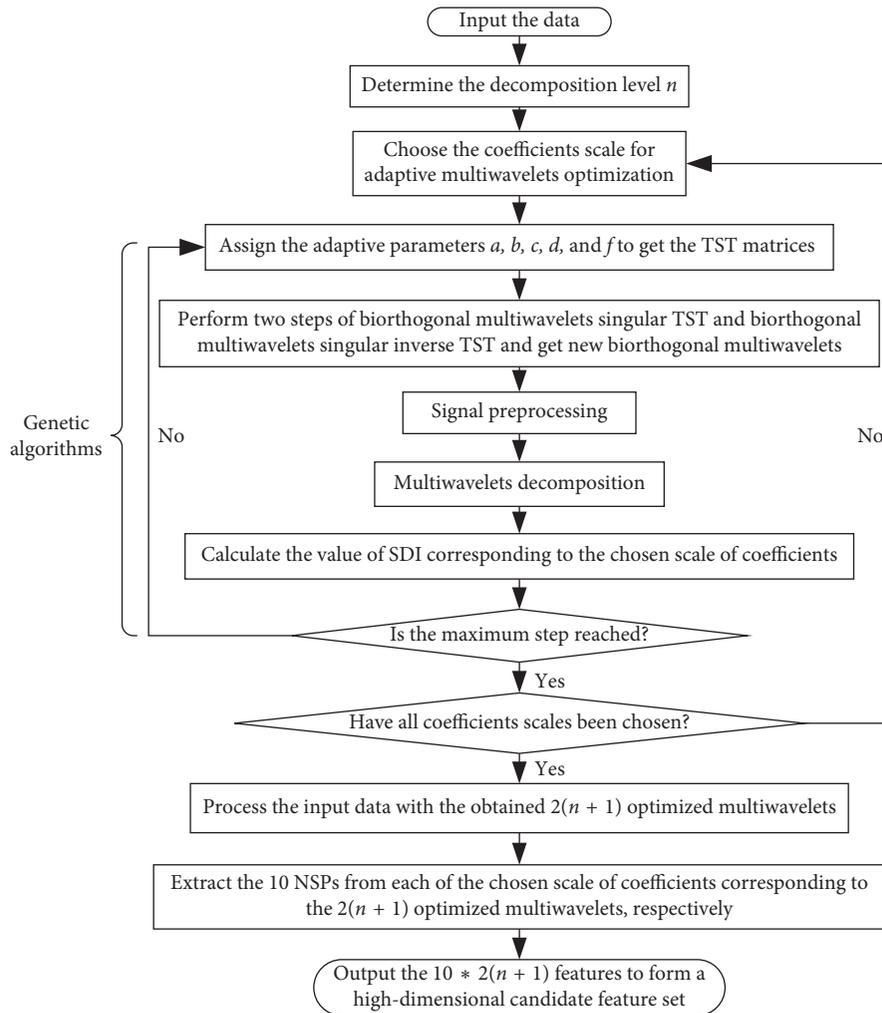


FIGURE 2: Flowchart of the algorithm for the construction of high-dimensional features set.

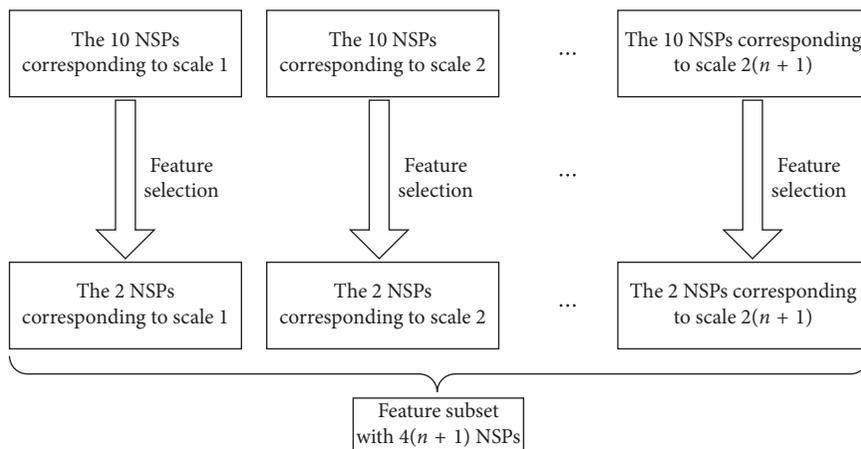


FIGURE 3: Feature selection scheme.

### 4. Application Experiments

To verify the feature extraction capability of the proposed method, two datasets, rotating machinery platform simulation data and hydroelectric generator unit data, are used for test.

*4.1. Rotating Machinery Platform Simulation Data.* The rotating machinery platform shown in Figure 5 is applied to collect vibration signals for the feature extraction experiment. The rotor of this platform is driven by a DC motor which is controlled by a speed controller. It is composed of 2

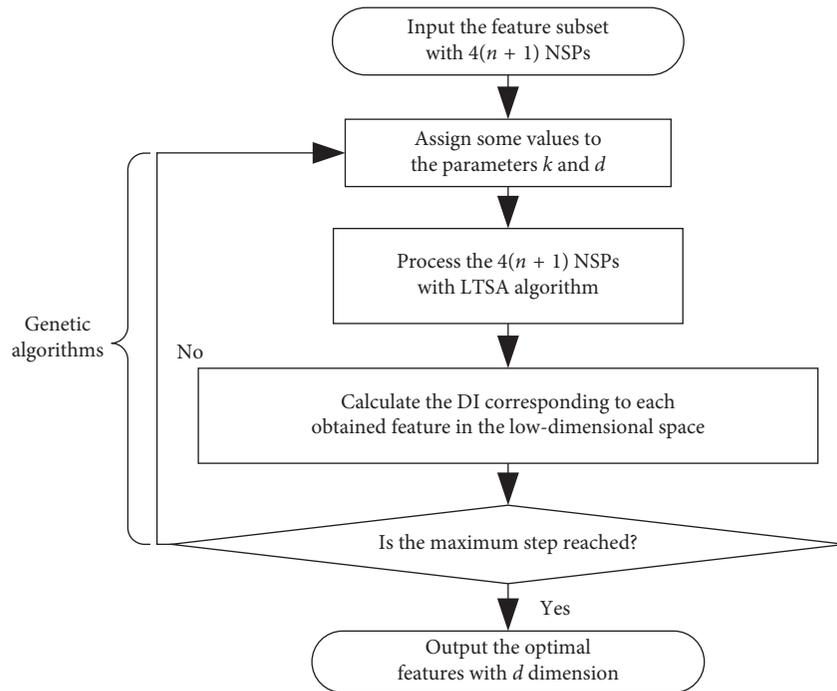


FIGURE 4: Flowchart of the feature fusion algorithm.

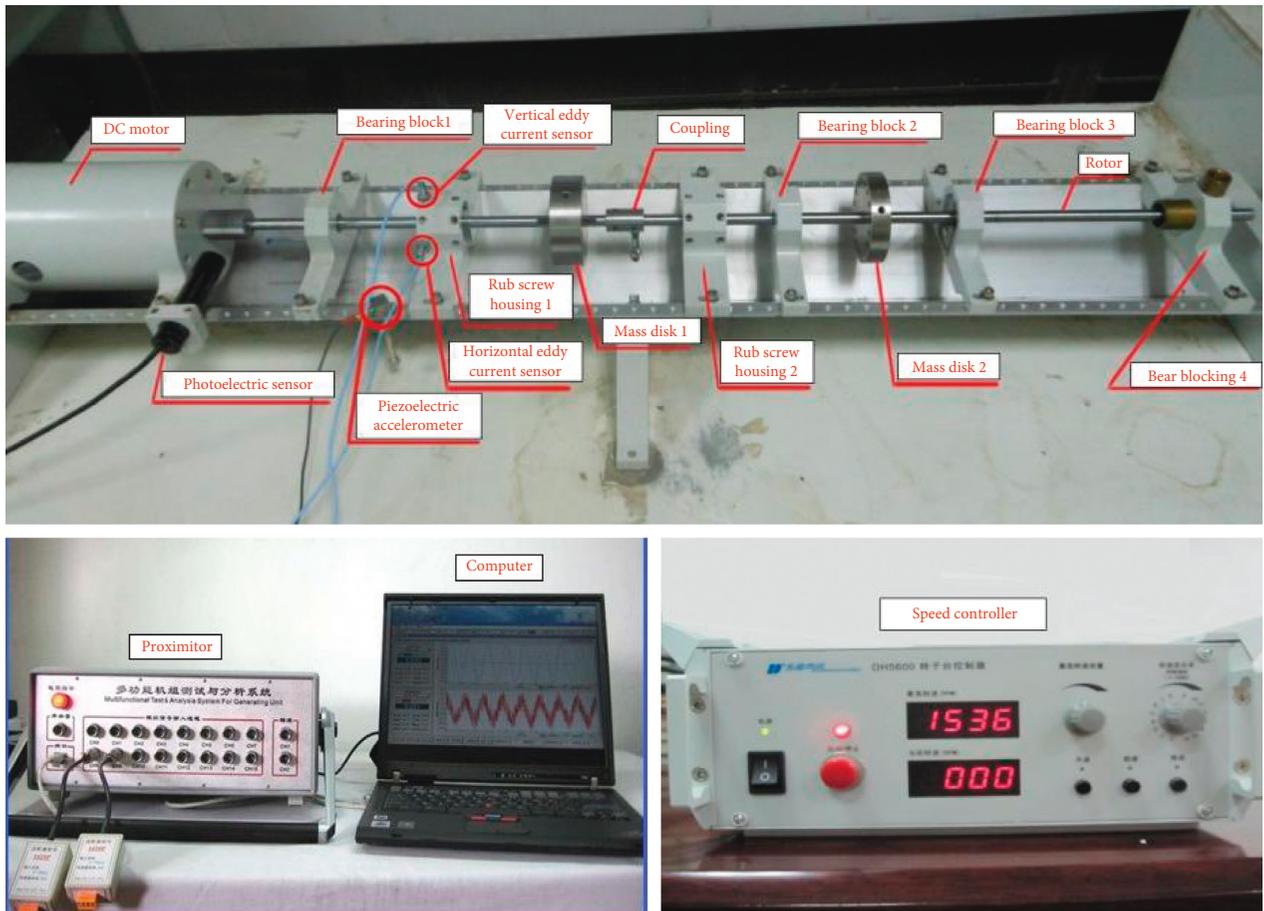


FIGURE 5: Rotating machinery platform [14].

single shafts coupled together and supported by 4 bearing blocks. 2 mass disks are fixed on the rotor and 2 rub screw housings are installed on the rack of this system. The vibration signals measured by the sensors are sent to a proximator for filtering and amplification and then transmitted to the computer for storage and analysis.

The test was performed on the condition that the speed of the platform was set at 1200 rpm, and the sample frequency was set to 2048 Hz. Four common classes of rotating machinery samples including normal, unbalance, misalignments, and rotor-to-stator rub samples were collected from the platform. For each of these classes, 20 samples with sample points 2048 were applied for the feature extraction method verification. Typical waveforms of the 4 kinds of samples are presented in Figure 6.

To verify the effectiveness of the proposed method, the acquired data were processed by GHM neighborhood coefficients denoising method first and then input into the algorithm in Section 3.1 to construct a high-dimensional candidate feature set. In this process, settings of some relevant parameters are listed in Table 1. After the high-dimensional feature set construction, 8 optimal multiwavelets and 80 features were obtained. The values of the SDI and free parameters  $a$ ,  $b$ ,  $c$ ,  $d$ , and  $f$  corresponding to these optimal multiwavelets are shown in Table 2.

To exclude the irrelevant features in the high-dimensional feature set, the obtained 80 features were processed by the feature selection scheme described in Section 3.2. Take the optimal multiwavelets corresponding to the A01 scale of multiwavelet coefficients as an example. The DI values corresponding to the A01 scale optimal multiwavelets are shown in Table 3. It can be seen in Table 3 that the DI values of NSP4 and NSP6 are significantly better than those of other features. Therefore, NSP4 and NSP6 are selected from the 10 features corresponding to the A01 scale optimal multiwavelets. Similarly, 2 features were selected from each of the 10 features corresponding to all the 8 optimal multiwavelets to form a 16-dimensional feature subset. The obtained 16 features corresponding to each 20 samples of the 4 machinery states are shown in Table 4.

Input the 16-dimensional feature subset into the method described in Section 3.3 for feature fusion. Set the neighborhood sample number  $k$ , the lower dimensionality  $d$ , and the relevant parameters in GA to be the values listed in Table 1. Then, the final fault features with the optimal dimensionality 3 and the optimal neighborhood sample number 6 were extracted. These features are illustrated in Figure 7, where  $\circ$ ,  $\times$ ,  $+$ , and  $*$  represent the features of normal samples, unbalance samples, misalignment samples, and rub samples, respectively. It can be seen from Figure 7 that with the proposed AMWT- and LTSA-based feature extraction method, the 4 rotating machinery states can be distinguished clearly.

Considering that GHM multiwavelets are the basis of the AMWT, features of rotating machinery were extracted based on GHM multiwavelets and LTSA for comparison. In this process, the AMWT was substituted with GHM multiwavelets in the high-dimensional feature set construction step. The feature selection step and feature fusion step are the

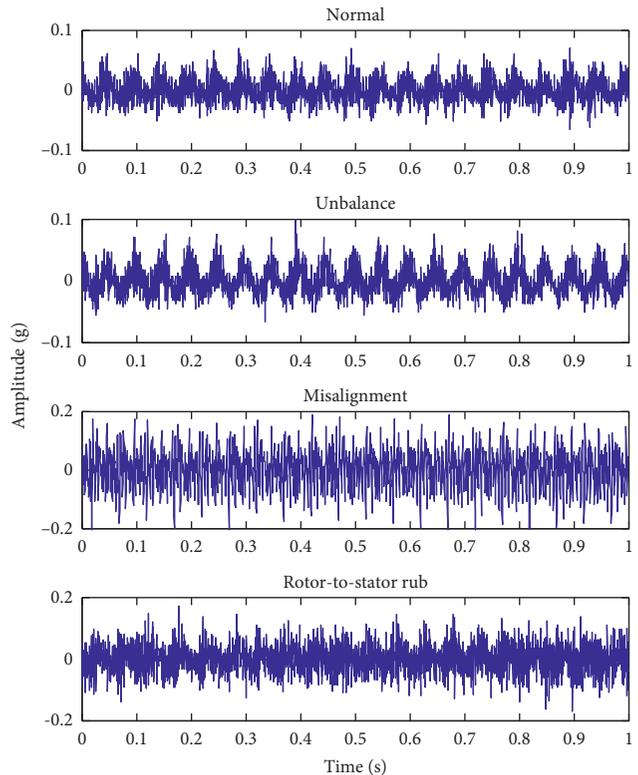


FIGURE 6: Acquired signals from the rotating machinery platform.

TABLE 1: Settings of the parameters used for rotating machinery platform experiment.

Parameters	Values
Decomposition level, $n$	3
Adaptive parameters, $a, b, c, d, f$	$[-10, 10]$ except 0
Neighborhood sample number, $k$	Integer within $[3, 16]$
Lower dimensionality, $d$	Integer within $[2, 6]$
Population scale in GA	30
Evolution generation in GA	60
Probability of crossover in GA	0.7
Probability of mutation in GA	0.02

same as those of the proposed method. By this method, the fault features with the optimal dimensionality 5 and the optimal neighborhood sample number 9 were extracted. Similarly, features were also extracted based on Db4 wavelet and LTSA for comparison. And, the fault features with the optimal dimensionality 5 and the optimal neighborhood sample number 13 were extracted.

In addition, other 2 comparative experiments were also carried out to verify the effectiveness of the feature extraction method based on AMWT and LTSA. One substituted LTSA with PCA, and the other substituted LTSA with MDS method in the proposed method. The three-dimensional extracted features by these two comparative experiments are illustrated in Figures 8 and 9, respectively. It can be seen from Figure 8, the features extracted by the AMWT and PCA methods under misalignment and rub conditions are mixed together, and the features under normal and unbalance conditions are very close to each

TABLE 2: Values of the SDI and free parameters  $a$ ,  $b$ ,  $c$ ,  $d$ , and  $f$  corresponding to the optimal multiwavelets for rotating machinery platform experiment.

Scale	A01	A02	D01	D02	D11	D12	D21	D22
SDI	125.67	119.70	135.81	76.93	175.49	108.34	166.59	105.30
$a$	-1.10	-0.61	-3.81	-8.26	-1.17	-2.98	-8.50	4.56
$b$	7.86	8.93	3.46	7.41	-4.16	-6.76	-3.46	-3.90
$c$	-3.17	-1.06	-4.74	-5.18	-3.00	0.96	-3.33	-5.68
$d$	9.48	-0.08	-4.64	-1.26	5.53	5.69	7.25	5.51
$f$	1.03	3.30	7.15	9.86	0.89	4.59	-1.17	1.85

TABLE 3: DI values corresponding to the A01 scale optimal multiwavelets for rotating machinery platform experiment.

DI	NSP <sub>1</sub>	NSP <sub>2</sub>	NSP <sub>3</sub>	NSP <sub>4</sub>	NSP <sub>5</sub>	NSP <sub>6</sub>	NSP <sub>7</sub>	NSP <sub>8</sub>	NSP <sub>9</sub>	NSP <sub>10</sub>
DI <sub>1-2</sub>	0.31	0.39	6.52	<b>9.51</b>	0.34	<b>7.39</b>	2.78	0.96	3.01	1.23
DI <sub>1-3</sub>	0.15	0.33	1.01	<b>8.99</b>	0.21	<b>3.04</b>	1.69	0.07	3.67	1.57
DI <sub>1-4</sub>	0.16	0.63	1.59	<b>3.30</b>	0.18	<b>1.19</b>	0.02	0.31	0.12	0.31
DI <sub>2-3</sub>	0.54	0.51	2.10	<b>5.47</b>	0.51	<b>27.96</b>	1.19	1.36	0.64	0.87
DI <sub>2-4</sub>	0.30	0.67	0.80	<b>0.52</b>	0.28	<b>3.84</b>	1.89	0.69	2.14	1.24
DI <sub>3-4</sub>	0.44	0.52	0.67	<b>2.99</b>	0.45	<b>0.72</b>	1.16	0.29	2.52	1.47

TABLE 4: Sixteen-dimensional feature subset obtained by feature selection scheme for rotating machinery platform experiment.

Features number	1	2	3	4	...	15	16	
Normal	1	1.39	1.68	1.39	3.01	...	0.59	0.49
	2	1.35	1.63	1.35	2.87	...	0.59	0.49
	...	...	...	...	...	...	...	...
Unbalance	20	1.34	1.66	1.37	2.77	...	0.58	0.47
	1	1.67	1.53	1.74	2.34	...	1.18	0.38
	2	1.67	1.51	1.76	2.04	...	1.32	0.39
Misalignment	...	...	...	...	...	...	...	...
	20	1.67	1.51	1.75	2.20	...	1.09	0.40
	1	2.25	2.19	2.68	2.55	...	0.83	0.72
Rub	2	2.16	2.05	2.64	2.25	...	0.88	0.72
	...	...	...	...	...	...	...	...
	20	2.04	2.49	2.64	2.35	...	0.93	0.75
Rub	1	1.59	1.66	1.73	3.34	...	0.44	0.59
	2	1.72	1.67	1.86	2.33	...	0.55	0.63
	...	...	...	...	...	...	...	...
20	1.77	1.77	1.87	2.83	...	0.44	0.66	

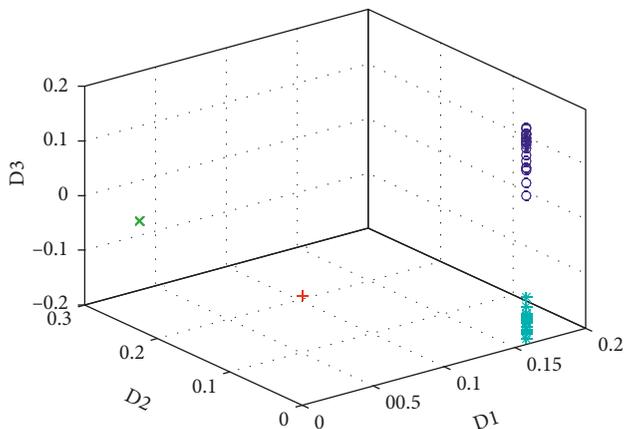


FIGURE 7: Extracted fault features by the proposed method based on AMWT and LTSA for rotating machinery platform experiment.

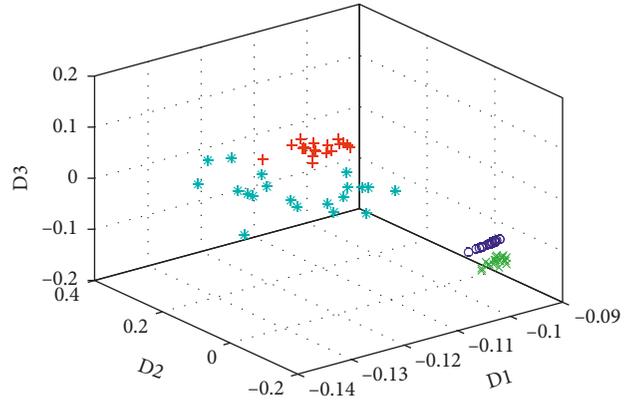


FIGURE 8: Extracted fault features by the method based on AMWT and PCA for rotating machinery platform experiment.

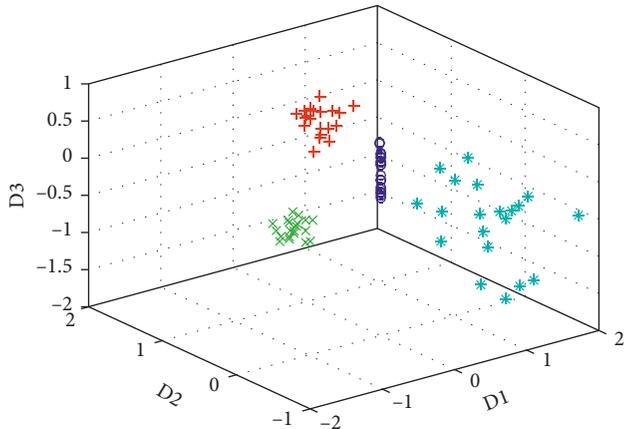


FIGURE 9: Extracted fault features by the method based on AMWT and MDS for rotating machinery platform experiment.

other. According to Figure 9, the 4 kinds of samples can be distinguished roughly; however, the distribution of the features is much scattered.

To verify the effectiveness of the proposed method further, the features extracted by the proposed method and the 4 comparative methods were input into clustering algorithm, respectively, for fault classification. According to [41], K-medoids clustering is more robust to noise and outliers than K-means clustering. Therefore, it is applied in this paper. The fault discrimination results are demonstrated in Table 5.

It can be seen from Table 5 that the 4 rotating machinery states can be classified by the proposed method, but some of these machinery states cannot be entirely identified by the other comparative methods. This indicates that the proposed method can be applied to extract higher sensitive fault features for rotating machinery.

**4.2. Hydroelectric Generator Unit Data.** The effectiveness of proposed method was also verified by extracting features from vibration signals of practical rotating machinery, a hydroelectric generator unit in a hydropower plant in China. The sketch of this hydroelectric generator unit structure is shown in Figure 10. It is composed of a generator, a hydroturbine, 2 brackets, a thrust bearing, and 3 guide bearings.

It was noticed on a certain day that the unit vibrated so intensely that the amplitude of the vibration signals exceeded the safety threshold. The specialist identified the fault pattern as unbalance fault. After performing counterweight process to the unit, the amplitude of the vibration signals returned to normal.

The signals acquired before and after the counterweight process by the accelerometer mounted on the lower bracket are taken as signals under unbalance and normal conditions, respectively. In the operation process of the unit, the speed was 500 rpm and the sampling frequency was set to 500 Hz. For each of the 2 machinery conditions, 16 samples with sample points 512 acquired from the unit were applied for the feature extraction method verification. Typical waveforms of the 2 kinds of samples are presented in Figure 11.

Similar to the rotating machinery platform experiment, the raw signals acquired from the hydroelectric generator unit were firstly input into the GHM neighborhood coefficients denoising method for noise cancellation and then processed by the algorithm in Section 3.1 to construct a high-dimensional candidate feature set with 80 features. In this process, setting of the parameters is the same as those for the rotating machinery platform shown in Table 1. The values of the SDI and optimal parameters  $a$ ,  $b$ ,  $c$ ,  $d$ , and  $f$  corresponding to the obtained 8 optimal multiwavelets are shown in Table 6.

By inputting these 80 features into the method described in Section 3.2 for irrelevant features cancellation, 16 features were obtained. These features for each 16 samples of the 2 machinery states are shown in Table 7.

The obtained 16 features were processed with the feature fusion method in Section 3.3 to remove the redundant features. Then, the final fault features with the optimal dimensionality 2 and the optimal neighborhood sample number 10 were extracted. They are shown in Figure 12. It

can be seen that the features of the two classes can be distinguished clearly.

The 4 comparative experiments were also carried out to verify the effectiveness of the feature extraction method. By the method based on GHM and LTSA, the fault features with the optimal dimensionality 2 and the optimal neighborhood sample number 15 were extracted. The 2-dimensional extracted features are shown in Figure 13. According to Figure 13, some of the obtained two kinds of features are mixed together so they cannot be distinguished clearly. By the method based on Db4 and LTSA, the fault features with optimal dimensionality 4 and the optimal neighborhood sample number 9 were extracted. By the other two methods based on AMWT and PCA and on AMWT and MDS, the 2-dimensional fault features shown in Figures 14 and 15, respectively, can be obtained. It can be seen in Figures 14 and 15 that the two classes of samples can be separated. However, the distribution of points is relatively dispersed.

Inputting the features extracted by the proposed method and the 4 comparative methods into the K-medoids method, the discrimination results shown in Table 8 can be obtained. According to Table 8, the discrimination rate by the proposed method and the two comparative methods based on AMWT and PCA and on AMWT and MDS are all 100%, which are higher than the results obtained by other two comparative experiments. Nevertheless, comparing Figure 12 with Figures 14 and 15, it can be seen that the two classes of samples in Figures 14 and 15 are relatively dispersed. Therefore, the proposed method can be applied to extract fault features for rotating machinery effectively.

## 5. Conclusions

Rotating machinery is commonly used in the modern industry. Developing an efficient feature extraction method is a real necessity for accurate diagnosis of the fault type and avoids catastrophe accidents and huge economic losses. In this paper, a novel feature extraction scheme based on AMWT and LTSA is proposed for fault diagnosis of rotating machinery. To obtain fault features with lower dimensionality and higher sensitivity, three main steps are implemented. Firstly, the high-dimensional NSPs are extracted from vibration signals with AMWT. Because the AMWT has the capability to change adaptively with the characteristics of the signals, the comprehensive sensitivity of the constructed high-dimensional feature set is improved. Then, the DI is used to evaluate the sensitivity of the NSPs in the original high-dimensional candidate feature set and remove the irrelevant features. By this step, the dimensionality of the feature set can be reduced substantially and features with higher sensitivity can be picked out. After that, LTSA is investigated to exclude the redundant features and reduce the dimensionality of the feature space further. In this process, GA is applied to optimize the free parameters in LTSA so as to obtain an optimal feature fusion result. In order to verify the effectiveness of the proposed method, fault feature extraction

TABLE 5: Fault discrimination rate for rotating machinery platform experiment.

Method	Normal (%)	Unbalance (%)	Misalignment (%)	Rotor-to-stator rub (%)	Total (%)
AMWT + LTSA	100	100	100	100	100
GHM + LTSA	100	100	60	5	66.25
Db4 + LTSA	55	5	45	95	50
AMWT + PCA	100	100	100	65	91.25
AMWT + MDS	100	100	100	85	96.25

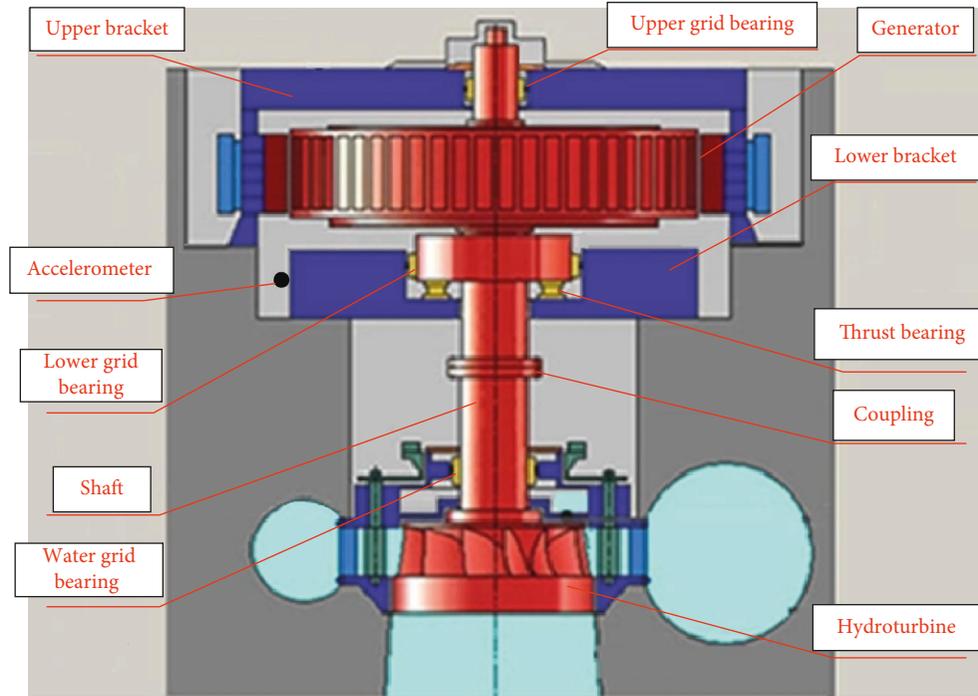


FIGURE 10: Sketch of the hydroelectric generator unit structure.

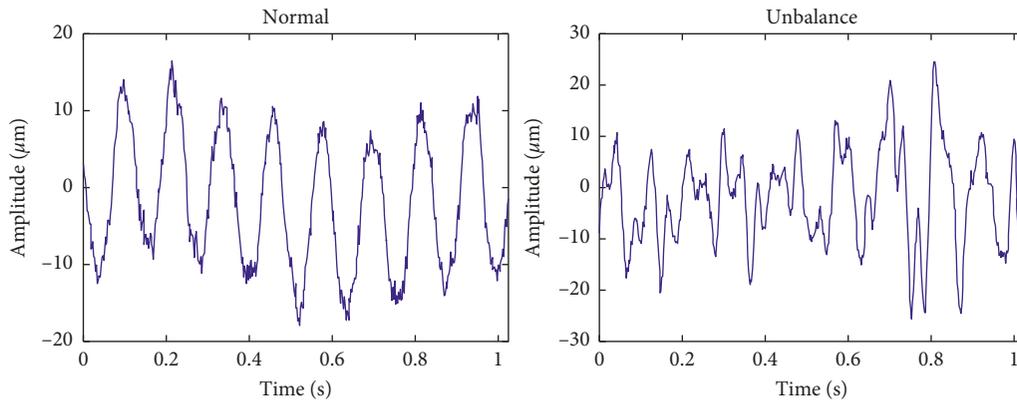


FIGURE 11: Signals of the hydroelectric generator unit.

TABLE 6: Values of the SDI and free parameters  $a$ ,  $b$ ,  $c$ ,  $d$ , and  $f$  corresponding to the optimal multiwavelets for hydroelectric generator unit experiment.

Scale	A01	A02	D01	D02	D11	D12	D21	D22
SDI	7.56	7.12	7.24	8.91	8.43	16.66	10.72	11.45
$a$	-2.67	4.32	-1.09	4.52	-9.48	-6.42	-6.80	3.03
$b$	5.13	-7.51	-8.82	-0.17	-0.20	-6.29	-7.62	1.18
$c$	-5.50	-1.29	-7.95	5.09	-8.94	-7.18	-0.16	9.60
$d$	-3.32	-9.87	-7.12	0.15	0.01	8.67	2.13	0.78
$f$	-1.39	-7.54	1.81	9.34	1.45	2.67	0.75	-0.00

TABLE 7: Sixteen-dimensional feature subset obtained by feature selection scheme for hydroelectric generator unit experiment.

Features number	1	2	3	4	...	15	16	
Normal	1	0.04	1.56	0.21	1.96	...	2.64	0.60
	2	0.05	1.66	0.19	1.92	...	3.14	0.55
	...	...	...	...	...	...	...	...
Unbalance	16	0.00	1.51	0.17	1.75	...	2.44	0.73
	1	0.44	3.11	0.25	3.05	...	2.84	0.99
	2	0.03	3.22	0.26	2.53	...	3.03	1.12
...	...	...	...	...	...	...	...	...
16	0.18	3.62	0.87	2.91	...	2.49	1.00	

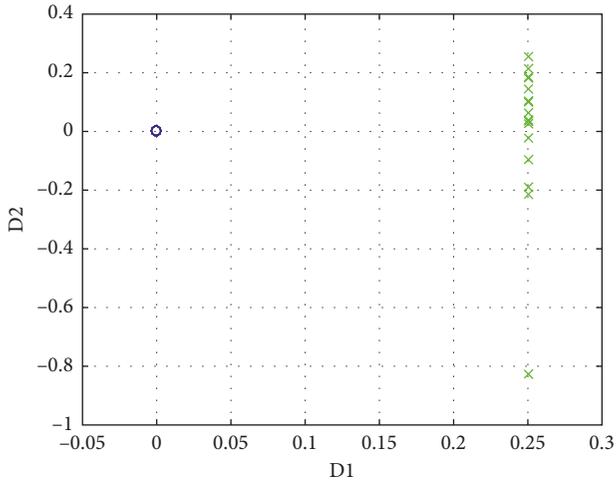


FIGURE 12: Extracted fault features by the proposed method based on AMWT and LTSA for hydroelectric generator unit experiment.

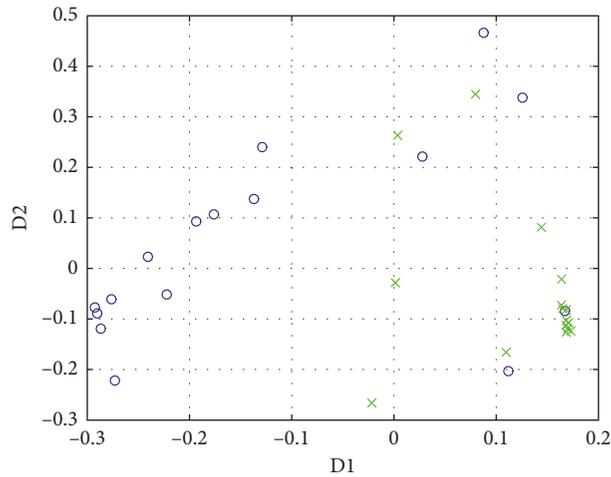


FIGURE 13: Extracted fault features by the method based on GHM and LTSA for hydroelectric generator unit experiment.

tests were carried out with the data acquired from a rotating machinery platform and a hydroelectric generator unit. Experimental results show that the proposed method has the capability to obtain fault features with lower dimensionality and higher sensitivity and thus can improve the fault classification performance.

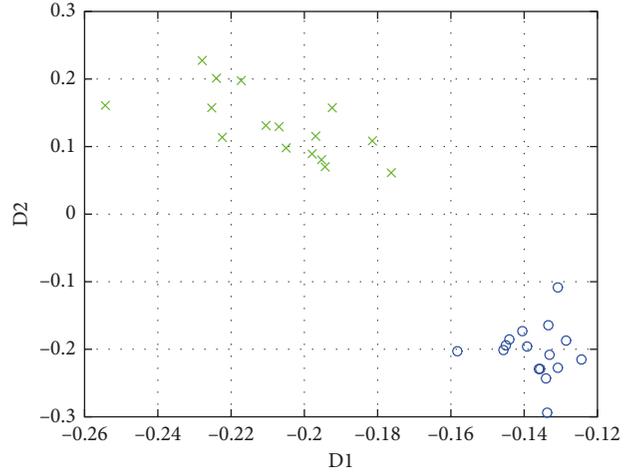


FIGURE 14: Extracted fault features by the method based on AMWT and PCA for hydroelectric generator unit experiment.

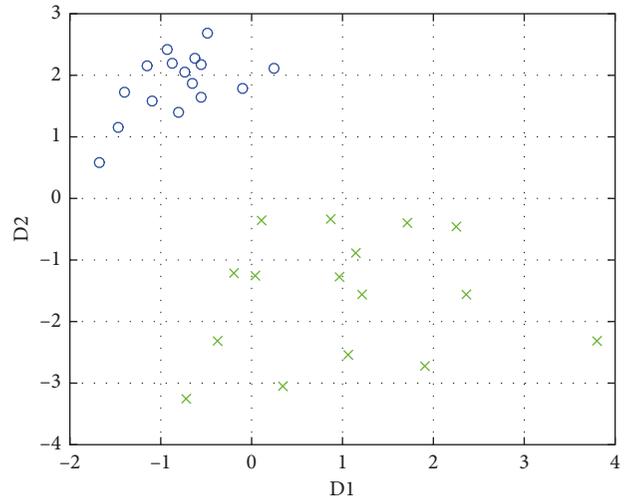


FIGURE 15: Extracted fault features by the method based on AMWT and MDS for hydroelectric generator unit experiment.

TABLE 8: Fault discrimination rate for hydroelectric generator unit experiment.

Method	Normal (%)	Unbalance (%)	Total (%)
AMWT + LTSA	100	100	100
GHM + LTSA	87.5	87.5	87.5
Db4 + LTSA	62.5	93.75	78.125
AMWT + PCA	100	100	100
AMWT + MDS	100	100	100

### Data Availability

The data are attached in the “Optional Supplementary Materials” named as “supp.1201084.xlsx.”

### Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

The authors acknowledge the financial support from the National Natural Science Foundation of China under Grant No. 51609203 and from the Xi'an Science and Technology Planning Project under Grant No. NC1504(1).

## Supplementary Materials

Four common classes of rotating machinery vibration signals including normal, unbalance, misalignment, and rotor-to-stator rub samples collected from the rotating machinery platform described in the paper are included in the data file. For each of these classes, there are 20 samples with sample points 2048. There are four sheets in this file which including the normal, unbalance, misalignment, and rotor-to-stator rub samples, respectively. (*Supplementary Materials*)

## References

- [1] A. Muszynska, *Rotor Dynamics*, CRC Press-Taylor & Francis Group, Boca Raton, FL, USA, 2005.
- [2] M. L. Adams, *Rotating Machinery Vibration from Analysis to Troubleshooting*, Marcel Dekker, New York, NY, USA, 2000.
- [3] L. Liang, F. Liu, M. Li, and G. Xu, "Manifold learning with self-organizing mapping for feature extraction of nonlinear faults in rotating machinery," *Mathematical Problems in Engineering*, vol. 2015, Article ID 873905, 11 pages, 2015.
- [4] S. Dong, L. Chen, B. Tang, X. Xu, Z. Gao, and J. Liu, "Rotating machine fault diagnosis based on optimal morphological filter and local tangent space alignment," *Shock and Vibration*, vol. 2015, Article ID 893504, 9 pages, 2015.
- [5] J. Li, Y. Zhang, and P. Xie, "A new adaptive cascaded stochastic resonance method for impact features extraction in gear fault diagnosis," *Measurement*, vol. 91, pp. 499–508, 2016.
- [6] W. Li, Z. Zhu, F. Jiang, G. Zhou, and G. Chen, "Fault diagnosis of rotating machinery with a novel statistical feature extraction and evaluation method," *Mechanical Systems and Signal Processing*, vol. 50-51, pp. 414–426, 2015.
- [7] Y. Lei and M. J. Zuo, "Gear crack level identification based on weighted K nearest neighbor classification algorithm," *Mechanical Systems and Signal Processing*, vol. 23, no. 5, pp. 1535–1547, 2009.
- [8] F. Liu, B. He, Y. Liu, S. Lu, Y. Zhao, and J. Zhao, "Phase space similarity as a signature for rolling bearing fault diagnosis and remaining useful life estimation," *Shock and Vibration*, vol. 2016, Article ID 6127479, 12 pages, 2016.
- [9] J. Rangel-Magdaleno, H. Peregrina-Barreto, J. Ramirez-Cortes, R. Morales-Caporal, and I. Cruz-Vega, "Vibration analysis of partially damaged rotor bar in induction motor under different load condition using DWT," *Shock and Vibration*, vol. 2016, Article ID 3530464, 11 pages, 2016.
- [10] J.-D. Wu and C.-K. Huang, "An engine fault diagnosis system using intake manifold pressure signal and Wigner-Ville distribution technique," *Expert Systems with Applications*, vol. 38, no. 1, pp. 536–544, 2011.
- [11] Y. Lei, J. Lin, Z. He, and M. J. Zuo, "A review on empirical mode decomposition in fault diagnosis of rotating machinery," *Mechanical Systems and Signal Processing*, vol. 35, no. 1-2, pp. 108–126, 2013.
- [12] X. Wang, Y. Zi, and Z. He, "Multiwavelet denoising with improved neighboring coefficients for application on rolling bearing fault diagnosis," *Mechanical Systems and Signal Processing*, vol. 25, no. 1, pp. 285–304, 2011.
- [13] J. Yuan, Z. He, Y. Zi, Y. Lei, and Z. Li, "Adaptive multiwavelets via two-scale similarity transforms for rotating machinery fault diagnosis," *Mechanical Systems and Signal Processing*, vol. 23, no. 5, pp. 1490–1508, 2009.
- [14] N. Lu, Z. Xiao, and O. P. Malik, "Feature extraction using adaptive multiwavelets and synthetic detection index for rotor fault diagnosis of rotating machinery," *Mechanical Systems and Signal Processing*, vol. 52-53, pp. 393–415, 2015.
- [15] L. Lu, J. Yan, C. W. de Silva, and 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.
- [16] Z. Liu, X. Zhao, M. J. Zuo, and H. Xu, "Feature selection for fault level diagnosis of planetary gearboxes," *Advances in Data Analysis and Classification*, vol. 8, no. 4, pp. 377–401, 2014.
- [17] B. Long, S. Tian, and H. Wang, "Feature vector selection method using Mahalanobis distance for diagnostics of analog circuits based on LS-SVM," *Journal of Electronic Testing*, vol. 28, no. 5, pp. 745–755, 2012.
- [18] 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.
- [19] K. Li, P. Chen, and H. Wang, "Intelligent diagnosis method for rotating machinery using wavelet transform and ant colony optimization," *IEEE Sensors Journal*, vol. 12, no. 7, pp. 2474–2484, 2012.
- [20] H. Sun, K. Li, H. Wang, P. Chen, and Y. Cao, "Intelligent mechanical fault diagnosis based on multiwavelet adaptive threshold denoising and MPSO," *Mathematical Problems in Engineering*, vol. 2014, Article ID 142795, 15 pages, 2014.
- [21] I. T. Jolliffe, *Principle Component Analysis*, Springer, New York, NY, USA, 1986.
- [22] T. F. Cox and M. A. Cox, *Multi-Dimensional Scaling*, Chapman & Hall, London, UK, 1994.
- [23] M. Gan, C. Wang, and C. a. Zhu, "Multiple-domain manifold for feature extraction in machinery fault diagnosis," *Measurement*, vol. 75, pp. 76–91, 2015.
- [24] Z. Su, B. Tang, J. Ma, and L. Deng, "Fault diagnosis method based on incremental enhanced supervised locally linear embedding and adaptive nearest neighbor classifier," *Measurement*, vol. 48, pp. 136–148, 2014.
- [25] B. Tang, T. Song, F. Li, and L. Deng, "Fault diagnosis for a wind turbine transmission system based on manifold learning and Shannon wavelet support vector machine," *Renewable Energy*, vol. 62, pp. 1–9, 2014.
- [26] F. Li, B. Tang, and R. Yang, "Rotating machine fault diagnosis using dimension reduction with linear local tangent space alignment," *Measurement*, vol. 46, no. 8, pp. 2525–2539, 2013.
- [27] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323–2326, 2000.
- [28] J. B. Tenenbaum, V. de Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, no. 5500, pp. 2319–2323, 2000.
- [29] Z. Zhang and H. Zha, "Principal manifolds and nonlinear dimensionality reduction via tangent space alignment," *SIAM Journal on Scientific Computing*, vol. 26, no. 1, pp. 313–338, 2005.

- [30] Y. Zhang, B. Li, W. Wang, T. Sun, X. Yang, and L. Wang, "Supervised locally tangent space alignment for machine fault diagnosis," *Journal of Mechanical Science and Technology*, vol. 28, no. 8, pp. 2971–2977, 2014.
- [31] X. Chen, D. Liu, G. Xu, K. Jiang, and L. Liang, "Application of wavelet packet entropy flow manifold learning in bearing factory inspection using the ultrasonic technique," *Sensors*, vol. 15, no. 1, pp. 341–351, 2014.
- [32] Y. Zhang, G. Xu, J. Wang, and L. Liang, "An automatic patient-specific seizure onset detection method in intracranial EEG based on incremental nonlinear dimensionality-reduction," *Computers in Biology and Medicine*, vol. 40, no. 11–12, pp. 889–899, 2010.
- [33] C. Wang, M. Gan, and C. a. Zhu, "Non-negative EMD manifold for feature extraction in machinery fault diagnosis," *Measurement*, vol. 70, pp. 188–202, 2015.
- [34] J. Wang and Q. He, "Exchanged ridge demodulation of time-scale manifold for enhanced fault diagnosis of rotating machinery," *Journal of Sound and Vibration*, vol. 333, no. 11, pp. 2450–2464, 2014.
- [35] T. N. T. Goodman and S. L. Lee, "Wavelets of multiplicity  $r$ ," *Transactions of the American Mathematical Society*, vol. 342, no. 1, pp. 307–324, 1994.
- [36] V. Strela and A. T. Walden, "Orthogonal and biorthogonal multiwavelets for signal denoising and image compression," in *Proceedings SPIE 3391 Wavelet Applications*, vol. 96, pp. 96–107, Orlando, FL, USA, March 1998.
- [37] V. Strela, *Multiwavelets, and Theory and Application*, Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 1996.
- [38] K. Fritz, *Wavelets and Multiwavelets*, Chapman & Hall, Ames, IA, USA, 2003.
- [39] J. S. Geronimo, D. P. Hardin, and P. R. Massopust, "Fractal functions and wavelet expansions based on several scaling functions," *Journal of Approximation Theory*, vol. 78, no. 3, pp. 373–401, 1994.
- [40] G. Y. Chen and T. D. Bui, "Multiwavelets denoising using neighboring coefficients," *IEEE Signal Processing Letters*, vol. 10, no. 7, pp. 211–214, 2003.
- [41] A. Khatami, S. Mirghasemi, A. Khosravi, C. P. Lim, and S. Nahavandi, "A new PSO-based approach to fire flame detection using K-Medoids clustering," *Expert Systems with Applications*, vol. 68, pp. 69–80, 2017.

