

Research Article

A Novel Fault Diagnosis Model for Bearing of Railway Vehicles Using Vibration Signals Based on Symmetric Alpha-Stable Distribution Feature Extraction

Yongjian Li,¹ Weihua Zhang,¹ Qing Xiong,² Tianwei Lu,³ and Guiming Mei¹

¹State Key Laboratory of Traction Power, Southwest Jiaotong University, Chengdu 610031, China

²School of Automobile & Transportation, Xihua University, Chengdu 610039, China

³School of Mechanical Engineering, Southwest Jiaotong University, Chengdu 610031, China

Correspondence should be addressed to Yongjian Li; guinong9339@163.com

Received 28 August 2016; Revised 18 November 2016; Accepted 28 November 2016

Academic Editor: Jussi Sopanen

Copyright © 2016 Yongjian Li 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.

Axle box bearings are the most critical mechanical components of railway vehicles. Condition monitoring is of great benefit to ensure the healthy status of bearings in the railway train. In this paper, a novel fault diagnosis model for axle box bearing based on symmetric alpha-stable distribution feature extraction and least squares support vector machines (LS-SVM) using vibration signals is proposed which is conducted in three main steps. Firstly, fast nonlocal means is used for denoising and ensemble empirical mode decomposition is applied to extract fault feature information. Then a new statistical method of feature extraction, symmetric alpha-stable distribution, is employed to obtain representative features from intrinsic mode functions. Additionally, the hybrid fault feature sets are input into LS-SVM to identify the fault type. To enhance the performance of LS-SVM in the case of small-scale samples, Morlet wavelet kernel function is combined with LS-SVM for the classification of fault type and fault severity and the particle swarm optimization is used for the optimization of LS-SVM parameters. Finally, the experimental results demonstrate that the proposed approach performs more effectively and robustly than the other methods in small-scale samples for fault detection and classification of railway vehicle bearings.

1. Introduction

Rolling element bearings have been widely used in industrial applications. Axle box bearings are one of the critical mechanical components of railway vehicles. The frequent failures including pitting, stripping, wear, crack, and abrasion of train bearings have a great influence on the traffic safety. Therefore, effective identification of bearing health status is indispensable to monitor the working condition of axle box bearings for train maintenance [1, 2]. Currently, vibration analysis and acoustic analysis are two main approaches for defect detection [3, 4]. Vibration-based diagnosis has become the most common monitoring technique because of its higher reliability.

In the process of fault diagnosis, extracting defect features from noisy vibration signals remains a great challenge. Many sources of signal contamination including additive noise,

the signals from shafts, gearboxes, and other mechanical components of railway vehicles overlap signals of interest in both time and frequency. Thus, it is vital to advance signal denoising method to get rid of the noises and extract the fault characteristics. For that reason, a lot of algorithms have been developed for vibration signal denoising. In recent years, the methods based on the discrete wavelet transform (DWT) [5–7] coefficient shrinkage, the empirical mode decomposition (EMD) [8–13], and the nonlocal means (NLMs) [14, 15] have been introduced as three popular methods to rotating mechanical fault diagnosis. DWT is characteristic of analyzing signals on multiple scales by discarding the lower magnitude, and the performance of the wavelet transform relies on the selection of the wavelet basis function. In the EMD method, the clean vibration signal is obtained by discarding first few intrinsic mode functions (IMFs). Mode mixing [16], resulting from signal intermittence, is

the disadvantage of EMD. To overcome this obstacle, the ensemble EMD (EEMD) was proposed by Wu and Huang [17]. The fast NLMs (FNLM) approach is a very successful image denoising method [14], which has been applied for rotating machinery fault diagnosis. In this study, the FNLM approach and the EEMD approach are combined to denoise the raw vibration signal.

After denoising step, the feature parameters which correctly represent the health status of the axle box bearing should be extracted. According to previous studies, the fault features such as permutation entropy [18, 19], subband energy [20], and statistical features (variance, kurtosis) [13] in frequency domain, time domain, and time frequency domain could be extracted. However, the methods above are not stable for complex signal. As a number of non-Gaussian signals have an impulsive property and heavy tail in engineering, alpha-stable distribution has been widely applied in various fields [21–23]. Three estimation methods for alpha-stable distribution were comparatively analyzed theoretically [21], and the results showed that the stability and estimation accuracy of empirical characteristic function method (ECF) ranked first. The kurtogram and stable parameter α have been proposed to detect incipient bearing faults in [22]. After the analysis on stability and sensitivity of parameters, optimal parameters are selected for bearing fault diagnosis [23].

However, the estimation of the symmetry parameter β and location parameter μ can be calculated with the estimated value of α and γ , resulting in cumulative error propagation of such β and μ . Meanwhile, because the characteristic function of α stable distribution is intermittent in $x = 1$, the estimation error is particularly serious in $\beta \neq 0$ and $\alpha \rightarrow 1$. Furthermore, the geometry of the bearing structure is symmetric, so the SaS distribution is a more accurate statistical model to describe the bearing signals. Therefore, to enhance the computational efficiency and recognition accuracy of rolling bearings diagnosis, it attempts to extract fault feature using symmetric α stable distribution in this paper.

After feature extraction and selection, the early fault of axle box bearing should be detected via the classification of the selected fault characteristics. Recently, based on statistical learning theory, support vector machine is widely used in pattern classification and fault diagnosis of rotating machinery due to its high classification accuracy [18, 19, 24–27]. For its low complexity and improved computational efficiency, LS-SVM has better performance in applications. The kernel function of LS-SVM is critical for a better classification result. Multiple kinds of kernel functions including Polynomial Kernel, Gaussian Kernel, and Sigmoid Kernel are applicable. In order to obtain better performance, WSVM is proposed here with the combination of Morlet wavelet kernel and SVM. Compared with RBF kernel, the Morlet wavelet kernel shows a more reasonable hyperplane. Thus, this article will employ LS-SVM with wavelet kernel function and optimized parameters by PSO to enhance the accuracy of fault diagnosis.

The remainder of this paper is organized as follows. We briefly describe the FNLM and EEMD denoising methods in Section 2. The introduction of the feature extraction based on symmetric alpha-stable (SaS) distribution is presented in Section 3. Section 4 describes the proposed PSO-LSWSVM

method. In Section 5 the proposed approach is validated by experimental data. The conclusion is drawn in Section 6.

2. Denoising and Feature Extraction

2.1. Fast Nonlocal Means Algorithm. With the additive noise models, the definition of noise signals can be expressed as $y(i) = u(i) + n$, where u is the true signal and n is additive noise. For a given sample i , the estimate of signal u is a weighted sum of values within their neighbourhood $N(i)$:

$$\hat{u}(i) = \frac{1}{Z(i)} \sum_{j \in N(i)} \omega(i, j) y(i), \quad (1)$$

where $Z(i) = \sum_j \omega(i, j)$, and the weights are [14]

$$\begin{aligned} \omega(i, j) &= \exp\left(-\frac{\sum_{\sigma \in \Delta} (y(i + \sigma) - y(j + \sigma))^2}{2L_{\Delta}\lambda^2}\right) \\ &= \exp\left(-\frac{d^2(i, j)}{2L_{\Delta}\lambda^2}\right). \end{aligned} \quad (2)$$

This similarity is measured via the weighted Euclidean distance. The weight $\omega(i, j)$ takes a large value if the patch i is similar to the patch j and vice versa. In (2), λ is a bandwidth parameter, while Δ stands for a local patch of samples surrounding i , with L_{Δ} samples included. To reduce the computing time, the fast NLM has been proposed. For a signal of length N , given a translation vector d_x , S_{d_x} corresponds to the discrete integration of the squared difference of the sample y and its translation by d_x .

$$S_{d_x}(p) = \sum_{k=0}^p (y(k) - y(k + d_x))^2, \quad p \in N. \quad (3)$$

Now let $d_x = j - i$ and define $q = i + \sigma$; the patch size is $\Delta = [-P, P]$. Thus, $d^2(i, j)$ can be rewritten as follows:

$$d^2(i, j) = \sum_{i-p}^{i+p} (y(q) - y(q + d_x))^2. \quad (4)$$

We split the sum and use the identity in (3); we obtain

$$d^2(i, j) = S_{d_x}(i + P) - S_{d_x}(i - P). \quad (5)$$

This is the key expression that computes the weight for a pair of pixels in constant time.

2.2. Ensemble Empirical Mode Decomposition. As an improved version of EMD, EEMD can decrease the mode mixing effect. The algorithm can be given as follows [17].

(1) Add white noise $n_i(t)$ with the given amplitude to the original signal $x(t)$ to generate a new signal:

$$x_i(t) = x(t) + n_i(t), \quad (6)$$

where $x_i(t)$ represents the noise-added signal of the i th trial, while $i = 1, 2, \dots, M$.

(2) With EMD algorithm, the signal $x_i(t)$ is decomposed into some IMFs.

$$x_i(t) = \sum_{s=1}^S c_{i,s}(t) + r_{i,s}(t), \quad (7)$$

where $c_{i,s}(t)$ represents the IMFs ($c_{i,1}, c_{i,2}, \dots, c_{i,s}, \dots, c_S$), $r_{i,s}(t)$ is the final residue, and S is the number of IMFs.

(3) Repeat steps (1) and (2) while $i < M$, with various white noise series every time to acquire an ensemble of IMFs.

(4) Ensemble means of the corresponding IMFs of the decomposition is calculated; the final result is as follows:

$$c_s(t) = \frac{1}{M} \sum_{i=1}^M c_{i,s}(t), \quad (8)$$

where $c_s(t)$ is the s th IMF decomposed by EEMD, while $i = 1, 2, \dots, M$, and $s = 1, 2, \dots, S$. The IMFs include different frequency bands ranging from high to low. In this study, the first five IMFs are chosen for analysis.

3. Symmetric Alpha-Stable Distribution

3.1. Alpha-Stable Distribution. It is found that alpha-stable distribution can provide useful models for non-Gaussian signals with impulsive waveform and heavy tail probability density. Since the probability density function of an alpha-stable random variable cannot be given in a closed-form, the characteristic function can always be given as follows:

$$\varphi(t) = \exp(i\mu t - \gamma^\alpha |t|^\alpha (1 - i\beta \text{sign}(t) \omega(t, \alpha))), \quad (9)$$

where

$$\omega(t, \alpha) = \begin{cases} -\tan\left(\frac{\pi\alpha}{2}\right), & \alpha \neq -1 \\ \frac{2 \log |t|}{\pi}, & \alpha = 1 \end{cases} \quad (10)$$

$$\text{sign}(t) = \begin{cases} 1, & t > 0 \\ 0, & t = 0 \\ -1, & t < 0. \end{cases}$$

Thus, the characteristic function is a four-parameter family of distribution and is denoted by $S(\alpha, \beta, \gamma, \mu)$. The first parameter α ($0 < \alpha \leq 2$) is the characteristic exponent which describes the tail of the density function. The second parameter β ($-1 \leq \beta \leq 1$) is called symmetric parameter controlling the skewness. The parameters γ ($\gamma > 0$) and μ ($-\infty < \mu < +\infty$) are the scale parameter and the location parameter, respectively.

3.2. Symmetric Alpha-Stable Distribution. In the case of $\beta = 0$, the distribution is symmetric about μ , called symmetric alpha-stable (S α S), which has a characteristic function such that

$$\varphi(t) = \exp(i\mu t - \gamma^\alpha |t|^\alpha). \quad (11)$$

Furthermore, as the large estimation error, the location parameter μ cannot describe the health condition of bearings. Thus, the parameter μ is set to be zero so as to improve processing speed; the characteristic function could be rewritten as

$$\varphi(t) = \exp(-\gamma^\alpha |t|^\alpha). \quad (12)$$

3.3. Empirical Characteristic Function Parameter Estimation Method. In practical applications of engineering, the real-time parameter estimation of random sequence is crucial in alpha-stable distribution. In the literature, there are three major methods used to obtain the parameter value: (1) quantiles method, (2) logarithmic moment method, and (3) empirical characteristic function method. By comparative analysis [21], the empirical characteristic function approach has the highest estimation accuracy for four parameters of alpha-stable distribution with best stability. The parameter estimation process based on ECF is described as follows [28]:

(1) Calculating the sample characteristic function is as follows:

$$\widehat{\varphi}(t) = \frac{1}{n} \sum_{k=1}^n \exp(jt x_k), \quad (13)$$

where x_k ($k = 1, 2, \dots, n$) is the sample of a random variable.

(2) Equation (14) can be easily achieved based on (12):

$$\log(-\log |\varphi(t)|^2) = \log(2\gamma^\alpha) + \alpha \log(t). \quad (14)$$

(3) The characteristic exponent α and the scale parameter γ are acquired by linear regression estimation:

$$y_k = m + \alpha w_k + \varepsilon_k \quad (15)$$

$$y_k = \log(-\log(\widehat{\varphi}(t_k))^2),$$

where $m = \log(2\gamma^\alpha)$, $w_k = \log(t_k)$, $t_k = \pi k/25$, ε_k is the random error, $k = 1, 2, \dots, K$, and parameter K is accessible judging by Table 1 in [28].

The S α S distribution densities with different α and γ value are shown in Figure 1. For the fault bearing signal, the defect characteristic parameters, such as the exponent α , the scale parameter γ , and the maximum PDF (MPDF) value which represent the healthy status of the bearing, can be gained by the S α S distribution method.

4. Bearing Defect Diagnosis Methodology Based on PSO-LS-WSVM

4.1. Morlet Wavelet Kernel-Based SVM. Support vector machines (SVM) have been shown to be effective for many classification problems. The SVM, a new supervised machine learning technique on the basis of statistical learning theory, is designed to find optimal hyperplanes among diverse kinds of input training data in high dimensional feature space. Subsequently, with separating hyperplanes, testing data could be sorted. In order to reduce computing time and enhance recognition accuracy, the least squares SVM (LS-SVM) was

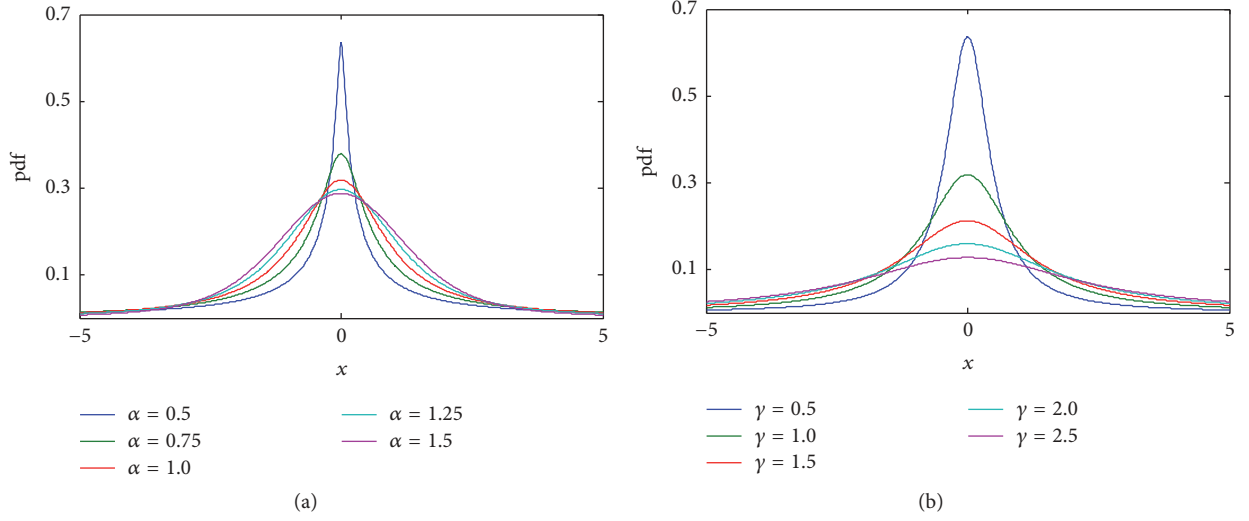


FIGURE 1: (a) The pdf of SαS with different α values; (b) the pdf of SαS with different γ values.

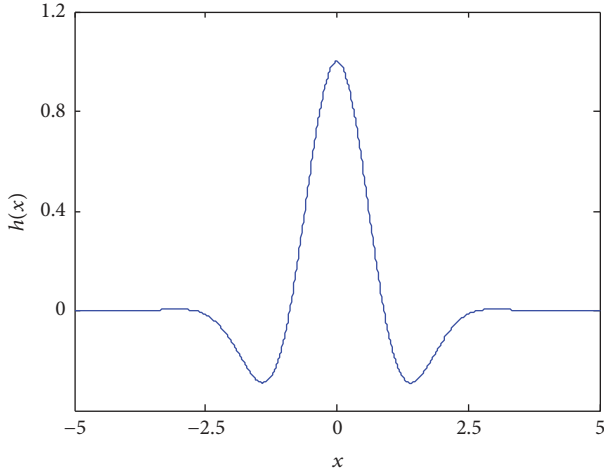


FIGURE 2: The Morlet wavelet function.

proposed. Kernel mapping is applied to map the data in input space to a high dimensional feature space, where the problem is linearly separable. Therefore, the kernel function is a critical factor for classification accuracy. Several types of kernel functions including Sigmoid Kernel, Gaussian (RBF) Kernel, and Polynomial Kernel are generally used in many applications; specifically, Gaussian Kernel has been widely used due to excellent performance. In recent years, the wavelet kernel as a type of multidimensional wavelet can approximate arbitrary nonlinear function, and Zhang et al. have proven that wavelet kernel is better than the Gaussian Kernel [26].

We consider that the wavelet analysis is a function with a family of functions emerging from dilating and translating of a mother wavelet function:

$$h_{a,c}(x) = |a|^{-1/2} h\left(\frac{x-c}{a}\right), \quad (16)$$

where $x, a, c \in R$, a is a dilation factor, c is a translation factor, and $h(x)$ is the mother wavelet. The product of one-dimensional wavelet function can be written as follows:

$$h(x) = \prod_{i=1}^N h(x_i), \quad (17)$$

where $\{x = (x_1, \dots, x_N) \in R^N\}$. If $x, x' \in R^N$, the dot-product wavelet kernels are

$$K(x, x') = \prod_{i=1}^N h\left(\frac{x_i - c_i}{a}\right) h\left(\frac{x'_i - c'_i}{a}\right). \quad (18)$$

And the translation invariant wavelet kernels are

$$K(x, x') = \prod_{i=1}^N h\left(\frac{x_i - x'_i}{a}\right). \quad (19)$$

With no loss of generality, people can construct Morlet wavelet functions as translation invariant wavelet kernel functions as follows:

$$h(x) = \cos(1.75x) \exp\left(-\frac{x^2}{2}\right). \quad (20)$$

The Morlet wavelet function is shown in Figure 2. Equation (19) defines the mother wavelet, of which the wavelet kernel can be described as follows:

$$\begin{aligned} K(x, x') &= \prod_{i=1}^N h\left(\frac{x_i - x'_i}{a}\right) \\ &= \prod_{i=1}^N \left(\cos\left(1.75 \times \frac{x_i - x'_i}{a}\right) \right. \\ &\quad \left. \cdot \exp\left(-\frac{\|x_i - x'_i\|^2}{2a^2}\right) \right). \end{aligned} \quad (21)$$

4.2. Particle Swarm Optimization for Parameter of LS-WSVM. Particle swarm optimization (PSO) is a population based on stochastic optimization technique inspired by social behavior of bird flocking or fish schooling [29]. Compared to genetic algorithm (GA) [30], the advantages of PSO are easy to implement and there are few parameters to adjust. Thus, it shows better performance with optimization problems.

In PSO, suppose that the search space is D -dimensional; there are n particles in the population. The position of the particle i in generation t is expressed as D -dimensional vector, $x_i(t) = (x_{i1}(t), x_{i2}(t), \dots, x_{iD}(t))$. The position represents the particle velocity vector, $v_i(t) = (v_{i1}(t), v_{i2}(t), \dots, v_{iD}(t))$. The position and velocity of each particle are replaced continuously according to the formula as follows:

$$\begin{aligned} v_{id}^{t+1} &= \beta \cdot v_{id}^t + c_1 r_1 \cdot (p_{best,id}^t - x_{id}^t) + c_2 r_2 \\ &\quad \cdot (g_{best,d}^t - x_{id}^t) \\ x_{id}^{t+1} &= x_{id}^t + v_{id}^{t+1}, \end{aligned} \quad (22)$$

where t is the updated iteration of the particle. The best position of the i th particle in D -dimensional search space can be recorded as follows: $p_{best,i}(t) = (p_{best,i1}(t), p_{best,i2}(t), \dots, p_{best,iD}(t))$, and the best position in whole swarm is recorded as follows: $G_{best,i}(t) = (g_{best,i1}(t), g_{best,i2}(t), \dots, g_{best,iD}(t))$. c_1 and c_2 are the acceleration constants. r_1 and r_2 are two independent random parameters which obey the uniform distribution ranging for $[0, 1]$. Parameter β , being applied to resolve the capabilities of global and local exploration, should be set as follows:

$$\beta_k = \beta_{\max} - \frac{\beta_{\max} - \beta_{\min}}{\text{iter}_{\max}} \times \text{iter}, \quad (23)$$

where β_{\min} is the minimal inertia weight and β_{\max} is the maximal inertia weight, iter is the current iteration number, and iter_{\max} is the maximum iteration number. The optimization procedure is illustrated in Figure 3.

5. The Proposed Intelligent Bearing Fault Diagnosis Method and Experimental Results

5.1. The Proposed Intelligent Bearing Fault Diagnosis Methodology. On the basis of the superiorities of FNLM, EEMD, SaS, and PSO-LS-WSVM, researchers put forward a new bearing fault diagnosis approach, with the purpose of sorting multiple and normal types of faulty bearing. Figure 4 shows the proposed procedure and the steps are displayed as follows.

- (1) Samples of vibration signals are taken by acceleration sensors at a particular sampling frequency under various operating conditions.
- (2) FNLM and EEMD methods are applied to preprocess the vibration-based signals, aiming at acquiring a range of IMF components. Subsequently, for feature extraction, the first five IMF components with more important state information are chosen.

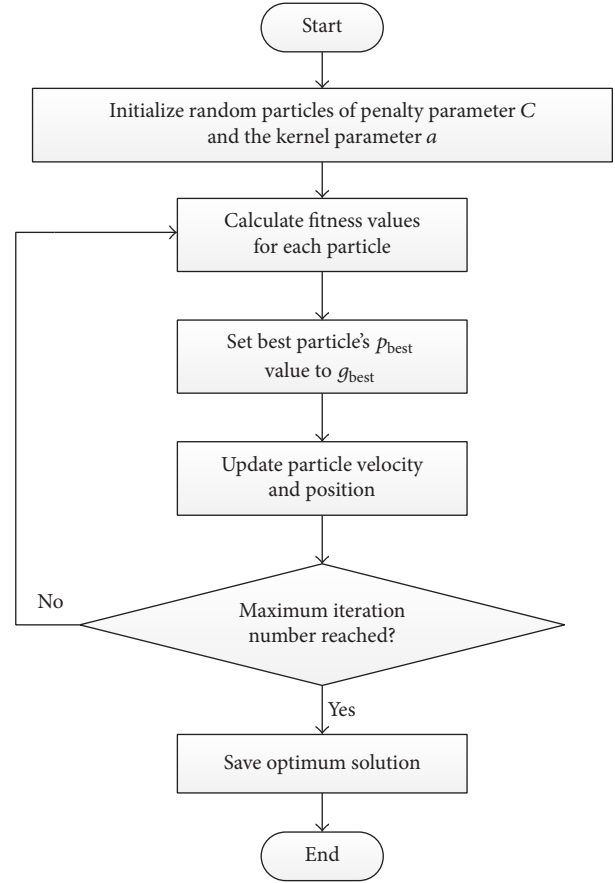


FIGURE 3: Flowchart of the PSO optimization procedure.

- (3) Extract the α , γ , and MPDF feature parameters using SaS; then choose the best IMF with feature parameters which describe the bearing health condition to construct the new fault feature vector.
- (4) The obtained feature data is partitioned into training samples and testing samples.
- (5) The Morlet wavelet kernel function is chosen for LS-WSVM classification and its parameters are optimized by PSO.
- (6) Put the training and testing samples into the classifier to conduct automatic fault diagnosis.

5.2. Experimental Results. With the purpose of examining the effectiveness of the proposed approach, the axle box bearing vibration data are used as an example. Figure 5 displays the experimental test on the axle box bearings of railway conducted in test rig. The test rig for data acquisition consists of two motors, two friction wheels, hydraulic loading installation, and control electronics (not shown). Experimental bearing is mounted to the wheelset which is fixed by installation of the test rig, and the wheelset is driven by a friction wheel. Figure 6 displays the experimental test on the axle box bearings of railway conducted in test rig. Every fault condition consists of two sizes: the width is set as 0.1 mm; the depth is set as 0.23 mm and 0.43 mm, respectively. An

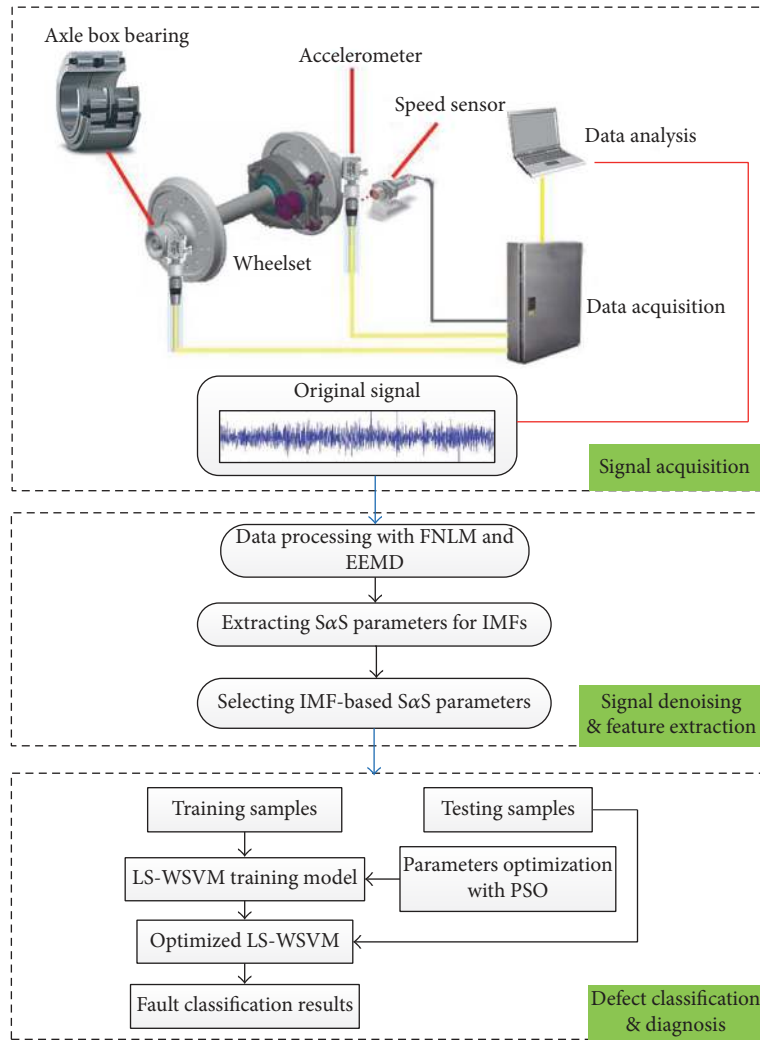


FIGURE 4: Structure diagram of the proposed fault diagnosis algorithm.

TABLE 1: Geometrical parameters of the bearings.

Axle box bearing	Outside diameter (mm)	Inside diameter (mm)	Pitch diameter (mm)	Thickness (mm)	Ball diameter (mm)	Ball thickness (mm)
NJ(P)3226X1	250	130	190	80	32	52

accelerometer, attached to the housing with glue, is applied to collect vibration data. As the occupation at six o'clock of housing, the accelerometers are located at the two o'clock position on the housing, with the sampling frequency of 25.6 kHz. The parameters of the axle box bearings are shown in Table 1; and defect frequencies of the axle box bearings are shown in Table 2.

As shown in Table 2, the present research needs to distinguish 7 classes in total. For each condition, 80 samples can be obtained. The gathered original signals are classified into training samples and testing samples for each condition, with each sample containing 5000 data points. The training samples are used to train the classifier model and the testing samples are used to evaluate the effectiveness of the proposed fault diagnosis methods.

Figure 7 shows the typical waveforms in time domain by FNLM denoising method and EEMD decomposing algorithm. Generally, defect information is contained in the first five IMF components, which could be utilized to extract defect features using S α S. The detailed steps to extract S α S features have already been discussed in Section 3.2.

Table 3 shows characteristic exponent values of all five modes, and the third mode (c3) performs a better fault indication, where three values observed under normal situation and abnormal situations with inner race fault (0.43 mm) and roller fault (0.43 mm) are 1.1757, 1.1618, and 1.1409, respectively. The above three values are supposed to be pretty close. Therefore, single alpha values of mode c3 fail to show the difference of bearing healthy status. Scale parameter and MPDF value of mode c3 in Tables 4 and 5 show the distinctions

TABLE 2: Specifications of the bearing defects.

Fault ID	Defect type	Defect size (mm)	Rotating speed (rpm)	Defect frequency (Hz)
F1	Normal	—	230	—
F2	Outer race (OR)	0.23	230	22.31
F3	Outer race (OR)	0.43	230	22.31
F4	Inner race (IR)	0.23	230	31.35
F5	Inner race (IR)	0.43	230	31.35
F6	Rolling element	0.23	230	11.04
F7	Rolling element	0.43	230	11.04

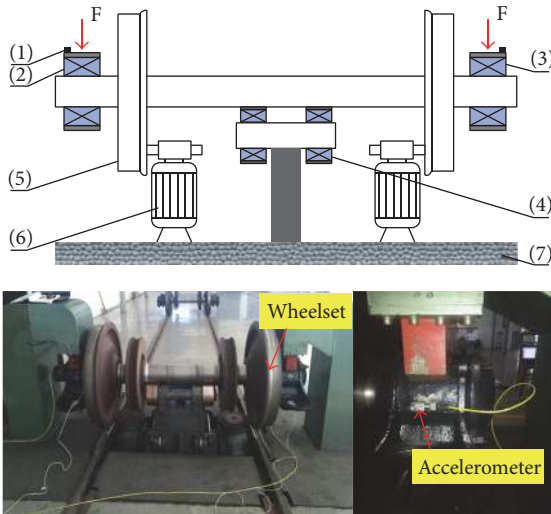


FIGURE 5: Experimental setup for vibration signal acquisition. (1) Accelerometer; (2) normal axle box bearing; (3) fault axle box bearing; (4) installation of wheelset; (5) friction wheel; (6) induction motor; (7) ground.

TABLE 3: Characteristic exponent α values of IMFs (c1–c5).

Fault ID	IMF1	IMF2	IMF3	IMF4	IMF5
F1	2	1.9838	1.1757	2	1.9888
F2	1.9946	0.9609	0.7953	2	1.9187
F3	2	1.3545	1.1618	2	1.9781
F4	2	1.5353	1.0183	2	2
F5	2	1.9849	1.8908	2	1.9614
F6	2	1.9612	1.2997	1.8759	2
F7	2	1.9279	1.1409	1.9920	2

under different bearing fault conditions. Thus, characteristic exponent, scale parameter, and MPDF value have been combined to describe the working condition of axle box bearing. The parameters α , γ and MPDF value of 40 training samples are aligned in Figures 8(a), 8(b), and 8(c), respectively. It can be seen that anyone of the three parameters cannot identify different fault types, but when the three parameters were combined, samples of the same class exhibit excellent clustering result in Figure 8(d).

TABLE 4: Scale parameter γ values of IMFs (c1–c5).

Fault ID	IMF1	IMF2	IMF3	IMF4	IMF5
F1	0.0028	0.0017	0.0182	0.0437	0.0113
F2	$9.4197e-04$	$9.9551e-04$	0.0152	0.0115	0.0040
F3	0.0015	0.0013	0.0258	0.0107	0.0038
F4	$9.2164e-04$	$6.7810e-04$	0.0094	0.0091	0.0037
F5	0.0035	0.0021	0.0097	0.0268	0.0110
F6	$8.8241e-04$	$5.6481e-04$	0.0054	0.0118	0.0135
F7	$6.8524e-04$	$4.5847e-04$	0.0073	0.0084	0.0047

TABLE 5: MPDF values of IMFs (c1–c5).

Fault ID	IMF1	IMF2	IMF3	IMF4	IMF5
F1	99.2876	162.8502	16.5710	6.4530	24.9091
F2	299.4869	325.4653	23.8567	24.6207	71.4203
F3	182.225	230.6710	11.7176	26.2765	74.7419
F4	306.0803	422.6269	33.5958	31.0184	77.1418
F5	80.7348	134.6213	29.1551	10.5104	25.6892
F6	319.6866	499.6525	54.0628	24.0175	20.8936
F7	411.6718	615.8217	41.7425	33.5575	59.5689

After feature extraction, the different feature sets, including α , γ , MPDF, and hybrid feature set, are used as input to the wavelet-based LS-SVM for fault diagnosis. Based on experience and experimental tests, computation complexity is taken into consideration in the experiment and the parameters of PSO optimization are set as follows: the number of particles is set as 20, the acceleration constants both are set as 2.0, and the evolutionary generation is set as 100. As shown in Table 6, when using 40 training samples and 40 testing samples as the input to proposed classifiers, the classifiers yield recognition rates, 88.57% for the α feature, 92.86% for the γ feature set, 93.57% for the MPDF feature set, and 95.71% for the hybrid feature set. It shows that the hybrid feature set contains more information characterizing the condition of axle box bearing. Moreover, Figure 9 shows that the consuming time of Morlet wavelet kernel is shorter than RBF kernel.

For the reason that the performance of diagnosis methods is closely related to the amount of training samples, we will study the recognition rates in different samples. In the testing

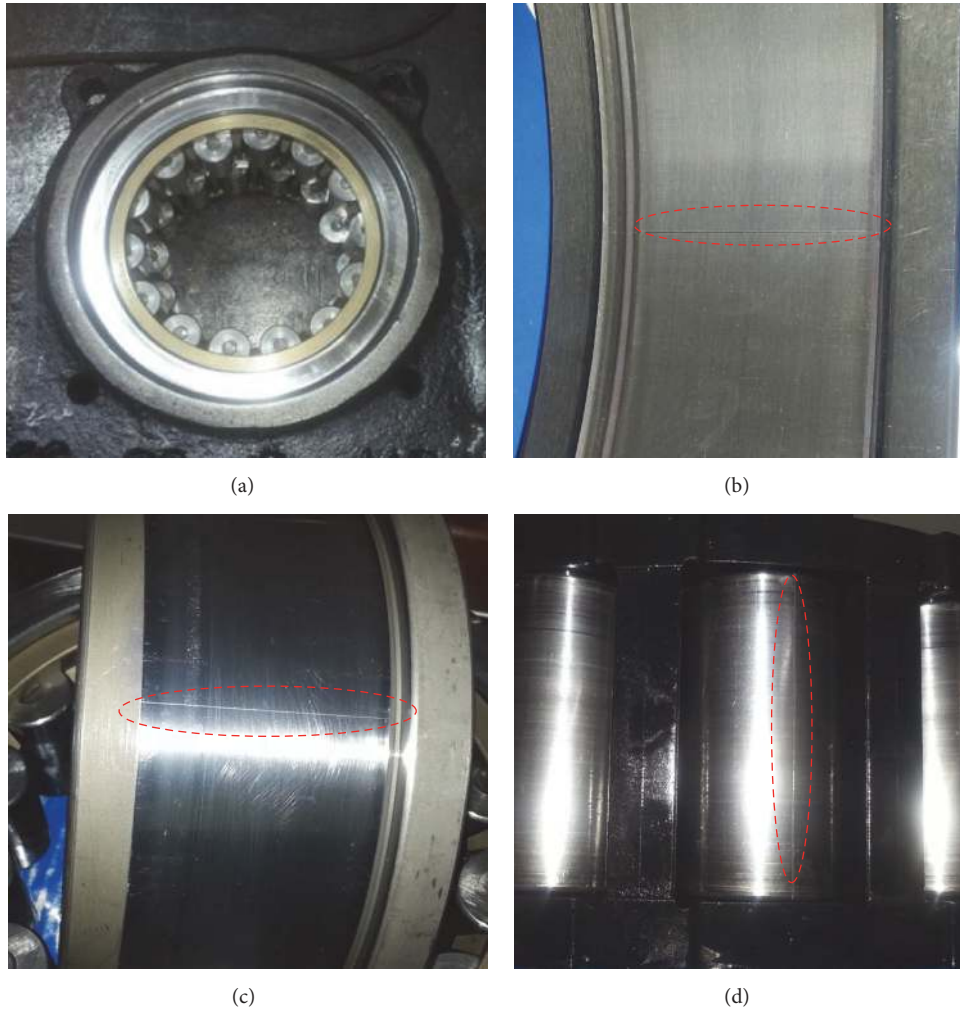


FIGURE 6: Artificial defects on the components of the axle box bearing: (a) axle box bearing; (b) defect on the outer race; (c) defect on the inner race; (d) defect on the rollers.

TABLE 6: Fault diagnosis result of the proposed fault diagnosis method based on S&S and PSO-LS-WSVM.

Fault feature	Faults classification result (%)							Average accuracy
	F1	F2	F3	F4	F5	F6	F7	
α	80	100	100	100	55	100	85	88.57
γ	100	80	85	100	85	100	100	92.86
MPDF	100	85	80	90	100	100	100	93.57
Hybrid feature set	100	90	100	95	85	100	100	95.71

experiment, 5, 10, 20, 30, and 40 samples for each class as the training and testing set were randomly selected, aiming at assessing the classification accuracy of diverse methods, respectively. To show the classification accuracy of wavelet-based LS-SVM and RBF-based LS-SVM in different numbers of samples, the above training sets were input to the classifiers,

respectively. The comparing recognition results are shown in Figure 10, which proves that the proposed method has reached higher recognition accuracy than that of RBF-based LS-SVM in different training sample. With the increase of the number of samples, the classification accuracy rate is also rising, and the proposed approach showed good performance in the case of a very small number samples.

To obtain the better recognition accuracy, many optimization algorithms including PSO, GA, and Grid Search (GS) were combined with LS-SVM classifier. In this paper, the PSO algorithm is used in our work; thus the GA and Grid Search algorithm are compared with PSO in optimizing parameters. The parameters of GA are set as follows: the population size is set as 20, the iteration number is set as 100, and the crossover probability and the mutation probability are set as 0.5 and 0.1, respectively. The comparison result of 40 samples for each fault class is shown in Table 7 and Figure 11; the classification result of PSO-LS-WSVM is 95.71%, in comparison with the 92.14% and 93.57% using GS-LS-WSVM

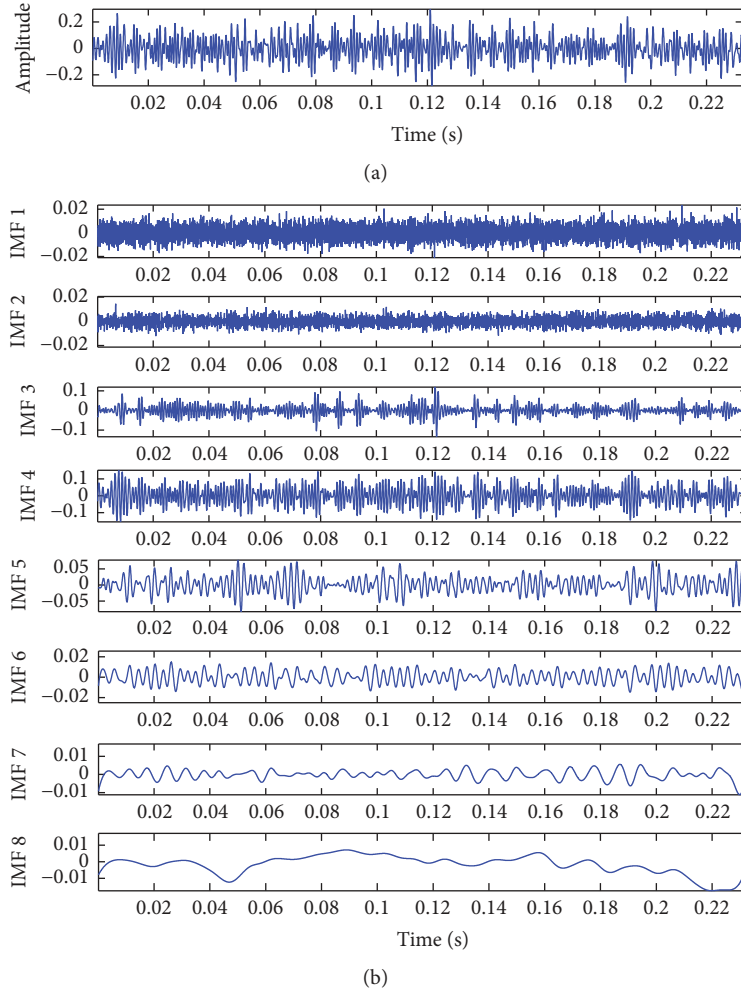


FIGURE 7: (a) Denoised signal after FNLM and (b) first 8 IMF components after EEMD.

TABLE 7: Recognition accuracies of the feature classification with different methods.

Fault classification method	Faults classification result (%)							Average accuracy
	F1	F2	F3	F4	F5	F6	F7	
GS-LS-WSVM	95	82.5	100	90	77.5	100	100	92.14
GA-LS-WSVM	100	87.5	100	85	82.5	100	100	93.57
PSO-LS-WSVM	100	90	100	92.5	87.5	100	100	95.71

and GA-LS-WSVM, respectively. As shown in Figure 11, the consuming time of PSO-LS-WSVM is longer than that of the other two approaches. The main reason is that PSO is not good at binary coding. Moreover, the average classification accuracy of wavelet-based LS-SVM optimized by PSO, GA, and GS in different number training sets and testing sets is compared. From Figure 12, it can be seen that the recognition rate of PSO algorithm is obviously higher than that of the other two methods. It can be concluded that the classification accuracy is affected by the number of training samples.

It should be noted that the Morlet-LSSVM method has better performance than RBF-LSSVM for small feature dataset. From the above analysis, the result of RBF-LSSVM is seriously affected by the number of training samples in small

size samples. The training time and the testing time of the classifiers rely on the sample size and coded programming. Hence, under the same condition, the smaller the sample size, the less the time it consumes. Furthermore, the consuming time of the Morlet-LSSVM method is less than that of the RBF-LSSVM. This phenomenon may be attributed to the fact that the Morlet kernel is approximately orthonormal, but the RBF kernel is not.

6. Conclusion

We proposed a novel bearing multifault diagnosis method based on FNLM and EEMD for denoising, symmetric alpha-stable distribution (S α S) for feature calculation, and an

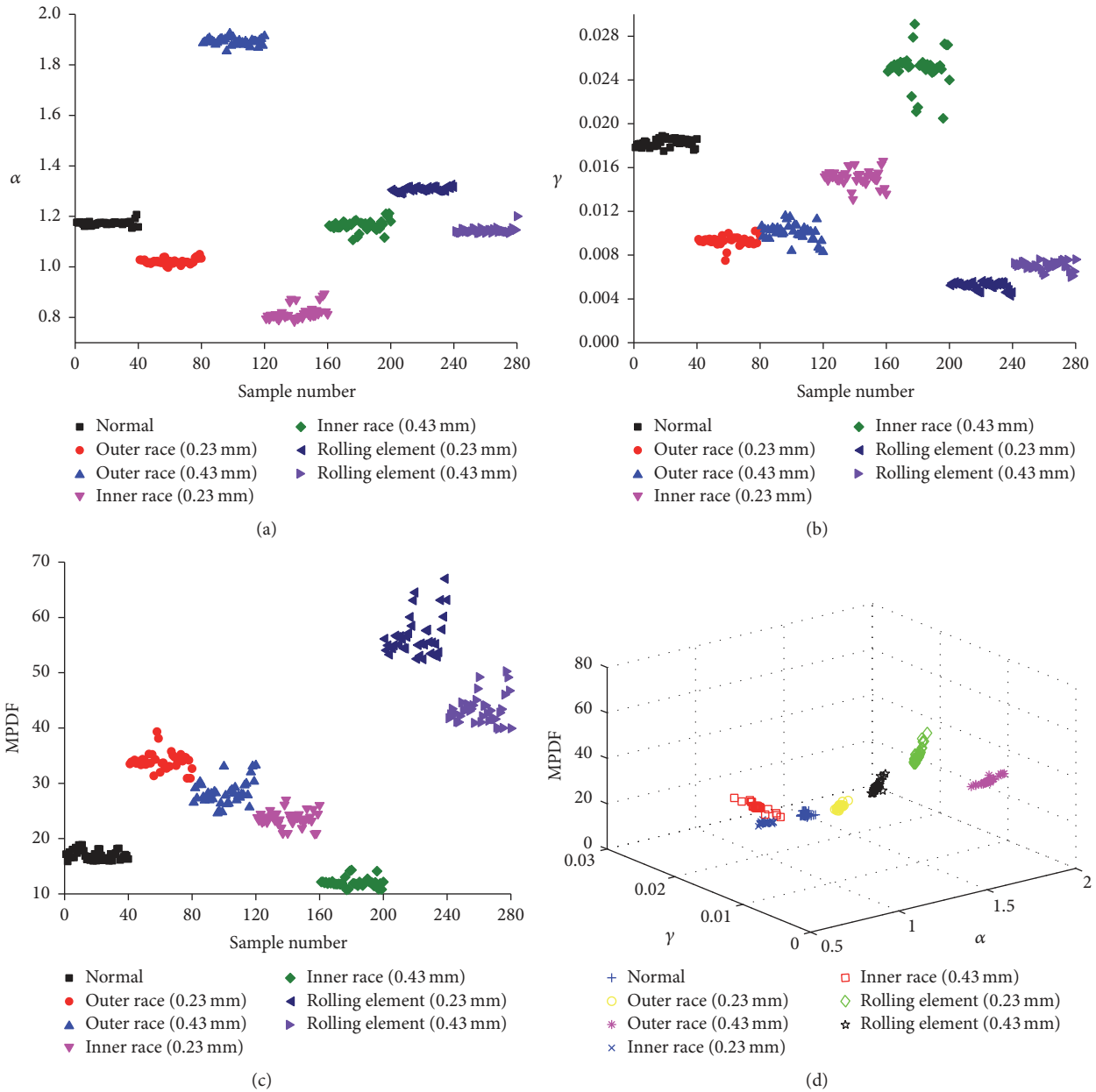


FIGURE 8: Illustration of S α S features of c3: (a) α feature parameter; (b) γ feature parameter; (c) MPDF feature parameter; (d) combined α - γ -MPDF clustering result.

appropriate PSO-LS-WSVM classifier. The results of experiment suggest that the denoised method FNLM-EEMD improves efficiency of defect feature extraction, and the proposed S α S parameter extraction method is capable of making the most discriminate and efficient features for fault diagnosis. By comparing combinations of S α S feature parameter with the LS-SVM based Morlet wavelet kernel and RBF-based classifiers and then optimizing with different algorithm, respectively, the classification capacity of the above classification methods has been studied under various sizes of training and testing samples. All results reveal that

the wavelet-LSSVM has better performance than the RBF-LSSVM when the size of data sample is very small. For its higher recognition accuracy and computational efficiency, the bearing fault diagnosis based on FNLM-EEMD, S α S, and the PSO-LS-WSVM classifier is an effective and powerful tool for monitoring the health status of axle box bearings.

Competing Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

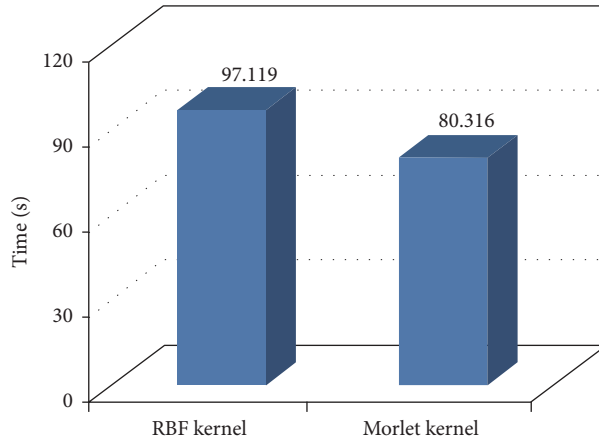


FIGURE 9: The consuming time of RBF kernel and Morlet kernel, respectively.

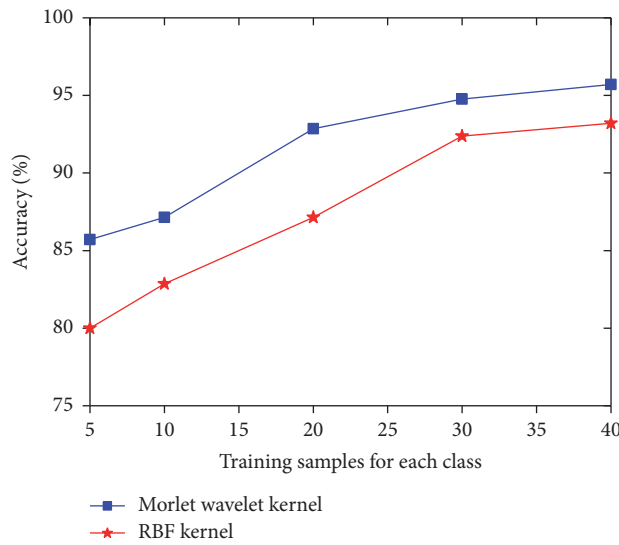


FIGURE 10: The diagnostic accuracy based on the Morlet wavelet kernel and RBF kernel LS-SVM optimized by PSO, respectively.

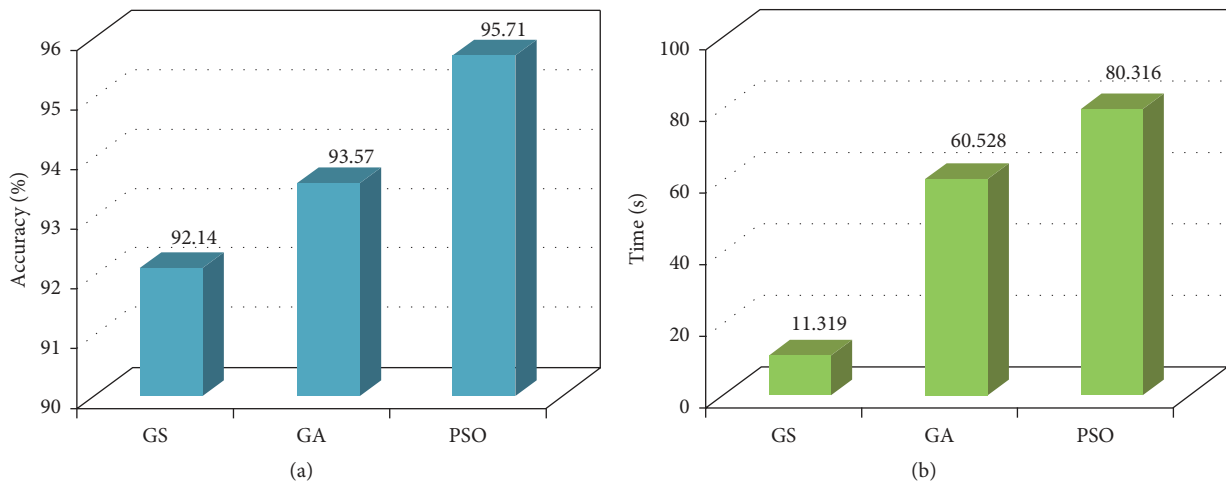


FIGURE 11: (a) The diagnostic accuracy based on the LS-WSVM optimized by PSO, GA, and GS, respectively; (b) the consuming time of PSO, GA, and GS, respectively.

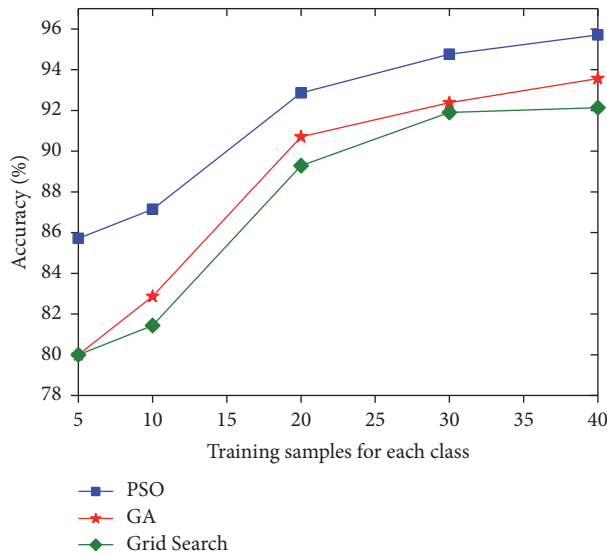


FIGURE 12: The diagnostic accuracy based on the LS-WSVM optimized by PSO, GA, and GS, respectively.

Acknowledgments

This work is supported by National Natural Science Foundation of China (no. U1234208) and the Fundamental Research Funds of the Traction Power State Key Laboratory, Southwest Jiaotong University (2013TPL-T04).

References

- [1] R. Lagneback, *Evaluation of wayside condition monitoring technologies for condition-based maintenance of railway vehicles [Licentiate thesis]*, Lulea University of Technology, Lulea, Sweden, 2007.
- [2] W. H. Sneed and R. L. Smith, "On-board real-time railroad bearing defect detection and monitoring," in *Proceedings of the ASME/IEEE Joint Railroad Conference*, pp. 149–153, IEEE, Philadelphia, Pa, USA, April 1998.
- [3] H. Zhang, S. Lu, Q. He, and F. Kong, "Multi-bearing defect detection with trackside acoustic signal based on a pseudo time-frequency analysis and Dopplerlet filter," *Mechanical Systems and Signal Processing*, vol. 70-71, pp. 176–200, 2016.
- [4] G. He, K. Ding, and H. Lin, "Fault feature extraction of rolling element bearings using sparse representation," *Journal of Sound and Vibration*, vol. 366, pp. 514–527, 2016.
- [5] J. Gao, H. Sultan, J. Hu, and W.-W. Tung, "Denoising nonlinear time series by adaptive filtering and wavelet shrinkage: a comparison," *IEEE Signal Processing Letters*, vol. 17, no. 3, pp. 237–240, 2010.
- [6] A. Bouzida, O. Touhami, R. Ibtouen, A. Belouchrani, M. Fadel, and A. Rezzoug, "Fault diagnosis in industrial induction machines through discrete wavelet transform," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 9, pp. 4385–4395, 2011.
- [7] J. S. Luo, D. J. Yu, and M. Liang, "A kurtosis-guided adaptive demodulation technique for bearing fault detection based on tunable-Q wavelet transform," *Measurement Science and Technology*, vol. 24, no. 5, Article ID 055009, 2013.
- [8] A. Soualhi, K. Medjaher, and N. Zerhouni, "Bearing health monitoring based on hilbert-huang transform, support vector machine, and regression," *IEEE Transactions on Instrumentation and Measurement*, vol. 64, no. 1, pp. 52–62, 2015.
- [9] M. Žvokelj, S. Zupan, and I. Prebil, "EEMD-based multiscale ICA method for slewing bearing fault detection and diagnosis," *Journal of Sound and Vibration*, vol. 370, pp. 394–423, 2016.
- [10] J.-H. Ahn, D.-H. Kwak, and B.-H. Koh, "Fault detection of a roller-bearing system through the EMD of a wavelet denoised signal," *Sensors*, vol. 14, no. 8, pp. 15022–15038, 2014.
- [11] Q. He, Y. Liu, and F. Kong, "Machine fault signature analysis by midpoint-based empirical mode decomposition," *Measurement Science and Technology*, vol. 22, no. 1, Article ID 015702, 2011.
- [12] V. K. Rai and A. R. Mohanty, "Bearing fault diagnosis using FFT of intrinsic mode functions in Hilbert-Huang transform," *Mechanical Systems and Signal Processing*, vol. 21, no. 6, pp. 2607–2615, 2007.
- [13] M. Amarnath and I. R. Praveen Krishna, "Empirical mode decomposition of acoustic signals for diagnosis of faults in gears and rolling element bearings," *IET Science, Measurement and Technology*, vol. 6, no. 4, pp. 279–287, 2012.
- [14] A. Buades, B. Coll, and J. M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Modeling & Simulation*, vol. 4, no. 2, pp. 490–530, 2005.
- [15] Y. Lv, Q. L. Zhu, and R. Yuan, "Fault diagnosis of rolling bearing based on fast nonlocal means and envelop spectrum," *Sensors*, vol. 15, no. 1, pp. 1182–1198, 2015.
- [16] R. Ricci, P. Pennacchi, M. Lombardi, and C. Mirabile, "Failure diagnostics of a spiral bevel gearbox using EMD and HHT," in *Proceedings of the 24th International Conference on Noise and Vibration Engineering (ISMA '10), in conjunction with the 3rd International Conference on Uncertainty in Structural Dynamics (USD '10)*, pp. 2965–2979, Leuven Belgium, September 2010.
- [17] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: a noise-assisted data analysis method," *Advances in Adaptive Data Analysis*, vol. 1, no. 1, pp. 1–41, 2009.
- [18] X. Zhang, Y. Liang, J. Zhou, and Y. Zang, "A novel bearing fault diagnosis model integrated permutation entropy, ensemble empirical mode decomposition and optimized SVM," *Measurement*, vol. 69, article 3321, pp. 164–179, 2015.
- [19] Y. Li, M. Xu, Y. Wei, and W. Huang, "A new rolling bearing fault diagnosis method based on multiscale permutation entropy and improved support vector machine based binary tree," *Measurement*, vol. 77, pp. 80–94, 2016.
- [20] D.-L. Zhao, W. Ma, W.-K. Liang, and X.-Q. Luo, "Data fusion fault diagnosis and simulation of hydroelectric units vibration," *Proceedings of the Chinese Society of Electrical Engineering*, vol. 25, no. 20, pp. 137–142, 2005.
- [21] X. Wang, K. Li, P. Gao, and S. Meng, "Research on parameter estimation methods for alpha stable noise in a laser gyroscope's random error," *Sensors*, vol. 15, no. 8, pp. 18550–18564, 2015.
- [22] G. Yu, C. Li, and J. Zhang, "A new statistical modeling and detection method for rolling element bearing faults based on alpha-stable distribution," *Mechanical Systems and Signal Processing*, vol. 41, no. 1-2, pp. 155–175, 2013.
- [23] Q. Xiong, W. Zhang, T. Lu, G. Mei, and S. Liang, "A fault diagnosis method for rolling bearings based on feature fusion of multifractal detrended fluctuation analysis and alpha stable distribution," *Shock and Vibration*, vol. 2016, Article ID 1232893, 12 pages, 2016.

- [24] R. Jegadeeshwaran and V. Sugumaran, "Fault diagnosis of automobile hydraulic brake system using statistical features and support vector machines," *Mechanical Systems and Signal Processing*, vol. 52-53, no. 1, pp. 436–446, 2015.
- [25] L. Si, Z. Wang, X. Liu, C. Tan, Z. Liu, and J. Xu, "Identification of shearer cutting patterns using vibration signals based on a least squares support vector machine with an improved fruit fly optimization algorithm," *Sensors*, vol. 16, no. 1, article 90, 2016.
- [26] L. Zhang, W. Zhou, and L. Jiao, "Wavelet support vector machine," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 34, no. 1, pp. 34–39, 2004.
- [27] F. Chen, B. Tang, and R. Chen, "A novel fault diagnosis model for gearbox based on wavelet support vector machine with immune genetic algorithm," *Measurement*, vol. 46, no. 1, pp. 220–232, 2013.
- [28] I. A. Koutrouvelis, "Regression-type estimation of the parameters of stable laws," *Journal of the American Statistical Association*, vol. 75, no. 372, pp. 918–928, 1980.
- [29] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of the IEEE International Conference on Neural Networks*, pp. 1942–1948, Perth, Australia, November 1995.
- [30] M. S. Elbisy, "Sea Wave Parameters Prediction by Support Vector Machine Using a Genetic Algorithm," *Journal of Coastal Research*, vol. 31, no. 4, pp. 892–899, 2015.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

