Mechanical Systems and Signal Processing 52-53 (2015) 393–415

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

journal homepage: www.elsevier.com/locate/ymssp

Feature extraction using adaptive multiwavelets and synthetic detection index for rotor fault diagnosis of rotating machinery

Na Lu [a], Zhihuai Xiao [a,*], O.P. Malik [b,1]

[a] Department of Power and Mechanical Engineering, Wuhan University, Wuhan 430072, Hubei Province, China
[b] Department of Electrical and Computer Engineering, University of Calgary, Calgary, Alberta, Canada T2N 1N4

ARTICLE INFO

Article history:
Received 13 July 2013
Received in revised form
6 May 2014
Accepted 30 July 2014
Available online 23 August 2014

Keywords:
Feature extraction
Fault diagnosis
Adaptive multiwavelets
Synthetic detection index
Genetic algorithms
Non-dimensional symptom parameters

ABSTRACT

State identification to diagnose the condition of rotating machinery is often converted to a classification problem of values of non-dimensional symptom parameters (NSPs). To improve the sensitivity of the NSPs to the changes in machine condition, a novel feature extraction method based on adaptive multiwavelets and the synthetic detection index (SDI) is proposed in this paper. Based on the SDI maximization principle, optimal multiwavelets are searched by genetic algorithms (GAs) from an adaptive multiwavelets library and used for extracting fault features from vibration signals. By the optimal multiwavelets, more sensitive NSPs can be extracted. To examine the effectiveness of the optimal multiwavelets, conventional methods are used for comparison study. The obtained NSPs are fed into K-means classifier to diagnose rotor faults. The results show that the proposed method can effectively improve the sensitivity of the NSPs and achieve a higher discrimination rate for rotor fault diagnosis than the conventional methods.

1. Introduction

Rotating machinery, as one of the most common types of mechanical equipment, plays an important role in industrial applications [1,2]. However, the severe working environment often inevitably causes the machinery to break down, and even results in significant economic losses and catastrophic accidents. Therefore, accurately detecting the fault type is of great significance for avoiding serious aftereffects and ensuring the safe and stable operation of the machinery.

The machinery fault diagnosis follows a roadmap of data acquisition, feature extraction and diagnostic decision making, in which feature extraction is the foundation and key to obtain an accurate diagnostic result. Various feature extraction methods have been developed to analyze the vibration signals, such as statistical method [3], time domain averaging [4], Wigner–Ville distribution [5], principle component analysis [6], independent component analysis [7], wavelet transform [8–12], and empirical mode decomposition [13,14]. A number of feature extraction methods for rotor systems fault diagnosis have been studied by researchers. Patel et al. [15] explored full spectra and orbit plots to reveal the unique nature of

* Corresponding author. Tel.: +86 13397109184; fax: +86 027 68772273.
E-mail address: xiaozhihuai@126.com (Z. Xiao).
[1] Life Fellow IEEE.

http://dx.doi.org/10.1016/j.ymssp.2014.07.024
0888-3270/

misalignment fault. Walker et al. [16] applied Butterworth filter, short term Fourier transform and moving average filter to the localization of unbalance fault. A cepstral analysis procedure to recover the excitations to transient components from vibration signal for abnormal conditions or damage detecting in rotor systems was developed by Zheng et al. [17]. Li et al. [18] investigated independent component analysis to extract features from multi-channel observations for fault diagnosis of rotor system during speed-up and speed-down process. Local mean decomposition method was applied by Wang et al. [19] to modulated signals analysis for rub-impact faults detection. Wavelet analysis was used by Peng et al. [20] as the feature extraction method for the rub-impact fault diagnosis. Bin et al. [21] combined wavelet packets and empirical mode decomposition method to extract IMFs energy moment features for early fault diagnosis of rotor systems. Lee et al. [22] introduced Hilbert–Huang transform method to get the time and frequency information for partial rub and looseness faults distinguishing. Among these methods, based on its perfect capability of focusing on localized structures in the time–frequency domain, wavelet transform is the most effective and widely used method at present [23]. However, like other signal processing techniques, wavelet transform also suffers from some defects. In essence, the wavelet transform is a kind of inner product operation. The similarity between the signals and the wavelet basis function has a decisive effect on the transform result [24]. Hence, much attention should be paid to the problem of wavelet basis function selection. Inappropriate basis function may lead to unsatisfactory feature extraction result. In addition, it has been proven that the scalar wavelet, except Haar wavelet, fails to possess the favorable properties of orthogonality, symmetry, short support and higher order of vanishing moments simultaneously [25].

Developed from scalar wavelet, multiwavelets not only preserve the advantage of advanced flexible time–frequency resolution, but also contain the properties of orthogonality, symmetry, short support and higher order of vanishing moments, simultaneously. Furthermore, with multiple basis functions, multiwavelets can match signal features more powerfully and flexibly. Since the first construction of polynomial multiwavelets was invented by the authors in [26], multiwavelets theory has developed rapidly [27–30] and achieved encouraging application effects in the aspects of noise reduction, data compression and image processing [31–33]. In recent years, multiwavelets have attracted attention from researchers in the field of machinery signal processing as well. Khadem and Rezaee [34] applied multiwavelets to the vibration analysis of a gearing system. Wang et al. [35] presented an improved multiwavelet denoising method and utilized it to the fault diagnosis of locomotive rolling bearings. Nevertheless, in these applications, the multiwavelets basis functions selected from a fixed multiwavelets library are usually independent of the vibration signal characteristics. Employing inappropriate basis functions likely reduces the accuracy of the fault diagnosis. To overcome this problem, various adaptive multiwavelets have been developed by scholars. Yuan et al. [36] constructed the adaptive multiwavelets via two similarity transforms [37] for rotating machinery fault diagnosis and showed its effectiveness in detecting the impulse feature components hidden in vibration signals. Lately, Yuan et al. [38] put forward another kind of adaptive multiwavelets based on the basis of Hermite spline interpolation and obtained a better result in gear fault detection compared to wavelet methods as well as spectral kurtosis. By symmetric lifting scheme, Wang et al. [39] also presented a kind of adaptive multiwavelets and achieved superior performance in roller bearing and gearbox vibration signals analysis. In contrast with the conventional multiwavelets, adaptive multiwavelets can change the shape of the basis functions adaptively according to the characteristics of the signals, and thus perform well in the aspect of feature extraction. Even so, the applications of adaptive multiwavelets are currently in its initial stage and more research needs to be done.

In many cases, the state identification to diagnose the condition of rotating machinery is converted to a classification problem of values of non-dimensional symptom parameters (NSPs). Therefore, the sensitivity of NSPs to the changes of machine condition is very critical to guarantee the validity of classification results. From this point of view, adaptive multiwavelets can be used to extract more sensitive NSPs in order to classify machine conditions better. In general, researchers often utilize a feature selection method to select superior NSPs from candidate NSPs set which is generated by some feature extraction technique. However, the inherent properties of previous feature extraction methods limit their abilities to adaptively change their basis functions with different types of data to obtain an optimal candidate NSPs set. Thus, the selected features may consequently possess low sensitivity and result in an inaccurate diagnosis result. To solve this problem, a novel feature extraction method based on adaptive multiwavelets and the synthetic detection index (SDI) is proposed in this paper. Adaptive multiwavelets can provide changeable basis functions according to the characteristics of signals and the SDI can evaluate the applicability of NSPs. Therefore, based on SDI maximization principle, the optimal multiwavelets, selected by genetic algorithms (GAs) from an adaptive multiwavelet library, have stronger ability of feature extraction and the NSPs generated by the optimal multiwavelets are more sensitive to changes in machine condition. To verify the effectiveness of the proposed method, the optimal multiwavelets are used for extracting features from signals measured from experimental and practical rotating machinery, and the extracted NSPs are then fed into *K*-means classifier for fault detection. The results show that, compared with GHM multiwavelets and Db4 wavelet, the proposed method can achieve higher discrimination rate for rotor fault diagnosis.

This paper is arranged in the following sequence. After an introduction of the background in Section 1, the theory of multiwavelets is briefly reviewed in Section 2. Section 3 gives introduction to the construction of adaptive multiwavelets and the way to obtain the optimal multiwavelets. Effectiveness of the optimal multiwavelets in feature extraction by application in rotor fault diagnosis is demonstrated in Section 4. Finally, conclusions are described in Section 5.

2. Multiwavelets

2.1. Multiscaling and multiwavelet functions

As in the scalar wavelet case, the theory of multiwavelets is based on the idea of multiresolution analysis (MRA) [40]. Whereas the multiscaling functions of the multiwavelets are vector functions formed by several scaling functions $\{\phi_r,\ r\in N\}$.

$$\mathbf{\Phi}(t)=[\phi_1(t),\ \phi_2(t),\ ...,\ \phi_r(t)]^T \tag{2.1}$$

and their translations and dilations generate the subspaces

$$V_j\equiv\overline{span\{2^{-j/2}\phi_l\left(\frac{t}{2^j}-k\right):1\le l\le r,\ k\in Z,\ j\in Z\}} \tag{2.2}$$

the multiwavelet functions are vector functions formed by several wavelet functions $\{\psi_r,\ r\in N\}$.

$$\mathbf{\psi}(t)=[\psi_1(t),\ \psi_2(t),\ ...,\ \psi_r(t)]^T \tag{2.3}$$

and their translations and dilations generate the subspaces

$$W_j\equiv\overline{span\{2^{-j/2}\psi_l(\frac{t}{2^j}-k):1\le l\le r,\ k\in Z,\ j\in Z\}} \tag{2.4}$$

where W_j denote the complementary subspaces of V_j in V_{j-1}, i.e. $V_{j-1}=V_j\oplus W_j$.

Two-scale relations are very important for scalar wavelet. Similarly, multiwavelet functions and multiscaling functions satisfy two-scale matrix equations

$$\mathbf{\Phi}(t)=\sqrt{2}\sum_{k=0}^{M}\mathbf{H_k}\mathbf{\Phi}(2t-k)\quad k\in Z \tag{2.5}$$

$$\mathbf{\psi}(t)=\sqrt{2}\sum_{k=0}^{M}\mathbf{G_k}\mathbf{\Phi}(2t-k)\quad k\in Z \tag{2.6}$$

where the coefficients $\{\mathbf{H_k}\}$ and $\{\mathbf{G_k}\}$ are $r\times r$ two-scale matrix sequences.

The corresponding forms of the two-scale matrix equations in frequency domain are

$$\hat{\mathbf{\Phi}}(\omega)=\mathbf{H}(\omega/2)\hat{\mathbf{\Phi}}(\omega/2) \tag{2.7}$$

$$\hat{\mathbf{\Psi}}(\omega)=\mathbf{G}(\omega/2)\hat{\mathbf{\Phi}}(\omega/2) \tag{2.8}$$

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

$$\mathbf{H}(\omega)=\frac{1}{2}\sum_{k=0}^{M}\mathbf{H}_k e^{-i\omega k} \tag{2.9}$$

$$\mathbf{G}(\omega)=\frac{1}{2}\sum_{k=0}^{M}\mathbf{G}_k e^{-i\omega k}. \tag{2.10}$$

2.2. Multiwavelet transforms

According to the two-scale matrix (Eqs. (2.5) and (2.6)), the decomposition of multiwavelets is

$$\mathbf{v}_{j,k}=\sum_n\mathbf{H}_{n-2k}\mathbf{v}_{j-1,n}\quad \mathbf{w}_{j,k}=\sum_n\mathbf{G}_{n-2k}\mathbf{v}_{j-1,n} \tag{2.11}$$

and the reconstruction of multiwavelets is

$$\mathbf{v}_{j-1,k}=\sum_n\mathbf{H}^*_{k-2n}\mathbf{v}_{j,n}+\sum_n\mathbf{G}^*_{k-2n}\mathbf{w}_{j,n} \tag{2.12}$$

where the superscript "*" represents the complex conjugate transpose.

Two-scale matrix sequences are $r\times r$ matrices. Hence, in order to perform the multiwavelet decomposition algorithm, preprocessing is needed to mapping a sequence of univariate data to r vectors (only the situation $r=2$ is considered in this paper). Among the proposed multiwavelets preprocessing methods, oversampling scheme has been proven to be better for feature extraction than other schemes [37]. Therefore, it is taken as the multiwavelet preprocessing method in this paper.

By this scheme, the bivariate vectors are formed from the sequence of original univariate data via

$$\mathbf{v}_{0,k} = \begin{bmatrix} v_{0,k}^{(0)} \\ v_{0,k}^{(1)} \end{bmatrix} = \begin{bmatrix} f[k] \\ \alpha f[k] \end{bmatrix} \tag{2.13}$$

where α is a constant. In the GHM case, $\alpha = 1/\sqrt{2}$ [41].

2.3. GHM multiwavelets

The GHM multiwavelets [27] used in this paper possess the properties of orthogonality, short support, symmetry and have an approximation order of 2. They consist of two multiscaling functions and two multiwavelet functions shown in Fig. 1.

The two-scale matrix sequences of the GHM multiwavelets are

$$\mathbf{H}_0 = \begin{bmatrix} \frac{3}{5} & \frac{4\sqrt{2}}{5} \\ -\frac{1}{10\sqrt{2}} & -\frac{3}{10} \end{bmatrix} \quad \mathbf{H}_1 = \begin{bmatrix} \frac{3}{5} & 0 \\ \frac{9}{10\sqrt{2}} & 1 \end{bmatrix}$$

$$\mathbf{H}_2 = \begin{bmatrix} 0 & 0 \\ \frac{9}{10\sqrt{2}} & -\frac{3}{10} \end{bmatrix} \quad \mathbf{H}_3 = \begin{bmatrix} 0 & 0 \\ -\frac{1}{10\sqrt{2}} & 0 \end{bmatrix} \tag{2.14}$$

$$\mathbf{G}_0 = \frac{1}{10}\begin{bmatrix} -\frac{1}{\sqrt{2}} & -3 \\ 1 & 3\sqrt{2} \end{bmatrix} \quad \mathbf{G}_1 = \frac{1}{10}\begin{bmatrix} \frac{9}{\sqrt{2}} & -10 \\ -9 & 0 \end{bmatrix}$$

$$\mathbf{G}_2 = \frac{1}{10}\begin{bmatrix} \frac{9}{\sqrt{2}} & -3 \\ 9 & -3\sqrt{2} \end{bmatrix} \quad \mathbf{G}_3 = \frac{1}{10}\begin{bmatrix} -\frac{1}{\sqrt{2}} & 0 \\ -1 & 0 \end{bmatrix}. \tag{2.15}$$

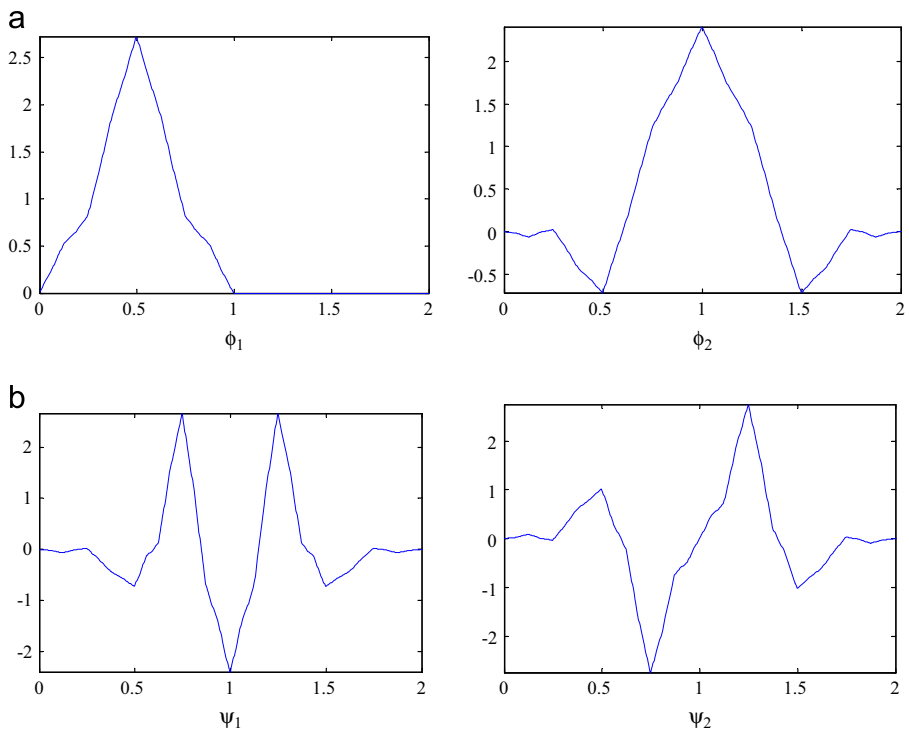

Fig. 1. GHM multiwavelets. (a) multiscaling functions, (b) multiwavelet functions.

3. Adaptive multiwavelets and optimal multiwavelets

3.1. Two-scale similarity transform

Two-scale similarity transform (TST) is a new, nonobvious construction for multiwavelets and is the theoretical basis of adaptive multiwavelets. It appeared first in [37], and then was extended to biorthogonal multiwavelets in [42]. According to the properties of the transform matrix, TST is divided into two categories: regular TST and singular TST. Compared with regular TST, singular TST can generate multiwavelets with approximation order one higher than the original multiwavelets. Therefore, it attracts more attention. The singular TST is

$$\mathbf{H}_{new}(\omega)=\tfrac{1}{2}\mathbf{M}(2\omega)\mathbf{H}(\omega)\mathbf{M}^{-1}(\omega) \tag{3.1}$$

where $\mathbf{H}(\omega)$ has approximation order $m\geq 1$, $\mathbf{H}_{new}(\omega)$ has approximation order $m+1$, and $\mathbf{M}(\omega)$ is called TST matrix which is a 2π-periodic, continuously differentiable matrix-valued function satisfying the following requirements: (1) $\mathbf{M}(\omega)$ is invertible for all $\omega\neq 2\pi k$, $k\in Z$; (2) $\mathbf{M}(0)$ has a simple eigenvalue $\lambda_{\mathbf{M}}(0)=0$ with left and right eigenvectors $\mathbf{l}$ and $\mathbf{r}$; and (3) the eigenvalue of $\mathbf{M}(\omega)$ satisfies $D(\lambda_{\mathbf{M}})(0)\neq 0$.

By singular TST, new multiwavelets with higher approximation order can be obtained, but unfortunately the orthogonality is destroyed. To make up for this drawback, the singular TST is extended to biorthogonal multiwavelets.

Theorem 3.1. [42] *Suppose that Φ, Ψ, $\tilde{\Phi}$ 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, respectively, $\tilde{\mathbf{H}}(\omega)$ has approximation order of at least one, and $\mathbf{M}(\omega)$ is a TST matrix. Then the biorthogonal multiwavelets singular TST (BI-TST) is*

$$\begin{aligned}&\mathbf{H}_{new}(\omega)=\tfrac{1}{2}\mathbf{M}(2\omega)\mathbf{H}(\omega)\mathbf{M}^{-1}(\omega)\\&\mathbf{G}_{new}(\omega)=\mathbf{G}(\omega)\mathbf{M}^{-1}(\omega)\end{aligned} \tag{3.2}$$

Correspondingly, the biorthogonal multiwavelets singular inverse TST (BI-ITST) is

$$\begin{aligned}&\tilde{\mathbf{H}}_{new}(\omega)=2\mathbf{M}^{-*}(2\omega)\tilde{\mathbf{H}}(\omega)\mathbf{M}^{*}(\omega)\\&\tilde{\mathbf{G}}_{new}(\omega)=\tilde{\mathbf{G}}(\omega)\mathbf{M}^{*}(\omega)\end{aligned} \tag{3.3}$$

By Eqs. (3.2) and (3.3), the approximation order of $\mathbf{H}_{new}(\omega)$ is increased by 1, and the approximation order of its dual counterpart $\tilde{\mathbf{H}}_{new}(\omega)$ is decreased by 1.

3.2. Adaptive multiwavelets construction

Adaptive multiwavelets are constructed by adding several adaptive parameters to the TST matrix [36]. The process consists of two steps of BI-TST and BI-ITST based on the GHM multiwavelets.

In the first step, the TST matrix added with adaptive parameters is

$$\mathbf{M}_1(\omega)=\begin{bmatrix}a(1+e^{-i\omega}) & -2\sqrt{2}a\\ b(1-e^{-i\omega}) & 0\end{bmatrix} \tag{3.4}$$

Substituting $\mathbf{M}_1(\omega)$ into Eqs. (3.2) and (3.3), the new biorthogonal multiwavelets can be obtained. Multiscaling functions, multiwavelet functions and their dual counterparts of the new biorthogonal multiwavelets on the condition $a=1$ and $b=1$ are shown in Fig. 2.

In the second step, $\tilde{\mathbf{H}}_{new}(\omega)$ and $\tilde{\mathbf{G}}_{new}(\omega)$ obtained in the first step are taken as $\mathbf{H}(\omega)$ and $\mathbf{G}(\omega)$ in Eq. (3.2), $\mathbf{H}_{new}(\omega)$ and $\mathbf{G}_{new}(\omega)$ are taken as $\tilde{\mathbf{H}}(\omega)$ and $\tilde{\mathbf{G}}(\omega)$ in Eq. (3.3), and the TST matrix is

$$\mathbf{M}_2(\omega)=\begin{bmatrix}c & 0\\ d(1+e^{-i\omega}) & f(1-e^{-i\omega})\end{bmatrix} \tag{3.5}$$

After the two steps, the final $\mathbf{H}_{new}(\omega)$ and $\tilde{\mathbf{H}}_{new}(\omega)$ with short support, symmetry properties, and approximation order of 2 are obtained. They are formed by adaptive parameters a, b, c, d and f. Therefore, the obtained GHM adaptive multiwavelets can adaptively change their basis functions with the characteristics of the data.

3.3. Optimal multiwavelets based on adaptive multiwavelets and the SDI

As is well known, certain NSPs are sensitive to certain changes of machine condition. Therefore, in order to make a comprehensive inspection on machine condition, multiple NSPs are needed to form a candidate NSPs set, from which several superior NSPs are selected for fault diagnosis. In this paper, 10 NSPs are taken into consideration as follows to form

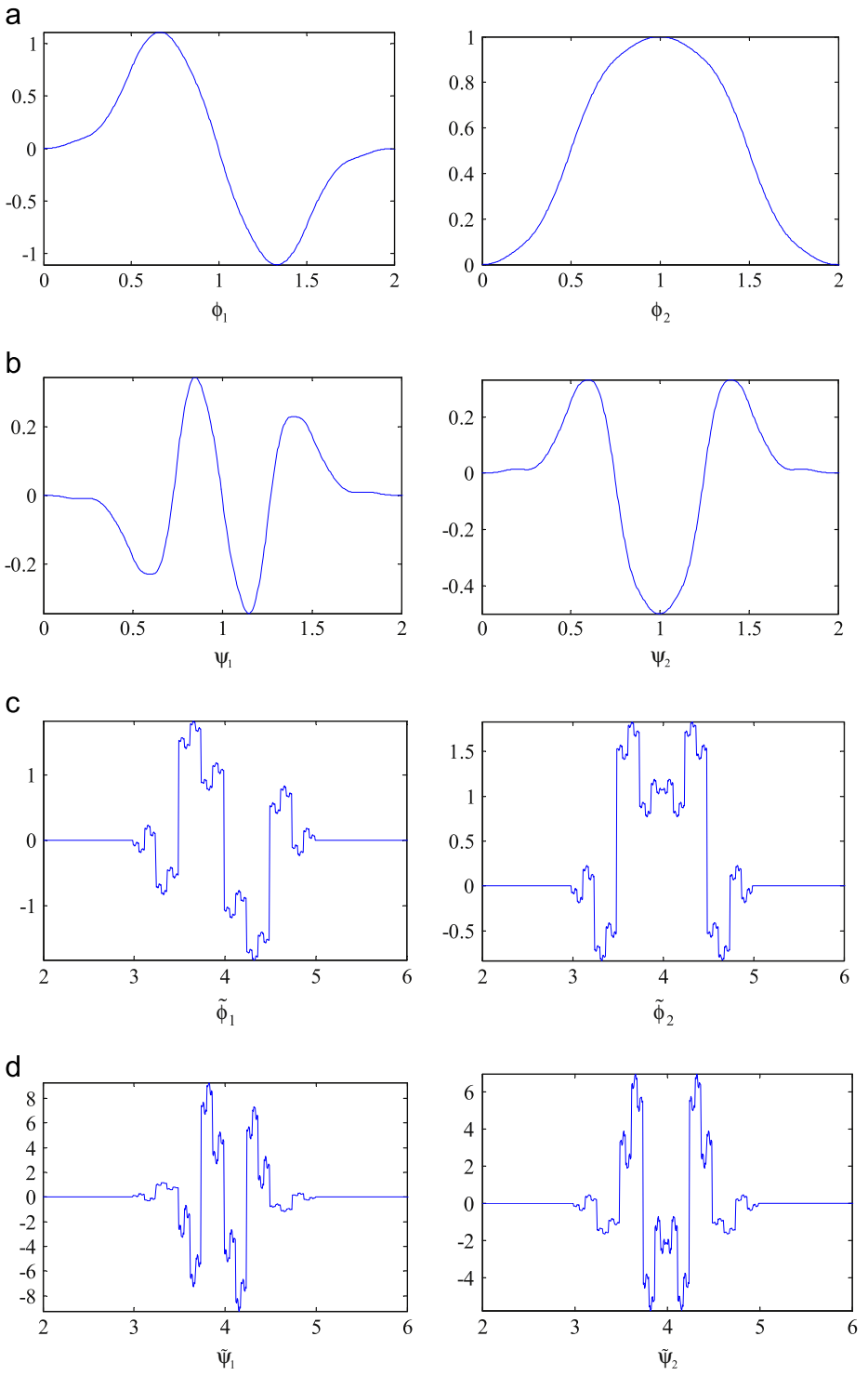

Fig. 2. Biorthogonal multiwavelets on the condition $a=1$ and $b=1$. (a) multiscaling functions, (b) multiwavelet functions, (c) dual multiscaling functions, (d) dual multiwavelet functions.

the candidate NSPs set [43]:

$$P_1 = \frac{\sigma}{\bar{\bar{f}}} \tag{3.6}$$

$$P_2 = \frac{\sum_{i=1}^{N} f_i^2}{\sigma^2} \tag{3.7}$$

$$P_3 = \sum_{i=1}^{N} \frac{(f_i - \bar{f})^3}{N\sigma^3} \tag{3.8}$$

$$P_4 = \sum_{i=1}^{N} \frac{(f_i - \bar{f})^4}{N\sigma^4} \tag{3.9}$$

$$P_5 = \frac{\bar{f}_p}{\bar{f}} \tag{3.10}$$

$$P_6 = \frac{\bar{f}_p}{\sigma} \tag{3.11}$$

$$P_7 = \frac{\left|\sum_{i=1}^{N_p} (f_{pi} - \bar{f}_p)^3\right|}{N_p \sigma_p^3} \tag{3.12}$$

$$P_8 = \frac{\left|\sum_{i=1}^{N_p} (f_{pi} - \bar{f}_p)^4\right|}{N_p \sigma_p^4} \tag{3.13}$$

$$P_9 = \frac{\left|\sum_{i=1}^{N_v} (f_{vi} - \bar{f}_v)^3\right|}{N_v \sigma_v^3} \tag{3.14}$$

$$P_{10} = \frac{\left|\sum_{i=1}^{N_v} (f_{vi} - \bar{f}_v)^4\right|}{N_v \sigma_v^4} \tag{3.15}$$

where f_i is the sampling point of the signal, f_{pi} and f_{vi} are the peak and valley values of f_i, respectively; $\bar{f}$ is the mean value of f_i, $\bar{f}_p$ and $\bar{f}_v$ are the mean values of f_{pi} and f_{vi}, respectively; σ is the standard deviation of f_i, σ_p and σ_v are the standard deviations of f_{pi} and f_{vi}, respectively.

To evaluate the sensitivity of NSPs, definitions of the detection index (DI) and the SDI are introduced as follows.

Definition 3.1. [43] Suppose s_1 and s_2 are values of an NSP calculated from the signals measured in state 1 and state 2, respectively, and conforming to the normal distributions $N(\mu_1, \sigma_1)$ and $N(\mu_2, \sigma_2)$,$\mu_2 \geq \mu_1$, then the DI corresponding to the two states is

$$\mathrm{DI} = \frac{\mu_2 - \mu_1}{\sqrt{\sigma_1^2 + \sigma_2^2}} \tag{3.16}$$

Suppose that

$$\mu = \frac{z - (\mu_2 - \mu_1)}{\sqrt{\sigma_1^2 + \sigma_2^2}}, \quad z = s_2 - s_1 \tag{3.17}$$

Then the value of the "discrimination rate (DR)" is

$$\mathrm{DR} = 1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-\mathrm{DI}} \exp\left(-\frac{\mu^2}{2}\right) d\mu \tag{3.18}$$

It can be concluded from Eqs. (3.16) and (3.18) that the more sensitive NSP is, the larger is the value of DI, and therefore, the larger the value of DR will be.

Definition 3.2. [43] Suppose the number of NSPs used for diagnosis and fault types are M and N. Then, the SDI is

$$\mathrm{SDI} = \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \sum_{k=1}^{M} \frac{|\mu_{ik} - \mu_{jk}|}{\sqrt{\sigma_{ik}^2 + \sigma_{jk}^2}} \tag{3.19}$$

where μ_{ik} is the mean value of the kth NSP related to the ith fault condition, σ_{ik} is the standard deviation of the kth NSP related to the ith fault condition.

In essence, the SDI is the sum of the DI over all of the NSPs and fault types. Therefore, better NSPs can be obtained by increasing the value of the SDI. From this perspective, the maximum SDI value is taken as the optimization object, and the commonly used generic algorithms [44] are taken as the optimization methods to search for the optimal multiwavelets from the GHM adaptive multiwavelets library with parameters a, b, c, d and f. The procedure is shown in Fig. 3. It is also summarized as Algorithm 3.1. In this algorithm, the GAs parameters are set as follows: the population scale is set to 30; the

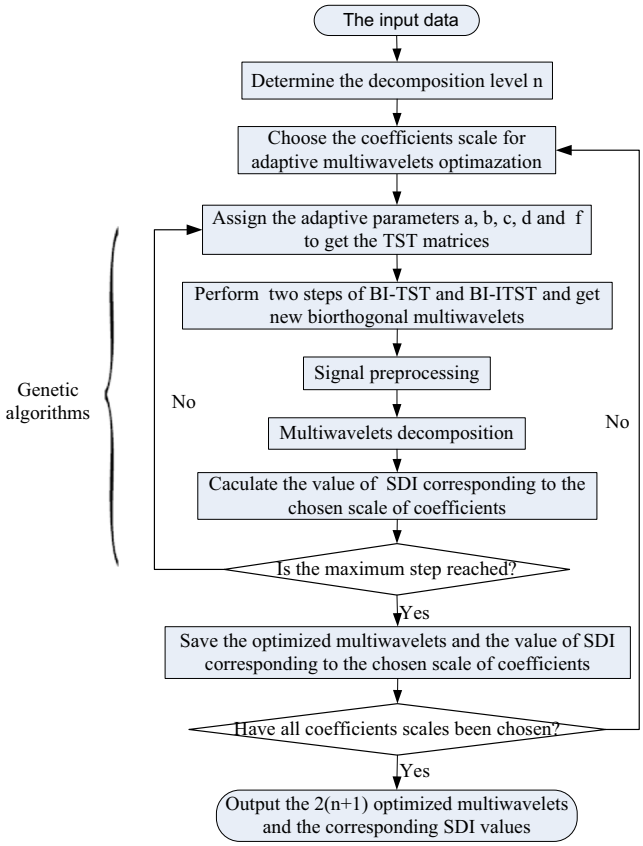

Fig. 3. Flow chart of the optimal multiwavelets obtaining procedure.

number of evolutional generations is set to 60; the probability of crossover is set to 0.7; the probability of mutation is set to 0.02; and the range of the adaptive parameters a, b, c, d, f is set to be $[-10,10]$ except zero.

Algorithm 3.1. Suppose $\mathbf{M}_1(\omega)$ and $\mathbf{M}_2(\omega)$ are TST matrices as shown in Eqs. (3.4) and (3.5), respectively. Then, the algorithm for obtaining the optimal multiwavelets for each scale of decomposed coefficients is as follows:

(1) Obtain the input data for analysis.
(2) Determine the decomposition level n.
(3) Choose the coefficients scale for adaptive multiwavelets optimization.
(4) Assign adaptive parameters: a, b, c, d, f in Eqs. (3.4) and (3.5) to get TST matrices $\mathbf{M}_1(\omega)$ and $\mathbf{M}_2(\omega)$.
(5) By employing the method introduced in Section 3.2, perform two steps of BI-TST and BI-ITST on the basis of GHM multiwavelets to obtain the two-scale matrix sequences $\mathbf{H}_{new}(\omega)$ and $\mathbf{G}_{new}(\omega)$ of the new GHM adaptive multiwavelets.
(6) Perform multiwavelets preprocessing to the input data by oversampling preprocessing scheme. Here, the parameter α is calculated by $\alpha=\left(2d/f^2c\right)+(6a/fbc)$.
(7) Plug the preprocessed signals $\mathbf{v}_{0,k}$ obtained in step (6) and the two-scale matrix sequences $\mathbf{H}_{new}(\omega)$ and $\mathbf{G}_{new}(\omega)$ obtained in step (5) into Eq. (2.11), and iterate Eq. (2.11) to get the decomposed coefficients.
(8) Get the coefficients of the scale chosen in step (3) from the decomposed coefficients, and calculate their SDI value.
(9) Compare the number of iteration with the assigned evolutional generations. Turn to step (4) if the number of iteration is smaller, otherwise save the optimized multiwavelets and the SDI value corresponding to the chosen scale of coefficients.
(10) Determine whether all the coefficients scales have been chosen. If so, finish the computation and output the $2(n+1)$ optimized multiwavelets and the corresponding SDI values, otherwise, turn to step (3).

Using Algorithm 3.1, $2(n+1)$ optimized multiwavelets are obtained. In order to select the most proper multiwavelets from these multiwavelets for feature extraction, the following procedure needs to be performed:

(1) Select the three optimized multiwavelets with the largest three SDI values from the $2(n+1)$ multiwavelets obtained by Algorithm 3.1.

(2) Use each of the selected optimized multiwavelets to decompose the input data in Algorithm 3.1 and calculate the values of the DI and NSPs by its corresponding scale of decomposed coefficients.
(3) From each of the three candidate NSPs set obtained in step (2), select two NSPs by the DI values and take them as the features extracted by each of the three optimized multiwavelets.
(4) Utilize the values of the DI to evaluate the six features obtained in step (3), and select the multiwavelets by which the two extracted features are the best as the final optimal multiwavelets.

4. Feature extraction applications for rotor fault diagnosis

To verify the effectiveness of the proposed method in characterizing fault signals, feature extraction applications to both experimental and practical rotating machinery are shown in this section.

4.1. Application to experimental rotating machinery

4.1.1. Data acquisition and denoising for experimental rotating machinery

The experimental rotating machinery system used in this paper is shown in Fig. 4. The system can simulate multiple types of rotating machinery faults. It is driven by a DC motor controlled with DH5600 speed controller. The rotor with a diameter of 10 mm and a length of 850 mm is composed of two single shafts coupled together by a coupling and supported by four bearing blocks. Two mass disks with a diameter of 75 mm are fixed on the rotor and two rub screw housings are installed on the rack of the system. The sensors for signal acquisition are comprised of two eddy current sensors for displacement measurement, a photoelectric sensor for speed measurement, and a piezoelectric accelerometer for vibration measurement. The signals measured by eddy current sensors and piezoelectric accelerometer are sent to a proximitor for filtering and amplification, and then transmitted to the computer for storage and analysis.

Vibration signals contain abundant information for rotating machinery health evaluation and are measured easily [45,46]. For these reasons, they are acquired in this paper for rotor fault detection. The considered machine conditions include normal, unbalance, misalignment, and rotor-to-stator rub condition. The unbalance condition is simulated by screwing a 2 g mass block into the threaded hole near the edge of the mass disk 1, while the misalignment condition is simulated by misaligning the coupling of the rotor. In the rotor-to-stator rub case, screw housing 1 is screwed into a rub screw. During the process of experiments, the speed of the system is assigned to 1200 rpm, and the sampling frequency is 2048 Hz. Under each of the four conditions, 40 vibration data files with 2048 points in each file are acquired, in which 20

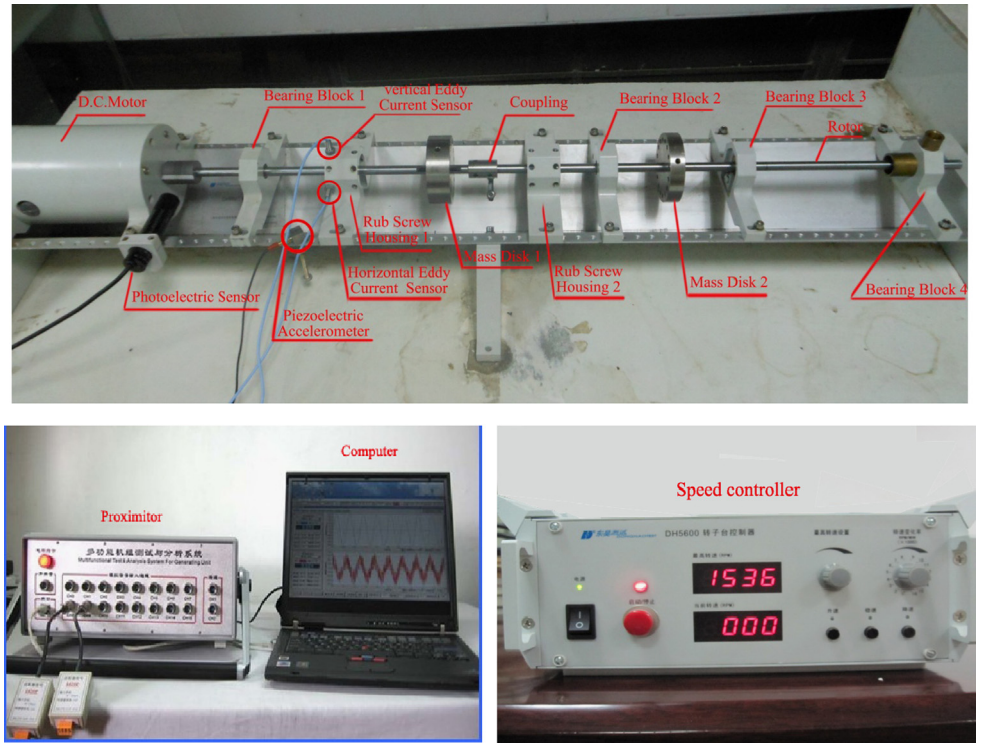

Fig. 4. Experimental rotating machinery system.

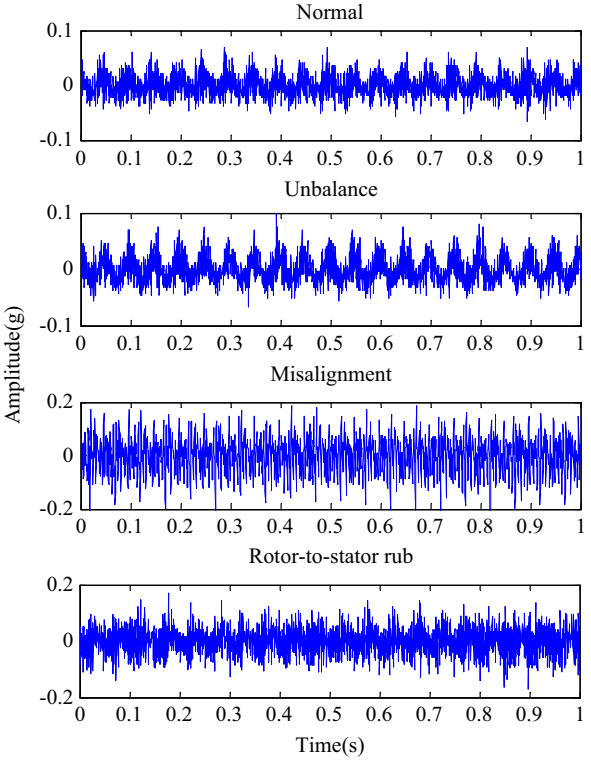

Fig. 5. Acquired signals for experimental rotating machinery.

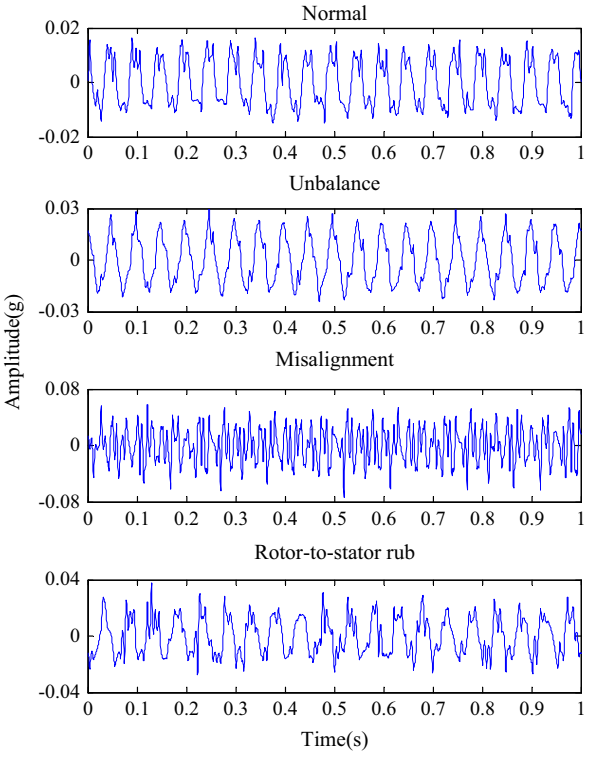

Fig. 6. Denoised signals for experimental rotating machinery.

data files are used for searching for the optimal multiwavelets and training *K*-means classifier, and the other 20 data files are used for testing. One group of data files is shown in Fig. 5.

It can be seen from Fig. 5 that the signals are submerged in substantial noise, thus their characteristics are blurred. For example, it is difficult to tell the difference between the signals acquired under the normal condition and unbalance condition. In order to remove the effects of noise on feature extraction, noise cancellation on the raw signals needs to be implemented.

Wavelet transform has become one of the most popular tools for signal denoising in recent years. However, it has been proven that the performance of multiwavelets transforms in this domain is superior to wavelet transform [32]. Particularly, the GHM multiwavelets denoising scheme using neighboring coefficients (Neighcoeff GHM) gets much better results [47]. Hence, Neighcoeff GHM denoising method is used in this paper for signal noise cancellation. Signals denoised by this method are shown in Fig. 6.

Obviously, the denoised signals in Fig. 6 can better display different dynamic responses raised by different types of machine conditions. However, for one who is not specialized in fault detection, it is still difficult to tell the type of fault from a certain denoised signal. As a result, some feature extraction strategy needs to be used to transform the denoised signals into NSPs for intelligent fault diagnosis.

4.1.2. Feature extraction for experimental rotating machinery

4.1.2.1. Feature extraction by optimal multiwavelets. Taking the denoised signals as the input data to Algorithm 3.1 and assigning the decomposition level to be 3, 8 scales of decomposed coefficients can be obtained. Denote *A*01 and *A*02 as the scales of decomposed multiscaling coefficients, and *D*01, *D*02, *D*11, *D*12, *D*21 and *D*22 as the scales of decomposed multiwavelet coefficients. Using Algorithm 3.1, 8 optimal multiwavelets corresponding to each scale of these decomposed coefficients are generated. The SDI values corresponding to the scales of decomposed coefficients obtained by each of the 8 optimal multiwavelets are shown in Fig. 7.

As seen in Fig. 7, the largest three SDI values are 175.49, 166.59 and 135.81, corresponding to the *D*11, *D*21 and *D*01 scale of decomposed coefficients. Therefore, the three optimized multiwavelets corresponding to these scales are selected to decompose the input data in Algorithm 3.1, and the values of DI and NSPs by its corresponding scale of coefficients are calculated. Use 1, 2, 3 and 4 to represent the four conditions: normal; unbalance; misalignment and rotor-to-stator rub, respectively. Then, the DI values corresponding to these scales are shown in Tables 1–3.

It can be seen from Table 1 that DI values of NSP4 and NSP6 are generally larger than others. Therefore, NSP4 and NSP6 are selected as the two NSPs corresponding to *D*11 scale of decomposed coefficients. Use the same approach to select the two NSPs corresponding to *D*21 and *D*01 scale of decomposed coefficients, and the results are NSP4 and NSP6. DI values of

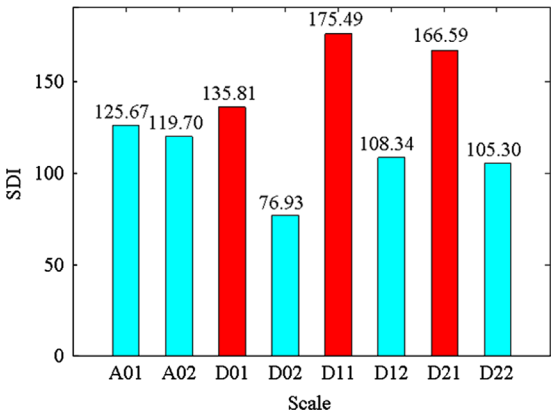

Fig. 7. SDI values corresponding to the scale of decomposed coefficients by adaptive multiwavelets for experimental rotating machinery.

Table 1
DI values corresponding to the *D*11 scale for experimental rotating machinery.

DI	NSP_1	NSP_2	NSP_3	NSP_4	NSP_5	NSP_6	NSP_7	NSP_8	NSP_9	NSP_{10}
DI_{1-2}	0.31	0.39	3.86	**11.91**	0.31	**12.50**	5.68	3.65	2.22	2.00
DI_{1-3}	0.19	0.24	0.168	**19.60**	0.32	**9.49**	2.14	4.64	5.42	1.36
DI_{1-4}	0.16	0.62	0.12	**3.22**	0.17	**0.52**	0.83	1.06	0.85	0.44
DI_{2-3}	0.54	0.48	2.90	**15.25**	0.50	**6.69**	6.25	3.07	3.32	1.87
DI_{2-4}	0.29	0.67	1.07	**1.21**	0.26	**7.56**	5.44	3.51	2.44	1.93
DI_{3-4}	0.45	0.56	0.19	**4.73**	0.46	**3.83**	0.66	2.82	1.59	0.54

Table 2
DI values corresponding to the *D*21 scale for experimental rotating machinery.

DI	NSP_1	NSP_2	NSP_3	NSP_4	NSP_5	NSP_6	NSP_7	NSP_8	NSP_9	NSP_{10}
DI_{1-2}	0.31	0.40	8.07	**17.10**	0.33	**6.19**	2.41	1.53	2.60	1.91
DI_{1-3}	0.14	0.36	1.70	**26.89**	0.23	**6.21**	2.05	1.99	2.40	3.11
DI_{1-4}	0.16	0.63	1.08	**2.97**	0.18	**0.99**	2.17	0.86	0.62	1.87
DI_{2-3}	0.53	0.53	2.16	**19.30**	0.50	**3.32**	2.82	1.30	3.01	1.79
DI_{2-4}	0.29	0.67	1.01	**0.62**	0.28	**5.76**	3.08	1.58	2.28	1.74
DI_{3-4}	0.44	0.50	0.13	**4.96**	0.46	**4.37**	1.04	2.44	1.78	0.515

Table 3
DI values corresponding to the *D*01 scale for experimental rotating machinery.

DI	NSP_1	NSP_2	NSP_3	NSP_4	NSP_5	NSP_6	NSP_7	NSP_8	NSP_9	NSP_{10}
DI_{1-2}	0.31	0.38	4.20	**9.93**	0.33	**8.58**	2.60	0.28	2.26	0.45
DI_{1-3}	0.19	0.25	0.21	**18.72**	0.21	**2.19**	4.06	1.23	3.81	1.88
DI_{1-4}	0.16	0.62	0.56	**3.53**	0.18	**1.33**	0.04	0.64	0.59	0.16
DI_{2-3}	0.54	0.48	2.58	**13.75**	0.50	**18.73**	0.78	0.86	0.12	1.11
DI_{2-4}	0.29	0.67	0.62	**1.02**	0.28	**4.96**	2.26	0.84	1.03	0.46
DI_{3-4}	0.45	0.55	0.62	**4.81**	0.45	**0.08**	3.29	1.46	1.17	1.19

the selected NSPs corresponding to *D*11, *D*21 and *D*01 scale of decomposed coefficients respectively are shown in bold in Table 1–3.

Comparing the DI values of the selected two NSPs in Tables 1–3, it can be observed that the DI_{3-4} value of the NSP6 in Table 3 is only 0.08 which means that the NSP6 has little effect on distinguishing fault types of misalignment and rotor-to-stator rub, and the lowest DI value in Table 1 is 0.52 which is slightly smaller than that in Table 2. Thus, the optimized multiwavelets corresponding to the *D*21 scale of decomposed coefficients are selected as the final optimal multiwavelets for fault feature extraction. These multiwavelets with the adaptive parameters (-8.50, -3.46, -3.33, 7.25, -1.17) are shown in Fig. 8.

Denoise the acquired training signals by Neighcoeff GHM denoising technique. Then set the decomposition level to be 3 and use the optimal multiwavelets to decompose the denoised signals into 8 scales of coefficients. Calculate the values of NSP4 and NSP6 by the *D*21 scale of decomposed coefficients. The extracted features are shown in Table 4.

4.1.2.2. Feature extraction by other methods for comparison. In this part, feature extraction by GHM multiwavelets and Db4 wavelet is performed for comparison.

4.1.2.2.1. Feature extraction by GHM multiwavelets. Use GHM multiwavelets to decompose the denoised signals into 8 scales of coefficients. The SDI values of these scales are shown in Fig. 9.

It can be seen from Figs. 7 and 9 that the SDI values are increased dramatically by adaptive multiwavelets method. Taking the *D*21 scale of decomposed coefficients as an example, the SDI value obtained by the adaptive multiwavelets method is almost 2 times larger than that by GHM multiwavelets.

Select three decomposed coefficients with the largest three SDI values from the decomposed coefficients by GHM multiwavelets. Calculate 10 NSPs values from these coefficients, and choose two better NSPs, respectively. By comparing the two NSPs corresponding to each of the three decomposed coefficients with DI values, it turns out that the NSP4 and NSP6 calculated from the *A*01 scale of coefficients are much better than the other. Therefore, the *A*01 scale of coefficients is used to extract features for fault diagnosis. The extracted features are shown in Table 5.

4.1.2.2.2. Feature extraction by Db4 wavelet. Being similar with GHM multiwavelets, Db4 wavelet is also used to extract features for comparison. The decomposition level is also set to be 3. After Db4 wavelet transformation, the signals are decomposed into 4 scales of coefficients, namely the scale of *A*0, *D*0, *D*1 and *D*2. The SDI values of these scales are shown in Fig. 10.

It can be seen in Fig. 10 that the SDI values by Db4 wavelet are significantly lower than that by adaptive multiwavelets. In the same approach with GHM multiwavelets, the two NSPs: NSP4 and NSP6 are extracted from the *A*0 scale of decomposed coefficients. The values of these features are shown in Table 6.

4.1.3. Fault diagnosis verification for experimental rotating machinery

It is difficult to tell just from the extracted features in Tables 4–6 which feature extraction method is the best. Thus, the simplest classifier, *K*-means classifier, is used to analyze these features. Cluster results by optimal multiwavelets, GHM multiwavelets and Db4 wavelet are shown in Figs. 11, 12 and 13, respectively.

It can be seen from Figs. 11 to 13 that the cluster results with fault features extracted by optimal multiwavelets can discriminate the four fault types correctly. However, as for GHM multiwavelets and Db4 wavelet method, the extracted features cannot be concentrated well, so some of the signals may be classified wrongly.

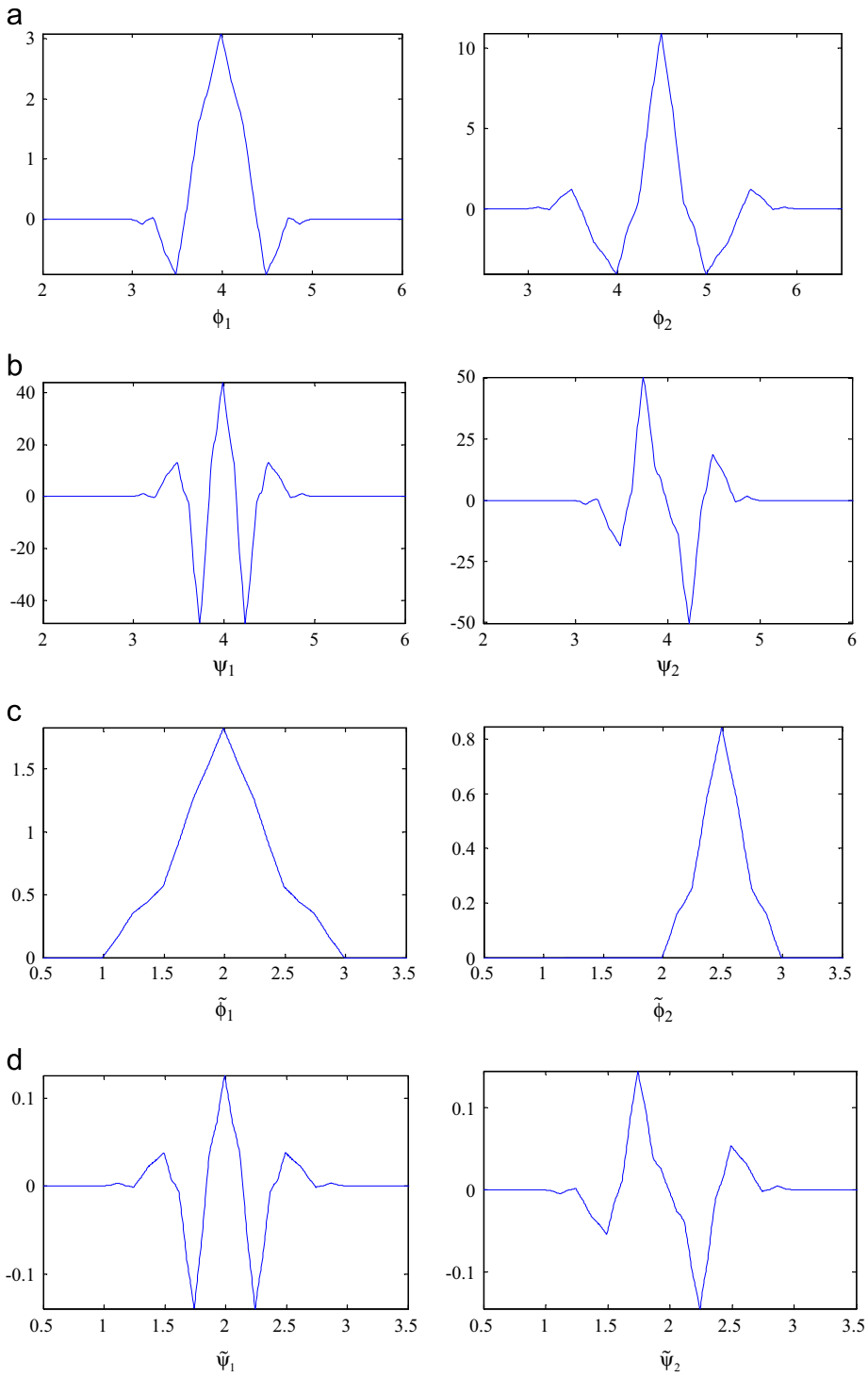

Fig. 8. Optimal multiwavelets for experimental rotating machinery. (a) multiscaling functions, (b) multiwavelet functions, (c) dual multiscaling functions, (d) dual multiwavelet functions.

In order to verify the feature extraction ability of the optimal multiwavelets more clearly, test signals are used to test the *K*-means classifiers obtained above. Taking optimal multiwavelets feature extraction method as an example (the process for GHM multiwavelets and Db4 wavelet feature extraction method is similar), the test process is as follows:

(1) The 20 test signals are denoised by the Neighcoeff GHM denoising method.

Table 4
Fault features extracted by optimal multiwavelets for experimental rotating machinery.

Signal	Normal		Unbalance		Misalignment		Rotor-to-stator rub	
	NSP4	NSP6	NSP4	NSP6	NSP4	NSP6	NSP4	NSP6
1	1.3658	0.5860	1.6466	1.1780	2.3836	0.8334	1.5425	0.4384
2	1.3402	0.5864	1.6743	1.3171	2.3072	0.8779	1.6690	0.5474
3	1.3460	0.5540	1.6194	1.3193	2.3553	0.8640	1.7665	0.4567
4	1.3458	0.5363	1.662	1.4510	2.3195	0.8968	1.7828	0.5163
5	1.3524	0.6371	1.6394	1.2891	2.3393	0.8377	1.7363	0.5241
6	1.3705	0.5424	1.6522	1.3153	2.3433	0.8524	1.5941	0.5173
7	1.3735	0.5930	1.6539	1.1120	2.3359	0.8647	1.6811	0.4047
8	1.3527	0.5552	1.6580	1.2741	2.4044	0.8501	1.7728	0.5527
9	1.3489	0.6187	1.6489	1.1422	2.3677	0.8617	1.7630	0.4237
10	1.3529	0.5849	1.6544	1.3537	2.4045	0.8844	1.7784	0.4244
11	1.3683	0.6178	1.6538	1.1738	2.4484	0.8832	1.8462	0.3486
12	1.3529	0.5379	1.6472	1.1647	2.3850	0.9167	1.8675	0.4009
13	1.3288	0.5223	1.6331	1.3574	2.3835	0.9811	2.0225	0.4473
14	1.3692	0.5777	1.6464	1.3688	2.4266	0.8721	1.8894	0.4248
15	1.3729	0.5779	1.6513	1.2380	2.377	0.8624	1.5843	0.5501
16	1.3545	0.5532	1.6649	1.1824	2.3374	0.9639	1.6837	0.5794
17	1.3468	0.5689	1.6427	1.4768	2.3886	0.8744	1.6005	0.5201
18	1.3542	0.5819	1.6441	1.2848	2.3979	0.9104	1.5909	0.5931
19	1.3521	0.5993	1.6510	1.3456	2.3633	0.9329	1.5853	0.6839
20	1.3315	0.5787	1.6473	1.0927	2.3794	0.9319	1.7942	0.4401

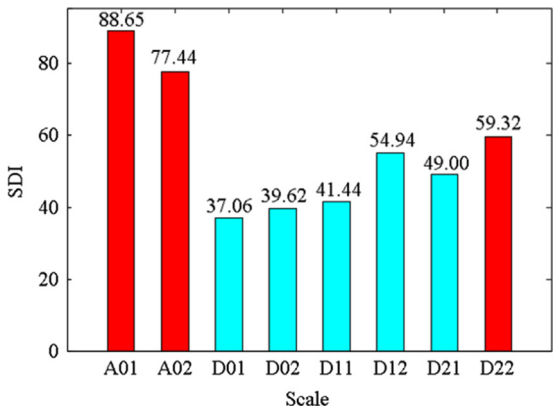

Fig. 9. SDI values of the scales of decomposed coefficients by GHM multiwavelets for experimental rotating machinery.

(2) As the feature extraction process of the training signals, certain NSPs (NSP4 and NSP6 for optimal multiwavelets feature extraction method) of the test signals are extracted by the optimal multiwavelets.
(3) For each test signal, the Euclidean distances between the extracted NSPs and the cluster centers of the cluster classifier obtained by the training signals are calculated.
(4) Find the nearest cluster center to the extracted NSPs. The fault type related to this test signal is discriminated as the fault type corresponding to this nearest cluster center.
(5) For each of the 20 test signals, compare the real fault type with the discriminated result to judge whether the discriminated result is correct or not. The number of correct ones in total number of test signals is the discrimination rate.

The test results are shown in Table 7. It can be seen from Table 7 that the fault diagnosis results using optimal multiwavelets as the feature extraction method are superior to those using GHM multiwavelets and Db4 wavelet. Specifically, the fault discrimination rate by the optimal multiwavelets is 100% under normal, unbalance, misalignment and rotor-to stator rub machinery conditions. By contrast, in the case of GHM multiwavelets, some of the signals acquired under misalignment and rotor-to-stator rub conditions cannot be discriminated correctly, and in the case of Db4 wavelet, the discrimination rates under unbalance and rotor-to-stator rub conditions are only 90% and 75%, respectively. Therefore, comparing with GHM multiwavelets and DB4 wavelet, optimal multiwavelets show stronger ability in feature extraction for rotor fault diagnosis of rotating machinery.

Table 5
Fault features extracted by GHM multiwavelets for experimental rotating machinery.

Signal	Normal		Unbalance		Misalignment		Rotor-to-stator rub	
	NSP4	NSP6	NSP4	NSP6	NSP4	NSP6	NSP4	NSP6
1	1.3997	0.9566	1.6921	1.4875	2.0271	1.0460	1.5837	0.9378
2	1.4181	1.0966	1.7531	1.5093	1.9068	1.0469	1.8316	0.9068
3	1.4233	0.8004	1.6428	1.4726	1.8698	1.0439	1.9424	1.0803
4	1.4057	0.8017	1.6939	1.4706	1.9191	1.0315	1.8873	0.9469
5	1.4200	0.9218	1.6798	1.4724	1.8922	1.0297	1.9855	0.8283
6	1.4690	1.0532	1.6522	1.4885	1.9522	1.0464	1.6980	1.0437
7	1.4256	0.9487	1.7163	1.5076	1.9291	1.0314	1.9566	1.0040
8	1.4387	1.0532	1.6795	1.4945	2.1515	1.0142	1.8515	0.9991
9	1.4489	0.9433	1.6694	1.5028	2.0315	1.0287	1.9589	1.0062
10	1.4280	0.8504	1.7009	1.4822	2.1196	1.0383	2.0435	1.1553
11	1.4692	0.8884	1.6971	1.5095	2.4632	1.0256	2.1959	1.2093
12	1.4518	1.0275	1.6656	1.4735	2.2260	1.0413	2.1651	1.1385
13	1.4029	0.7306	1.6683	1.4770	2.1084	1.0440	2.2409	1.0946
14	1.4445	0.9983	1.6746	1.4580	2.0807	1.0267	2.0457	1.1893
15	1.4958	0.8846	1.7023	1.4632	2.1235	1.0546	1.6592	1.0740
16	1.4323	0.9807	1.6743	1.490	2.2315	1.0289	1.6773	1.1243
17	1.4327	0.9526	1.6702	1.4831	2.2339	1.0111	1.6325	0.8774
18	1.4801	0.8136	1.6533	1.4671	2.2441	1.0637	1.6038	1.0090
19	1.4360	0.9612	1.6917	1.4914	2.1507	1.0131	1.6590	1.1039
20	1.4152	0.8931	1.6858	1.5026	2.1852	1.0375	1.9202	1.1788

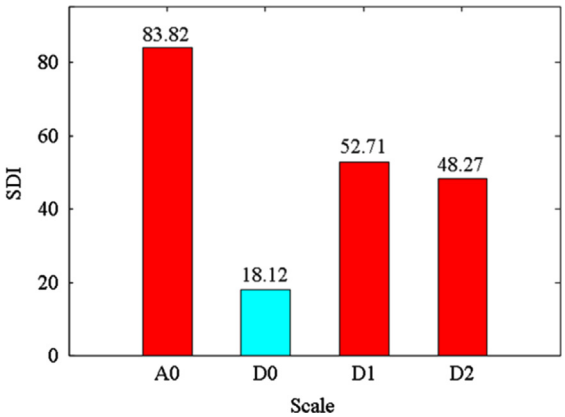

Fig. 10. SDI values of the scales of decomposed coefficients by Db4 wavelet for experimental rotating machinery.

4.2. Application to practical rotating machinery

4.2.1. Data acquisition and denoising for practical rotating machinery

Electric power generator is important equipment for supplying energy. Therefore, fault diagnosis of the power generator is of great significance. The structure of a power generator in a hydropower plant in China is sketched in Fig. 14. It is composed of a generator, a hydro-turbine, two brackets, a thrust bearing, and 3 guide bearings. The shafts of the generator and the hydro-turbine are coupled together by a coupling.

It was noticed on a certain day that the power generator 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 power generator, 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. During the process of operation, the speed of the power generator was 500 rpm, and the sampling frequency was 500 Hz. Under each of the two conditions, 32 vibration data files with 512 points in each file are acquired, in which 16 data files are used for searching for the optimal multiwavelets and training *K*-means classifier, and the other 16 data files are used for testing. One group of data files is shown in Fig. 15a. In order to remove the effects of noise on feature extraction, noise cancellation is also performed on the raw signals. The denoised signals are shown in Fig. 15b.

Table 6
Fault features extracted by Db4 wavelet for experimental rotating machinery.

Signal	Normal		Unbalance		Misalignment		Rotor-to-stator rub	
	NSP4	NSP6	NSP4	NSP6	NSP4	NSP6	NSP4	NSP6
1	1.5296	0.7735	1.6875	0.9783	2.2893	0.9981	1.7518	0.5506
2	1.5102	0.7439	1.7231	0.9643	2.1687	1.0204	2.2685	0.6792
3	1.4752	0.6820	1.6327	0.9581	2.1617	0.9757	1.9735	0.5255
4	1.4902	0.6974	1.7440	0.9290	2.1209	0.9700	2.0730	0.6499
5	1.4894	0.7521	1.7496	0.9448	2.1449	1.0298	2.0274	0.6327
6	1.5328	0.6804	1.7526	0.6977	2.2786	1.0071	1.8860	0.5642
7	1.5333	0.7686	1.7576	0.9484	2.1484	1.0327	1.8867	0.5832
8	1.5137	0.6910	1.7822	1.0550	2.1661	1.0288	2.0193	0.6440
9	1.5005	0.6025	1.7561	0.8755	2.1101	0.9846	2.2537	0.6051
10	1.5258	0.6613	1.7735	1.1528	2.3048	0.9760	2.3639	0.6319
11	1.5495	0.6881	1.7611	0.8587	2.5947	0.9441	2.4794	0.6392
12	1.5194	0.6590	1.7455	0.7992	2.3129	1.0091	2.4071	0.6171
13	1.4860	0.6502	1.7205	1.0100	2.3585	0.9695	2.4776	0.7244
14	1.5556	0.7315	1.7623	1.1441	2.2428	0.9513	2.3548	0.7564
15	1.6417	0.7483	1.7767	0.8558	2.2946	1.0278	1.7468	0.6064
16	1.5110	0.6512	1.7591	0.8967	2.3377	1.0135	1.9801	0.5747
17	1.5254	0.7917	1.7440	0.9719	2.2666	0.9974	1.7915	0.5303
18	1.5465	0.7096	1.7402	0.8692	2.2146	0.9688	1.8100	0.5558
19	1.4937	0.6611	1.7414	0.9212	2.2565	1.0050	1.7829	0.6331
20	1.4817	0.7301	1.7300	0.7380	2.2273	0.9744	2.1804	0.6724

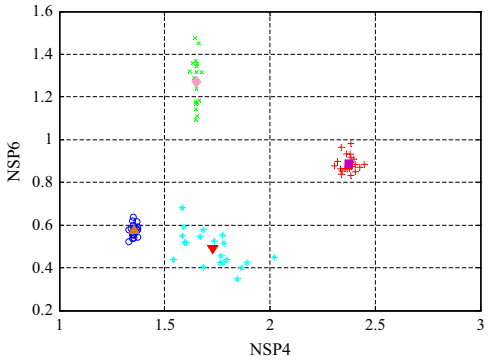

Fig. 11. Cluster results with fault features extracted by optimal multiwavelets for experimental rotating machinery.

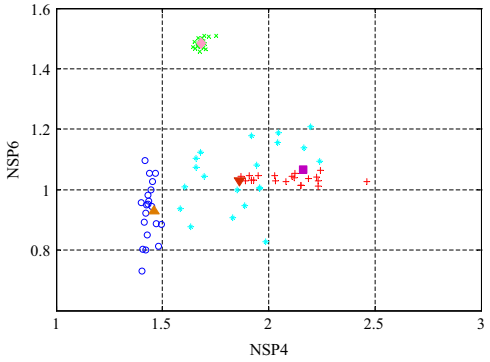

Fig. 12. Cluster results with fault features extracted by GHM multiwavelets for experimental rotating machinery.

4.2.2. Feature extraction for practical rotating machinery

The process of feature extraction for practical rotating machinery is the same as that for the experimental rotating machinery. Details are as follows.

4.2.2.1. Feature extraction by optimal multiwavelets. Taking the denoised signals as the input data to Algorithm 3.1 and assigning the decomposition level to be 3, 8 optimal multiwavelets are generated. The SDI values obtained by each of the 8

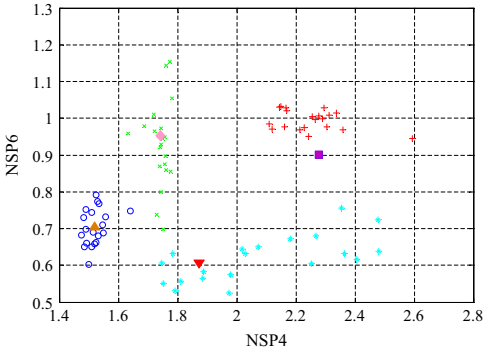

Fig. 13. Cluster results with fault features extracted by Db4 wavelet for experimental rotating machinery.

Table 7
Fault discrimination rate for experimental rotating machinery.

Method	Normal (%)	Unbalance (%)	Misalignment (%)	Rotor-to-stator rub (%)	Total (%)
Optimal multiwavelets	100	100	100	100	100
GHM multiwavelets	100	100	90	80	92.5
Db4 wavelet	100	90	100	75	91.25

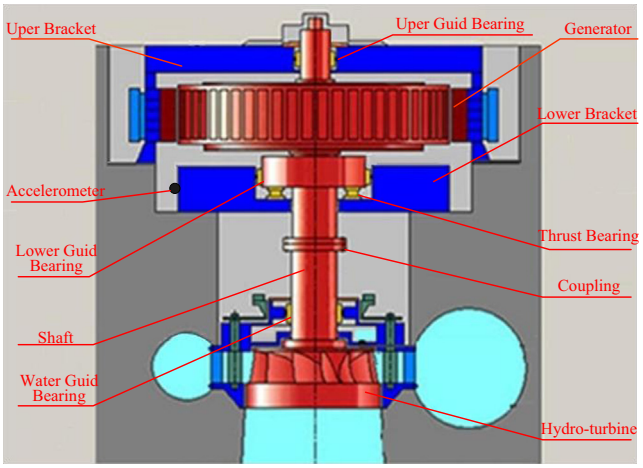

Fig. 14. Sketch of the power generator structure.

optimal multiwavelets are shown in Fig. 16. It can be seen from Fig. 16 that the largest three SDI values are 16.66, 10.72 and 11.45, corresponding to the $D12$, $D21$ and $D22$ scale of decomposed coefficients. Therefore, the three optimized multiwavelets corresponding to these scales are selected to decompose the input data in Algorithm 3.1 and the values of DI and NSPs by its corresponding scale of coefficients are calculated. The DI values corresponding to these scales are shown in Table 8. According to Table 8, NSP4 and NSP6 are selected as the two NSPs corresponding to $D12$ and $D22$ scale of decomposed coefficients, and NSP4 and NSP9 are selected as the two NSPs corresponding to $D21$ scale of decomposed coefficients. The DI values of the selected NSPs corresponding to, $D12$, $D21$ and $D22$, scale of decomposed coefficients respectively are shown in bold in Table 8. It can be seen from the selected NSPs that DI values of the selected two NSPs corresponding to $D12$ scale of decomposed coefficients are generally larger than those corresponding to $D21$ and $D22$ scale of decomposed coefficients. Therefore, the optimized multiwavelets corresponding to the $D12$ scale of decomposed coefficients are selected as the final optimal multiwavelets for fault feature extraction. These multiwavelets with the adaptive parameters $(-6.42,\ -6.29,\ -7.18,\ 8.67,\ 2.67)$ are shown in Fig. 17. The extracted features by the optimal multiwavelets are shown in Table 9.

4.2.2.2. Feature extraction by other methods for comparison

4.2.2.2.1. Feature extraction by GHM multiwavelets. Use GHM multiwavelets to decompose the denoised signals into 8 scales of coefficients. The SDI values of these scales are shown in Fig. 18. Using the same approach as that used for

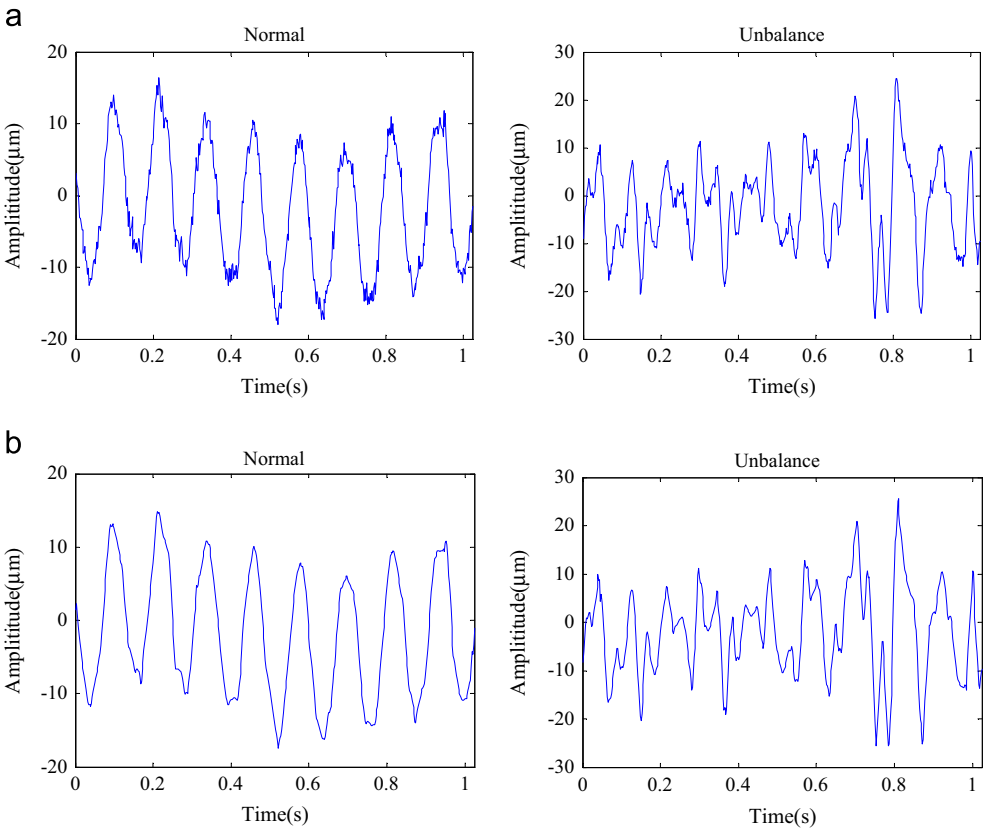

Fig. 15. Signals of the power generator. (a) acquired signals, (b) denoised signals.

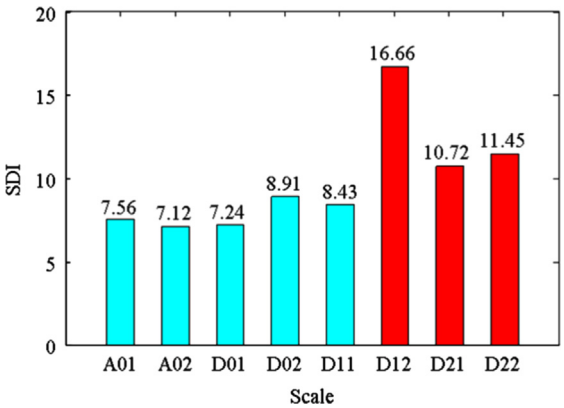

Fig. 16. SDI values corresponding to the scale of decomposed coefficients by adaptive multiwavelets for practical rotating machinery.

Table 8
DI values corresponding to the *D*12, *D*21 and *D*22 scale for practical rotating machinery.

Scale	NSP_1	NSP_2	NSP_3	NSP_4	NSP_5	NSP_6	NSP_7	NSP_8	NSP_9	NSP_{10}
D12	0.31	0.15	1.15	**3.12**	0.30	**9.26**	0.80	0.59	0.59	0.38
D21	0.15	0.37	1.01	**2.80**	0.03	1.49	1.04	0.53	**2.30**	0.99
D22	0.26	0.81	0.43	**1.51**	0.26	**4.96**	0.84	0.87	0.45	1.06

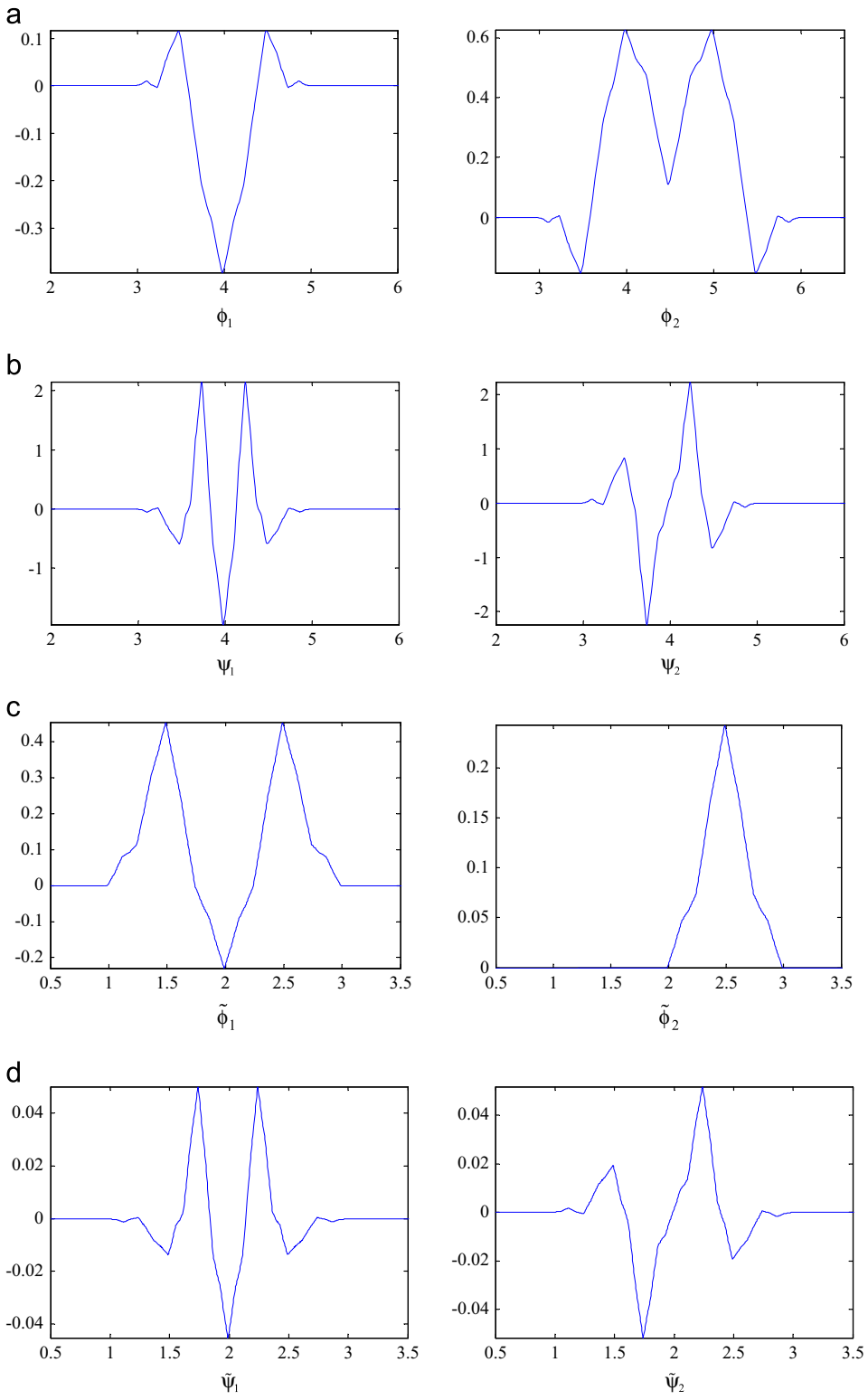

Fig. 17. Optimal multiwavelets for practical rotating machinery. (a) multiscaling functions, (b) multiwavelet functions, (c) dual multiscaling functions, (d) dual multiwavelet functions.

Table 9
Fault features extracted by optimal multiwavelets for practical rotating machinery.

Signal	Normal		Unbalance		Signal	Normal		Unbalance	
	NSP4	NSP6	NSP4	NSP6		NSP4	NSP6	NSP4	NSP6
1	1.6975	1.3437	2.6202	0.8265	9	1.6081	1.3837	3.0258	0.8337
2	1.6476	1.3701	2.7775	0.8442	10	1.5553	1.3547	3.2432	0.8935
3	1.5975	1.2902	2.7546	0.9298	11	1.5515	1.3409	3.6041	0.9067
4	1.653	1.3947	3.4196	0.7773	12	1.6868	1.3234	3.2574	0.8839
5	1.6372	1.3478	3.5575	0.8452	13	1.726	1.4174	4.373	0.8163
6	1.6443	1.306	3.4784	0.8875	14	1.6522	1.3616	3.5665	0.8176
7	1.6127	1.3688	2.6464	0.775	15	1.5324	1.3277	2.6896	0.8298
8	1.6608	1.3819	2.607	0.8463	16	1.6344	1.3739	3.0544	0.8595

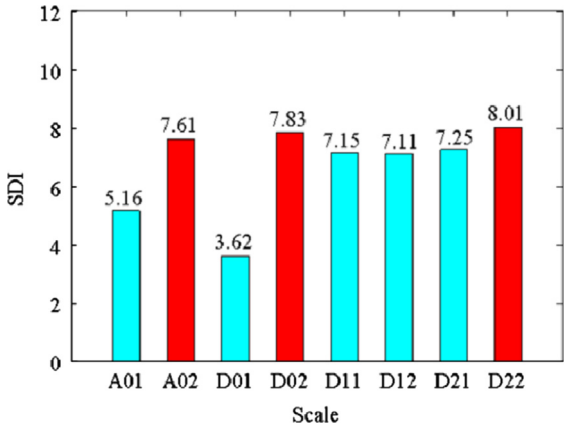

Fig. 18. SDI values of the scales of decomposed coefficients by GHM multiwavelets for practical rotating machinery.

Table 10
Fault features extracted by GHM multiwavelets for practical rotating machinery.

Signal	Normal		Unbalance		Signal	Normal		Unbalance	
	NSP3	NSP4	NSP3	NSP4		NSP3	NSP4	NSP3	NSP4
1	3.4194	17.1450	0.0883	2.9708	9	3.1620	15.5760	0.1917	3.1015
2	3.1027	15.5360	0.0601	2.7272	10	3.5197	17.6860	0.7208	3.6424
3	2.6800	14.2170	0.1564	2.0378	11	2.5327	11.7690	0.1830	2.3595
4	1.9900	10.3070	0.0129	2.9275	12	1.2529	7.0592	0.1330	2.3299
5	0.6164	3.3649	0.0841	2.9847	13	0.0948	4.0417	0.5478	2.6063
6	0.8066	3.4434	0.0268	2.2190	14	0.3976	2.8067	0.5652	2.7442
7	2.1612	11.4490	0.1984	2.0222	15	0.0366	1.9178	0.0212	2.3133
8	1.9081	8.2750	0.3360	2.1280	16	2.4280	11.7900	0.0833	2.4898

experimental rotating machinery, the NSP3 and NSP4 of *D*02 scale of coefficients are extracted for fault diagnosis. Their values are shown in Table 10.

4.2.2.2.2. Feature extraction by Db4 wavelet. In the Db4 wavelet case, the signals are decomposed into 4 scales of coefficients. The SDI values of the scale of *A*0, *D*0, *D*1 and *D*2 decomposed coefficients are shown in Fig. 19. The NSP4 and NSP6 of *A*0 scale of coefficients are extracted for fault diagnosis. Values of these features are shown in Table 11.

4.2.3. Fault diagnosis verification for practical rotating machinery

In order to compare the efficiency of the feature extraction methods, the extracted features in Tables 9–11 are fed into *K*-means classifier. The cluster results by optimal multiwavelets, GHM multiwavelets and Db4 wavelet are shown in Figs. 20, 21 and 22, respectively.

It can be seen from Figs. 20 to 22 that the fault features extracted by optimal multiwavelets can be concentrated better than those extracted by GHM multiwavelets and Db4 wavelet method.

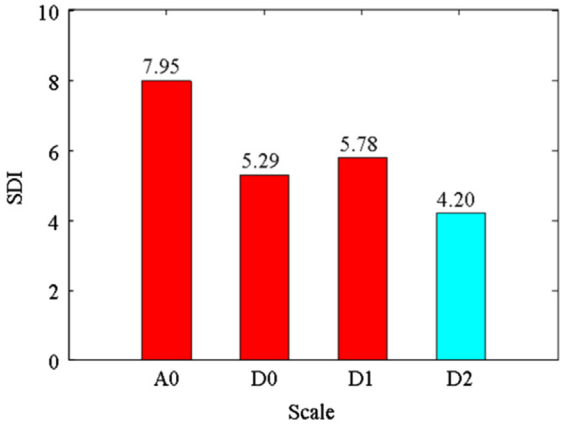

Fig. 19. SDI values of the scales of decomposed coefficients by Db4 wavelet for practical rotating machinery.

Table 11
Fault features extracted by Db4 wavelet for practical rotating machinery.

Signal	Normal		Unbalance		Signal	Normal		Unbalance	
	NSP4	NSP6	NSP4	NSP6		NSP4	NSP6	NSP4	NSP6
1	1.4580	1.2740	3.1175	0.6808	9	1.6975	1.2870	4.2830	0.9414
2	1.4759	1.3265	3.6586	0.6297	10	1.4282	1.0772	2.4666	0.6680
3	1.5800	0.8346	2.6995	0.7735	11	1.5184	1.3470	3.3144	0.7405
4	1.6621	1.3186	2.8703	0.9753	12	2.1813	0.9563	3.1100	0.7053
5	1.6156	1.1474	3.5254	0.6269	13	1.8402	1.2141	2.8434	0.7145
6	1.9065	0.9356	2.5201	0.6401	14	1.6719	1.1772	3.7590	0.7635
7	1.5730	1.2627	4.0956	0.7501	15	1.9455	1.1208	2.3673	0.8164
8	1.5466	0.9815	4.1269	0.8053	16	1.5354	1.2666	3.4828	0.8887

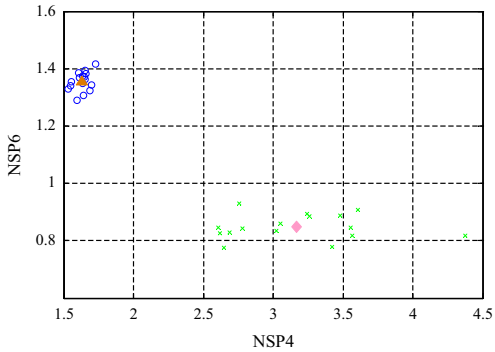

Fig. 20. Cluster results with fault features extracted by optimal multiwavelets for practical rotating machinery.

In order to verify the feature extraction ability of the optimal multiwavelets more clearly, the test signals are denoised by Neighcoeff GHM denoising method, and then used to test the obtained *K*-means classifiers. The test results are shown in Table 12.

It can be seen from Table 12 that fault diagnosis results using optimal multiwavelets as the feature extraction method is superior to those using GHM multiwavelets and Db4 wavelet. Specifically, by the optimal multiwavelets, the discrimination rates are 100% under normal and unbalance conditions, but the discrimination rate under normal condition is only 68.75% when using GHM multiwavelets as the feature extraction method, and the discrimination rate under unbalance condition is only 62.5% when using Db4 wavelet as the feature extraction method.

5. Conclusions

In recent years, adaptive multiwavelets have been applied for machinery fault diagnosis and achieved good results. However, they have not been used for improving the sensitivity of signals to the changes in machine conditions. Investigations in this area are conducted in this paper by combining adaptive multiwavelets with SDI, and the proposed

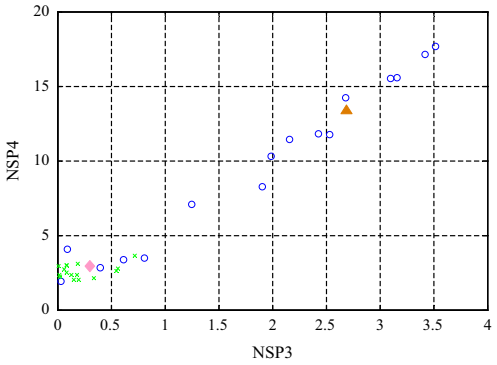

Fig. 21. Cluster results with fault features extracted by GHM multiwavelets for practical rotating machinery.

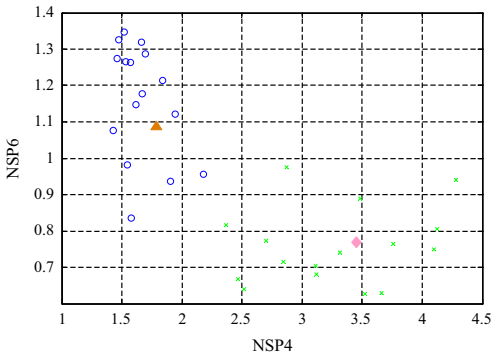

Fig. 22. Cluster results with fault features extracted by Db4 wavelet for practical rotating machinery.

Table 12
Fault discrimination rate for practical rotating machinery.

Method	Normal (%)	Unbalance (%)	Total (%)
Optimal multiwavelets	100	100	100
GHM multiwavelets	68.75	100	84.375
Db4 wavelet	100	62.5	81.25

method is applied to feature extraction of signals measured from rotating machinery. The results obtained by feeding the extracted features into the *K*-means classifier show that the proposed method is better than the conventional GHM multiwavelets and Db4 wavelet methods in respect of fault feature extraction.

Acknowledgments

The authors acknowledge the support from the National Natural Science Foundation of China under Grant no. 51179135 and the Fundamental Research Funds for the Central Universities under the Contract no. 201120802020004.

References

[1] A. Muszynska, Rotordynamics, CRC Press – Taylor & Francis Group, Boca Raton, 2005.
[2] M.L. Adams, Rotating Machinery Vibration from Analysis to Troubleshooting, Marcel Dekker, New York, 2000.
[3] S.J. Loutridis, Gear failure prediction using multiscale local statistics, Eng. Struct. 30 (2008) 1214–1223.
[4] Enayet B. Halim, M.A.A. Shoukat Choudhury, Sirish L. Shah, Ming J. Zuo, Time domain averaging across all scales: a novel method for detection of gearbox faults, Mech. Syst. Signal Process. 22 (2008) 261–278.
[5] F.K. Choy, D.H. Mugler, J. Zhou, Damage identification of a gear transmission using vibration signatures, Trans. ASME 125 (2003) 394–403.
[6] Q. He, R. Yan, F. Kong, R. Du, Machine condition monitoring using principal component representations, Mech. Syst. Signal Process. 23 (2009) 446–466.
[7] Q. He, R. Du, F. Kong, Phase space feature based on independent component analysis for machine health diagnosis, J. Vib. Acoust. 134 (2012) 1–11.

[8] J. Wu, J. Kuo, An automotive generator fault diagnosis system using discrete wavelet transform and artificial neural network, Expert Syst. Appl. 36 (2009) 9776–9783.
[9] J. Rafiee, M.A. Rafiee, P.W. Tse, Application of mother wavelet functions for automatic gear and bearing fault diagnosis, Expert Syst. Appl. 37 (2010) 4568–4579.
[10] Amir Hosein Zamanian, Abdolreza Ohadi, Gear fault diagnosis based on Gaussian correlation of vibrations signals and wavelet coefficients, Appl. Soft Comput. 11 (2011) 4807–4819.
[11] P.K. Kankar, Satish C. Sharma, S.P. Harsha, Rolling element bearing fault diagnosis using wavelet transform, Neurocomputing 74 (2011) 1638–1645.
[12] S. Wang, W. Huang, Z.K. Zhu, Transient modeling and parameter identification based on wavelet and correlation filtering for rotating machine fault diagnosis, Mech. Syst. Signal Process. 25 (2011) 1299–1320.
[13] X. Zhao, Tejas H. Patel, Ming J. Zuo, Multivariate EMD and full spectrum based condition monitoring for rotating machinery, Mech. Syst. Signal Process. 27 (2012) 712–728.
[14] Y. Lei, J. Lin, Z. He, Ming J. Zuo, A review on empirical mode decomposition in fault diagnosis of rotating machinery, Mech. Syst. Signal Process. 35 (2013) 108–126.
[15] Tejas H. Patel, Ashish K. Darpe, Experimental investigations on vibration response of misaligned rotors, Mech. Syst. Signal Process. 23 (2009) 2236–2252.
[16] R.B. Walker, R. Vayanat, S. Perinpanayagam, I.K. Jennions, Unbalance location through machine nonlinearities using an artificial neural network approach, Mech. Mach. Theory 75 (2014) 54–66.
[17] G.T. Zheng, W.J. Wang, A new cepstral analysis procedure of recovering excitations for transient components of vibration signals and applications to rotating machinery condition monitoring, Trans. ASME 123 (2001) 222–229.
[18] Zhinong Li, Yongyong He, Fulei Chu, Jie Han, Wei Hao, Fault recognition method for speed-up and speed-down process of rotating machinery based on independent component analysis and Factorial Hidden Markov Model, J. Sound Vib. 291 (2006) 60–71.
[19] Yanxue Wang, Zhengjia He, Yanyang Zi, A demodulation method based on improved local mean decomposition and its application in rub-impact fault diagnosis, Meas. Sci. Technol. 20 (2009) 1–10.
[20] Z. Peng, Y. He, Q. Lu, F. Chu, Feature extraction of the rub-impact rotor system by means of wavelet analysis, J. Sound Vib. 259 (2003) 1000–1010.
[21] G.F. Bin, J.J. Gao, X.J. Li, B.S. Dhillon, Early fault diagnosis of rotating machinery based on wavelet packets-empirical mode decomposition feature extraction and neural network, Mech. Syst. Signal Process. 27 (2012) 696–711.
[22] Seung-Mock Lee, Yeon-Sun Choi, Fault diagnosis of partial rub and looseness in rotating machinery using Hilbert–Huang transform, J. Mech. Sci. Technol. 22 (2008) 2151–2162.
[23] Ruqiang Yan, Robert X. Gao, Xuefeng Chen, Wavelets for fault diagnosis of rotary machines: A review with applications, Signal Process. 96 (2014) 1–15.
[24] B. Chen, Z. Zhang, C. Sun, B. Li, Y. Zi, Z. He, Fault feature extraction of gearbox by using over complete rational dilation discrete wavelet transform on signals measured from vibration sensors, Mech. Syst. Signal Process. 33 (2012) 275–298.
[25] I. Daubechies, Ten Lectures on Wavelets, SIAM, Philadelphia, 1992.
[26] Bradley K. Alpert, A class of bases in L^2 for the sparse representation of integral operators, SIAM J. Math. Anal. 24 (1993) 246–262.
[27] Jeffery S. Geronimo, Douglas P. Hardin, Peter R. Massopust, Fractal function and wavelet expansions based on several scaling functions, J. Approx. Theory 78 (1994) 373–401.
[28] Jeffery S. Geronimo, Douglas P. Hardin, Peter R. Massopust, Construction of orthogonal wavelets using fractal interpolation functions, SIAM J. Math. Anal. 27 (1996) 1158–1192.
[29] Charles K. Chui, J. Lian, A study of orthonormal multi-wavelets, Appl. Numer. Math. 20 (1996) 273–298.
[30] Douglas P. Hardin, Jeffrey A. Marasovich, Biorthogonal multiwavelets on [−1,1], Appl. Comput. Harmon. Anal. 7 (1999) 34–53.
[31] Tien D. Bui, G. Chen, Translation-invariant denoising using multiwavelets, IEEE Trans. Signal Process. 46 (1998) 3414–3420.
[32] Vasily Strela, Peter Niels Heller, Gilbert Strang, Pankaj Topiwala, Christopher Heil, The application of multiwavelet filter banks to image processing, IEEE Trans. Image Process. 8 (1999) 548–563.
[33] Vasily Strela, Andrew T. Walden, Signal and Image Denoising via Wavelet Thresholding: Orthogonal and Biorthogonal, Scalar and Multiple Wavelet Transforms, Imperial College, Statistics Section, Technical Report TR-98-01, 1998.
[34] S.E. Khadem, M. Rezaee, Development of vibration signature analysis using multiwavelet systems, J. Sound Vib. 261 (2003) 613–633.
[35] X. Wang, Y. Zi, Z. He, Multiwavelet denoising with improved neighboring coefficients for application on rolling bearing fault diagnosis, Mech. Syst. Signal Process. 25 (2011) 285–304.
[36] J. Yuan, Z. He, Y. Zi, Y. Lei, Z. Li, Adaptive multiwavelets via two-scale similarity transforms for rotating machinery fault diagnosis, Mech. Syst. Signal Process. 23 (2009) 1490–1508.
[37] Vasily Strela, Multiwavelets: Theory and Application (Ph.D. thesis), Massachusetts Institute of Technology, Cambridge, 1996.
[38] J. Yuan, Z. He, Y. Zi, Gear fault detection using customized multiwavelet lifting schemes, Mech. Syst. Signal Process. 24 (2010) 1509–1528.
[39] X. Wang, Y. Zi, Z. He, Multiwavelet construction via an adaptive symmetric lifting scheme and its applications for rotating machinery fault diagnosis, Meas. Sci. Technol. 20 (2009) 1–17.
[40] T.N.T. Goodman, S.L. Lee, Wavelets of multiplicity r, Trans. Am. Math. Soc. 342 (1994) 307–324.
[41] Vasily Strela, Andrew T. Walden, Orthogonal and biorthogonal multiwavelets for signal denoising and image compression, Proc. SPIE 3391 Wavelet Appl. 96 (1998) 96–107.
[42] Fritz Keinert, Wavelets and Multiwavelets, Chapman & Hall, Ames, 2003.
[43] K. Li, P. Chen, H. Wang, Intelligent diagnosis method for rotating machinery using wavelet transform and ant colony optimization, IEEE Sens. J. 12 (2012) 2474–2484.
[44] Z. Michalewicz, Genetic Algorithms+Data Structures=Evolution Programs, Springer, Charlotte, 1999.
[45] C. Duan, Z. He, H. Jiang, A sliding window feature extraction method for rotating machinery based on the lifting scheme, J. Sound Vib. 299 (2007) 774–785.
[46] Y. Lei, Ming J. Zuo, Gear crack level identification based on weighted K nearest neighbor classification algorithm, Mech. Syst. Signal Process. 23 (2009) 1535–1547.
[47] G.Y. Chen, T.D. Bui, Multiwavelets denoising using neighboring coefficients, IEEE Signal Process. Lett. 10 (2003) 211–214.