

## Research Article

# Mutual Information-Assisted Wavelet Function Selection for Enhanced Rolling Bearing Fault Diagnosis

Ruqiang Yan,<sup>1,2</sup> Mengxiao Shan,<sup>1</sup> Jianwei Cui,<sup>1</sup> and Yahui Wu<sup>2</sup>

<sup>1</sup>*School of Instrument Science and Engineering, Southeast University, Nanjing, Jiangsu 210096, China*

<sup>2</sup>*Changcheng Institute of Metrology and Measurement, Aviation Industry Corporation of China, Key Laboratory of Science and Technology on Metrology & Calibration, Beijing 100095, China*

Correspondence should be addressed to Ruqiang Yan; [ruqiang@seu.edu.cn](mailto:ruqiang@seu.edu.cn)

Received 28 February 2015; Revised 13 April 2015; Accepted 15 April 2015

Academic Editor: Changjun Zheng

Copyright © 2015 Ruqiang Yan 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.

This paper presents an enhanced rolling bearing fault diagnosis approach, based on optimized wavelet packet transform (WPT) assisted with quantitative wavelet function selection. Mutual information is utilized as a quantitative measure to select the most suitable wavelet function for the WPT-based vibration analysis. Energy features from coefficients of an optimal set of orthogonal wavelet subspaces which resulted from the WPT-based vibration analysis are input to different classifiers. The fault states of the rolling bearings can then be identified. Experiment studies conducted on a rolling bearing test system have verified the effectiveness of the proposed approach for rolling bearing fault diagnosis.

## 1. Introduction

In modern industry, the rotating machine is one of the most important mechanical pieces of equipment and has been widely used in each field of industrial production. The performance of the key rotating components, such as bearings and gears, has a severe influence on the working status of the whole rotating machine [1] so that a weak defect of these components may eventually lead to the machine collapse or breakdown of the production. Therefore, fault diagnosis of these key rotating components has become current research trend.

An important prerequisite for accurate diagnosis is to effectively extract characteristic features, which are indicative of faults, from signals measured on the rotating machine. As an effective tool for nonstationary signal analysis, wavelet transform and its extension, wavelet packet transform, have been widely used in feature extraction for fault diagnosis of rotating machines [2, 3]. For example, the statistical parameters were extracted from the collected signals obtained via the WPT at different decomposition depths and a support vector regressive (SVR)-based generic multiclass solver was proposed to identify the different fault patterns of rotating

machines [4]. The WPT was conducted to decompose multiclass signals into a library of time-frequency subspaces and the wavelet packet energy in each subspace was calculated to produce a feature vector for each signal for classification [5]. Renyi entropy values from subband coefficients of vibration signals were obtained to detect mechanical faults in rotational drives [6]. Such features were found to be sensitive to fault occurrence and robust to varying operating conditions. Sensitive subband feature set from the WPT-based signal decomposition was also extracted for classifying bearing faults [7]. In addition, the WPT was used to clean noisy signal before EEMD is applied to extracting informative feature vectors for early damage detection of rolling bearings [8]. An alternative empirical mode decomposition (EMD) method improved by WPT was developed to process the fault signals [9]. A preprocessing model of the bearing using WPT-EMD was constructed for feature extraction. Then it used self-organization mapping (SOM) for the condition assessment of the performance degradation [10]. Combined with manifold learning, the WPT was implemented to extract weak transient signal features for rolling bearing fault diagnosis [11]. Recently, a comprehensive review on wavelet transform for fault diagnosis of rotating machines has been conducted [3], where

the applications were summarized according to the following categories: continuous wavelet transform-based fault diagnosis, discrete wavelet transform-based fault diagnosis, wavelet packet transform-based fault diagnosis, and second generation wavelet transform-based fault diagnosis [3]. In addition, some new research trends have been discussed, among which wavelet function selection is considered as one of the important factors that affect the results of wavelet applications. This issue has also attracted more and more attention by the research community. For example, Rafiee et al. [12] used genetic algorithm to select proper Daubechies (Db) wavelet for gear fault diagnosis. They also compared 324 wavelet functions and found that Daubechies 44 wavelet showed the most similar shape across both gear and bearing vibration signals [13]. Jiang and Liu [14] applied  $t$ -test to validating the correlation between features extracted from DWT coefficients and the original signal and then applied the estimated probability from the  $t$ -test to guiding selection of wavelet functions. Yang and Ren [15] and Chen [16] utilized waveform matching method and similarity coefficient to select the optimal wavelet function for impact signal analysis. Schukin et al. [17] made use of the time-frequency window resolution and the estimation error of impact signal parameters to choose suitable wavelet functions. Li et al. [18–20] treated Shannon entropy as a measure to determine optimal wavelet function for lamb wave analysis and damage detection. Furthermore, Yan and Gao [21] investigated the energy measure for wavelet function selection in bearing vibration analysis. Later an energy-to-Shannon entropy ratio measure was also developed by Yan and Gao as a wavelet selection criterion for rotating machine fault diagnosis [22]. This criterion, together with maximum relative wavelet energy, was used by Kankar et al. to select appropriate wavelet function from six candidate wavelet functions for bearing defect-related feature extraction [23]. The energy-to-Shannon entropy ratio was also chosen by Wu et al. to guide the selection of wavelet functions for fault diagnosis of rolling bearings [24]. From these efforts, it can be seen that some of these researchers used qualitative method, such as waveform similarity method, to select wavelet functions, which largely relies on subjective assessment of the researchers. Other researchers designed quantitative measures calculated from wavelet coefficients directly to evaluate the similarity between signals and wavelet functions. In fact, the study of the relationship between wavelet coefficients and raw signals may be another outstanding angle for wavelet function selection, which has rarely been considered in previous studies in the field of fault diagnosis. Therefore, taking this relationship into consideration, this paper is intended to investigate a quantitative wavelet function selection method based on mutual information measure, with specific application for enhanced rolling bearing fault diagnosis.

This paper is organized as follows. After introducing the basic knowledge of the wavelet packet decomposition, which is used as the tool for analyzing bearing vibration signals, the mutual information-based wavelet function selection is illustrated with numerical verification in Section 2. Then Section 3 presents the fault diagnosis method based on an optimized wavelet packet transform. After that, experimental studies are performed on a bearing test system to verify

the effectiveness of the proposed method in Section 4. Finally, conclusions are drawn in Section 5.

## 2. Wavelet Function Selection

The essence of wavelet transform is to compare the signal  $x(t)$  with a set of known functions  $\{\psi_n(t)\}_{n \in \mathbb{Z}}$  and quantifies the similarity between the signal and the function. Such an operation can be mathematically expressed as

$$wt_n = \int_{-\infty}^{\infty} x(t) \psi_n^*(t) dt, \quad (1)$$

where  $(\cdot)^*$  stands for the complex conjugate of the function  $(\cdot)$ . The set of functions  $\{\psi_n(t)\}_{n \in \mathbb{Z}}$  are derived from a single wavelet function through a series of *scaling* (i.e., dilation and contraction, as denoted by the parameter  $s$ ) and *shifting* (i.e., time translation, denoted by the parameter  $\tau$ ) operations, as is expressed mathematically by

$$\psi_{s,\tau}(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-\tau}{s}\right); \quad (2)$$

then (1) can be represented as

$$wt(s, \tau) = \frac{1}{\sqrt{s}} \int x(t) \psi^*\left(\frac{t-\tau}{s}\right) dt. \quad (3)$$

Through variations of the scales and time shifts of the wavelet function, the wavelet transform can extract signal features over the entire signal by using small scales for high frequency components and large scales for low frequency components. Generally, the wavelet transform can be represented in continuous (i.e., continuous wavelet transform) as well as in discrete forms (i.e., discrete wavelet transform). Dyadic discretization of the scaling parameter  $s$  and shifting parameter  $\tau$  leads to the formulation of orthogonal basis for the set of function  $\psi_{s,\tau}(t)$ ; thus the discrete wavelet transform can be realized.

**2.1. Wavelet Packet Decomposition.** Wavelet packet decomposition is an extension of discrete wavelet transform and can be obtained by a generalization of the fast pyramidal algorithm [25]. Mathematically, a wavelet packet consists of a set of linearly combined wavelet functions, which are generated using the following recursive relationships:

$$\begin{aligned} \psi^{2k}(t) &= \sqrt{2} \sum_n h(n) \psi^k(2t-n), \\ \psi^{2k+1}(t) &= \sqrt{2} \sum_n g(n) \psi^k(2t-n), \end{aligned} \quad (4)$$

where  $\psi^0(t) = \phi(t)$  is the scaling function and  $\psi^1(t) = \psi(t)$  is the wavelet function. The symbols  $h(n)$  and  $g(n)$  represent coefficients of a pair of quadrature mirror filters (QMF) associated with the scaling function and wavelet function. Furthermore,  $h(n)$  and  $g(n)$  are related to each other by

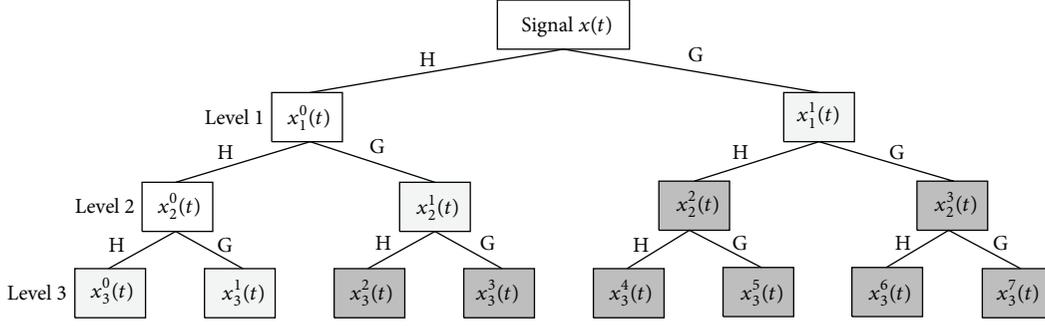


FIGURE 1: Illustration of a 3-level WPT (H-low pass filter, G-high pass filter).

$g(n) = (-1)^n h(1-n)$ . Using the QMF, a time-domain signal  $x(t)$  can be decomposed recursively as

$$\begin{aligned} x_{j+1}^{2k}(t) &= \sum_m h(2k-m) x_j^k(t), \\ x_{j+1}^{2k+1}(t) &= \sum_m g(2k-m) x_j^k(t), \end{aligned} \quad (5)$$

where  $x_j^k(t)$  denotes the wavelet packet coefficients at the  $j$ th level,  $k$ th denotes sub-frequency band, and  $x_0^0(t) = x(t)$ . The symbol  $m$  represents the number of the wavelet coefficients at the  $k$ th sub-frequency band within the level  $j$ . Figure 1 illustrates a 3-level decomposition of the signal  $x(t)$ .

**2.2. Wavelet Selection Criterion.** When the WPT is used to analyze vibration signals for rolling bearing fault diagnosis, a critical issue is to choose the most suited wavelet function for signal decomposition and feature extraction [22]. As it is known, the wavelet transform or wavelet packet transform can be considered as a kind of correlation calculation between the signals and the wavelet function in different scales. The more similar the signal and the wavelet function are, the more accurate the feature will be extracted by the wavelet analysis. In information theory, mutual information measures the information that the variables  $X$  and  $Y$  share: it measures how much knowing one of these variables reduces uncertainty about the other. In other words, the mutual information measures the degree of similarity between two groups of data sequences; it is presented in this study to direct the selection of wavelet function. Mathematically, the mutual information is described as

$$\begin{aligned} I(X; Y) &= \sum_{x \in X} \sum_{y \in Y} p(x, y) \log \frac{p(x, y)}{p(x)p(y)} \\ &= \sum_{x \in X} \sum_{y \in Y} p(x, y) \log p(x, y) \\ &\quad - \sum_{x \in X} p(x) \log p(x) - \sum_{y \in Y} p(y) \log p(y), \end{aligned} \quad (6)$$

where  $p(x, y)$  indicates the joint probability density of data sequences  $X$  and  $Y$ ,  $p(x)$  stands for the probability density of data sequence  $X$ , and  $p(y)$  represents the probability

density of data sequence  $Y$ . The mutual information determines how similar the joint distribution  $p(x, y)$  is to the products of factored marginal distribution  $p(x)p(y)$ . Furthermore, in (6),  $-\sum_{x \in X} \sum_{y \in Y} p(x, y) \log p(x, y)$  is the joint entropy of data sequences  $X$  and  $Y$ .  $-\sum_{x \in X} p(x) \log p(x)$  and  $-\sum_{y \in Y} p(y) \log p(y)$  represent the Shannon entropy of data sequence  $X$  and the Shannon entropy of data sequence  $Y$ , respectively. Therefore, (6) can be expressed in another form:

$$I(X; Y) = -H(X, Y) + H(X) + H(Y). \quad (7)$$

This equation indicates that the mutual information is the sum of the entropies  $H(X)$  and  $H(Y)$  minus the joint entropy  $H(X, Y)$ . The relationships among entropies and mutual information are illustrated in a Venn diagram as shown in Figure 2.

In Figure 2, it can be seen that the mutual information  $I(X, Y)$  is the intersection part of two data sequences. The greater the mutual information is, the more similar the two groups of data sequence will be. Such a relationship is applicable to wavelet selection for rolling bearing fault diagnosis by taking the vibration signal and wavelet packet coefficients as data sequences  $X$  and  $Y$ , respectively. It is expected that defect-induced transient vibration can be fully represented by the wavelet packet coefficients. In this study, raw vibration signals of rolling bearings are viewed as data sequence  $X$ . The reconstruction signals of the subband which contains the defect-induced transient components are viewed as data sequence  $Y$ . Then the mutual information entropy can be used to evaluate the similarity between original vibration signals and the reconstruction signals using different wavelet functions. By comparison, the wavelet function that maximizes the mutual information between the vibration signal and the reconstruction signal represents the most appropriate wavelet for defect-induced transient vibration extraction.

To evaluate the applicability of the mutual information for wavelet function selection, a simulation study is presented here. In the simulation, data sequence  $x$  is the Daubechies 10 wavelet (denoted as Db 10 wavelet). The sample signal  $x(t)$  is obtained by adding Gaussian white noise with signal-to-noise ratio (SNR) at 20 dB, which is viewed as data sequence  $X$  and decomposed by the wavelet packet transform, while the reconstruction signal is viewed as data sequence  $Y$ . The waveforms of Daubechies 10 wavelet, data sequence  $X$ , and data sequence  $Y$  are shown in Figure 3, respectively.

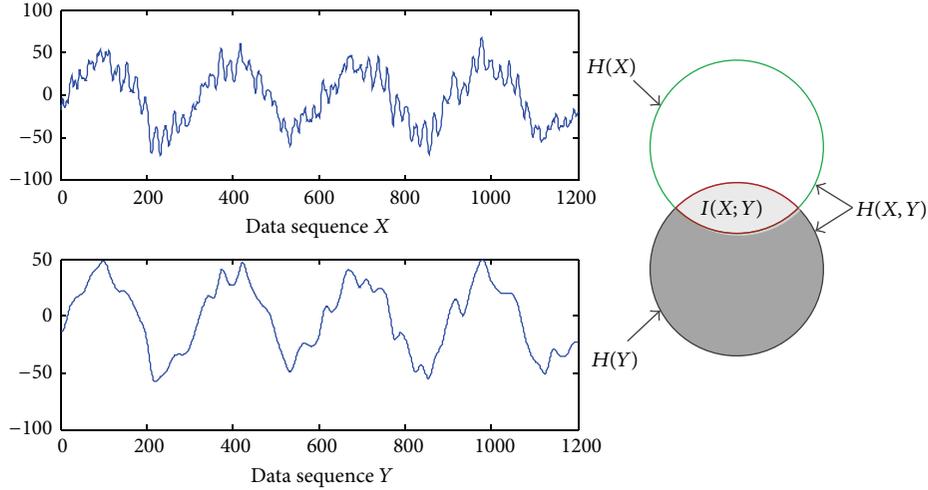


FIGURE 2: Relationships among entropies and mutual information.

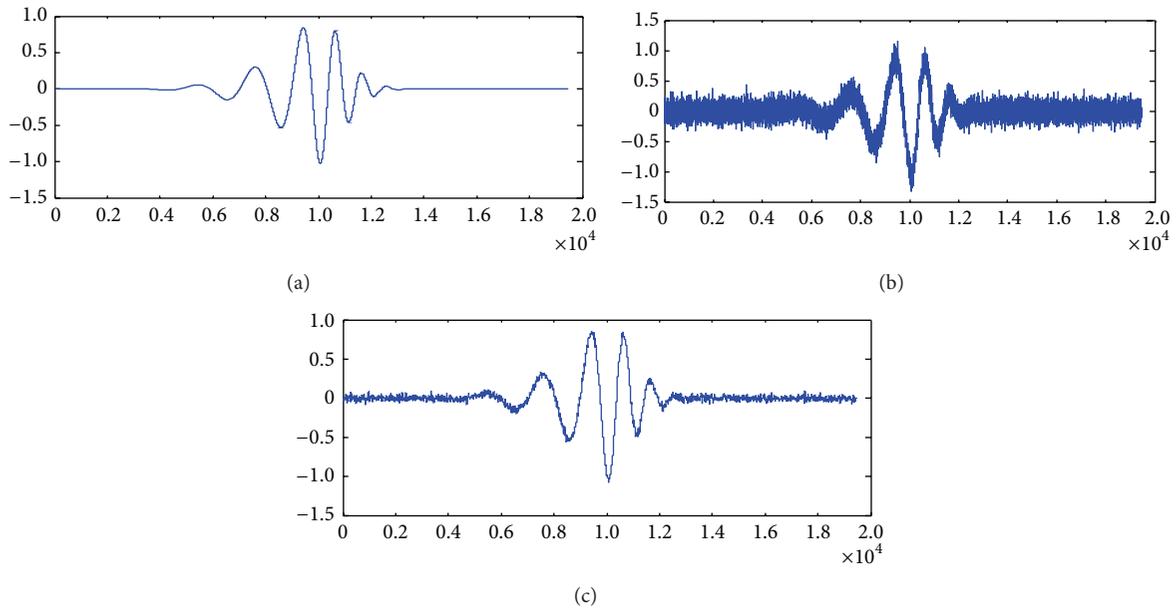


FIGURE 3: Waveform of (a) Daubechies 10 wavelet, (b) data sequence X, and (c) data sequence Y.

Then the mutual information for each wavelet function can be calculated and the results are listed in Table 1. Since the sample signal is constructed by adding Gaussian white noise to the waveform of the Db 10 wavelet, in principle, it should be better characterized by the Db 10 wavelet itself. This is verified by the results shown in Table 1, as the Db 10 wavelet maximizes the mutual information between the sample signal and the reconstruction signal, which validate the effectiveness of the proposed quantitative wavelet selection method.

### 3. Fault Diagnosis Based on Optimized Wavelet Packet Transform

With the wavelet function being chosen, the WPT can then be used to decompose the vibration signal for extracting features

to characterize rolling bearing faults. It can be seen from Figure 1 that there are multiple ways of decomposing a signal with the WPT. For the purpose of rolling bearing fault diagnosis, the challenge is to find a better representation of the signals that yields high discriminatory information among different bearing conditions. Therefore, an optimized wavelet packet transform (termed as OWPT) is applied to identifying the set of best subspaces that can provide maximum dissimilarity information between different classes of the signals, based on the local discriminant bases (LDB) algorithm [26]. The LDB algorithm is a pruning algorithm that can select an optimal set of complete orthogonal subspaces derived from the WPT. This algorithm has been successfully used in various engineering domains, such as audio signal analysis [27], physiological signal classification [28], and vibration data

TABLE 1: Mutual information resulted from the simulation signal for different wavelets.

Wavelet function	Mutual information	Wavelet function	Mutual information	Wavelet function	Mutual information
Haar	20.325	Coif5	24.302	Bior2.6	24.284
Db2	24.268	Sym2	24.288	Bior4.4	24.241
Db4	24.297	Sym3	24.273	Bior5.5	24.279
Db6	24.307	Sym4	24.257	Bior6.8	24.264
Db8	24.294	Sym5	24.274	rBio1.3	24.277
Db10	<b>24.313</b>	Sym6	24.273	rBio2.4	24.254
Coif1	24.265	Sym7	24.280	rBio2.6	24.274
Coif2	24.245	Sym8	24.284	rBio4.4	24.257
Coif3	24.286	Bior1.3	20.317	rBio5.5	24.280
Coif4	24.265	Bior2.4	24.280	rBio6.8	24.296

processing [29, 30]. Specifically, if  $A_{j,k}$  is used to represent the desired local discriminant basis restricted to the span of  $B_{j,k}$ , which is a set of basis vectors at  $(j, k)$  subspace, then, for a given training dataset consisting of  $L$  classes of signals  $\{\{x_i^{(l)}\}_{i=1}^{N_l}\}_{l=1}^L$  with  $N_l$  being the total number of training signals in class  $l$ , the LDB algorithm can be summarized as follows.

- (a) *The wavelet packet transform is used to decompose the signals contained in the training dataset, and the time-frequency energy maps  $C_l$  for  $l = 1, \dots, L$  on the wavelet packet coefficients can be constructed using the following equation:*

$$C_l(j, k, m) \equiv \frac{\sum_{i=1}^{N_l} \sum_m (x_j^k(m)_i^{(l)})^2}{\sum_{i=1}^{N_l} \|x_i^{(l)}\|^2}. \quad (8)$$

- (b) *Suppose that  $A_{J,k} = B_{J,k}$  and set  $\Delta_{j,k} = D(\{C_l(j, k, m)\}_{l=1}^L)$ , where  $D(\cdot)$  represents the dissimilarity measure to characterize the signal classification capability; then the best subspaces  $A_{j,k}$  for  $j = J - 1, \dots, 0$  and  $k = 0, \dots, 2^j - 1$  can be obtained through a bottom-up search as follows.*

*If  $\Delta_{j,k} \geq \Delta_{j+1,2k} + \Delta_{j+1,2k+1}$ , namely, the dissimilarity measure of the parent subspace is greater than those of cumulative the children subspace. Then  $A_{j,k} = B_{j,k}$ .*

*Else  $A_{j,k} = A_{j+1,2k} \oplus A_{j+1,2k+1}$  and set  $\Delta_{j,k} = \Delta_{j+1,2k} + \Delta_{j+1,2k+1}$ .*

- (c) *After a complete set of orthogonal subspaces are found in the decomposition results, they can be ranked from higher to lower based on their discrimination power.*

It should be noted that the optimal choice of the complete subspaces for a given dataset is significantly affected by the dissimilarity measures used to distinguish between classes, as the dissimilarity measure indirectly controls the classification accuracy. Therefore, a good dissimilarity measure should be able to differentiate various classes as much as possible. Considering that a single dissimilarity measure for the optimal subspace selection may not be able to capture all the characteristic information of the signals, both the relative entropy, which describes the difference of energy distribution

in different classes of signals, and the normalized energy difference between different classes of signals are used as the dissimilarity measures in this study, which have been identified as good measures for classification. The details of these two measures can be seen in [30]. Features from these subspaces can then be extracted to distinguish different classes in a given set of data that belong to several classes. Therefore, an enhanced rolling bearing fault diagnosis scheme can be designed as shown in Figure 4. Vibration signals from each of the classes are decomposed using the WPT with the selected wavelet function. Then the LDB algorithm is used to identify the set of best subspaces that provide maximum dissimilarity information between different classes of the signals. Once the optimal discriminatory subspaces are identified, the energy feature of wavelet packet coefficients from each subspace is then calculated. These features have better discriminatory capability and are chosen as inputs to a diagnostic classifier for characterizing rolling bearing faults.

## 4. Experimental Study

In order to verify the effectiveness of the proposed rolling bearing fault diagnosis approach, experimental study was conducted, where the data were collected on a bearing test system (Figure 5) from the Case Western Reserve University Bearing Data Center [31]. The system consists of a 2-horse power (hp) motor, a torque transducer, a dynamometer, and control electronics. Single point faults with the size of 0.18 mm, 0.36 mm, and 0.53 mm were set on the drive-end bearings (Type 6205-2RS JEM SKF) at the location of outer raceway, inner raceway, and rolling element (ball), respectively, using electrodischarge machining technique. The vibration data were measured by using an accelerometer being attached to the motor housing with the sampling frequency of 12 kHz.

*4.1. Diagnosis of Rolling Bearing Faults under Different Severity Levels.* The first study is to diagnose rolling bearing faults under different severity levels. The vibration signals of rolling bearings with different sizes of inner raceway defect are illustrated in Figure 6. The mutual information for each wavelet function is calculated and the results are listed in Table 2. Considering that the reverse Biorthogonal wavelet

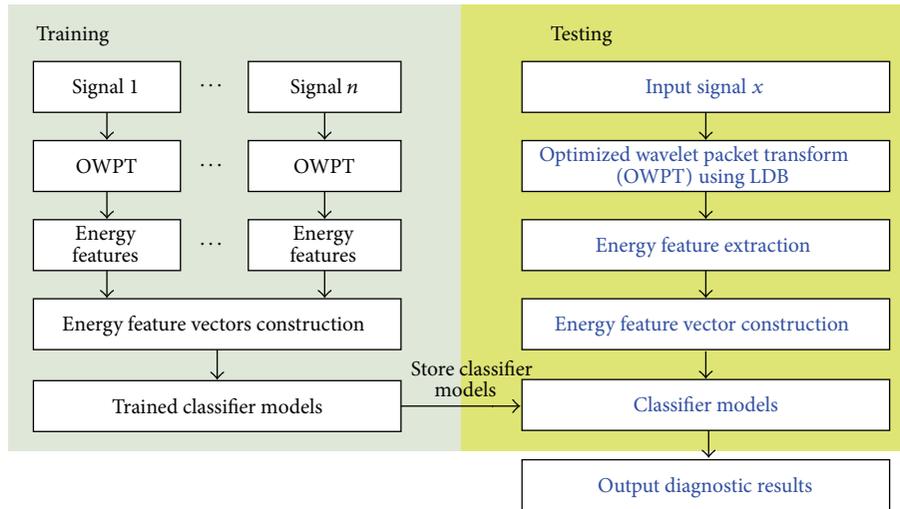


FIGURE 4: Flowchart of the rolling bearing fault diagnosis scheme.

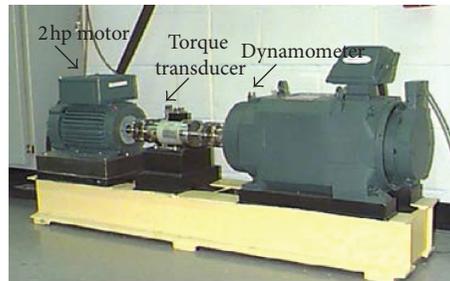


FIGURE 5: Bearing test system.

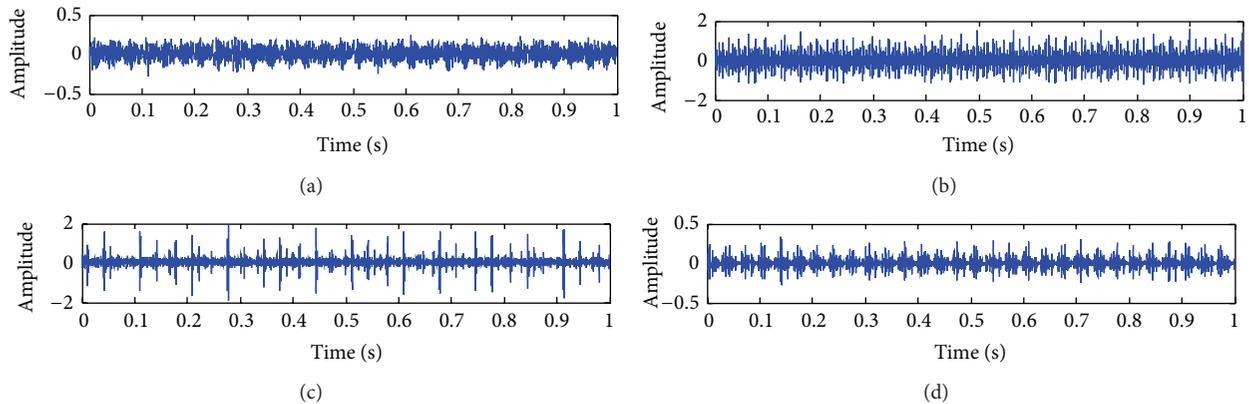


FIGURE 6: Vibration signals of rolling bearings under different defect sizes: (a) normal, (b) 0.18 mm hole, (c) 0.36 mm hole, and (d) 0.53 mm hole (speed: 1797 rpm, load: 0 hp).

1.3 (denoted as rBio1.3) possesses the highest value, thus it is identified as the most appropriate wavelet to analyze the rolling bearing signals through an optimized wavelet packet transform.

The energy values calculated from the optimal set of orthogonal wavelet subspaces (40 groups of training signals and 20 groups of testing signals, each containing 1,024 data points) are then chosen as inputs to a support vector machine (SVM) classifier for characterizing the rolling bearing defect

severity levels. The classification results are listed in Table 3. It can be seen that these features bring about much high classification accuracy score at 100%, which indicates that the developed fault diagnosis method is effective for classifying rolling bearing defect severity.

For the purpose of comparison, the energy features are also input to different classifiers, and Table 4 presents the classification results achieved by four different classifiers. It can be seen that, except for the HMM classifier, the other

TABLE 2: Mutual information of the extracted bearing vibration signal using different wavelets.

Wavelet function	Mutual information	Wavelet function	Mutual information	Wavelet function	Mutual information
Haar	17.007	Coif5	20.737	Bior2.6	20.801
Db2	20.781	Sym2	20.781	Bior4.4	20.762
Db4	20.740	Sym3	20.783	Bior5.5	20.765
Db6	20.771	Sym4	20.788	Bior6.8	20.773
Db8	20.772	Sym5	20.797	rBio1.3	<b>20.879</b>
Db10	20.762	Sym6	20.760	rBio2.4	20.741
Coif1	20.766	Sym7	20.783	rBio2.6	20.805
Coif2	20.795	Sym8	20.801	rBio4.4	20.760
Coif3	20.777	Bior1.3	16.953	rBio5.5	20.776
Coif4	20.764	Bior2.4	20.770	rBio6.8	20.789

TABLE 3: Classification results of rolling bearing faults under different severity levels.

Target classes	Sample size	Classification results				Classification rate [%]	Total [%]
		Normal	0.18 mm hole	0.36 mm hole	0.53 mm hole		
Normal	20	20	0	0	0	100	
0.18 mm hole	20	0	20	0	0	100	100
0.36 mm hole	20	0	0	20	0	100	
0.53 mm hole	20	0	0	0	20	100	

TABLE 4: Comparison for classification results using different classifiers.

Classifier types	Classification rate [%]
SVM classifier	100.00
Bayes classifier	100.00
HMM classifier	93.75
BP-NN classifier	100.00

three classifiers perform well for the selected features in characterizing the rolling bearing faults.

Furthermore, control experiments using three different wavelet functions are conducted. It can be seen that the test using the reverse Biorthogonal wavelet 1.3 which is the most suitable wavelet to analyze the rolling bearing signals acquires higher classification accuracy than those using other wavelet functions as shown in Table 5. This verifies the effectiveness of the mutual information measures for wavelet selection.

*4.2. Diagnosis of Rolling Bearing Faults with Different Locations.* The second study is to diagnose rolling bearing faults at different locations. Figure 7 illustrates the waveform of the vibration signals of rolling bearings with different fault locations under varying load and rotating speed.

Following the same procedure, the energy features extracted from the selected subspaces (40 groups of training signals and 20 groups of testing signals, each containing 1,024 data points) are chosen as inputs to the SVM classifier for distinguishing different faults of rolling bearing. The classification results for rolling bearing with different fault locations are shown in Table 6. It can be seen that the selected features produce the high accurate classification rate with 98.75%, where only one inner raceway fault is misclassified.

This indicates that the developed fault diagnosis approach is also effective for bearing fault location classification.

The effects of different classifiers on test results are also studied in the experiment to identify different fault locations. The results shown in Table 7 indicate that the SVM classifier is again good for rolling bearing fault location classification.

Similarly, control experiments using three different wavelets are conducted. In comparison, using the reverse Biorthogonal wavelet 1.3, which is the optimized wavelet function, to analyze the rolling bearing vibration signals has produced a higher classification rate than those using other two wavelet functions as shown in Table 8. This verifies the effectiveness of the quantitative wavelet selection criterion proposed in this paper.

## 5. Conclusions

This paper presents a quantitative wavelet selection approach for rolling bearing vibration signal analysis. Based on the quantitative wavelet selection and the optimized wavelet packet transform using LDB algorithm, a new effective approach for rolling bearing fault diagnosis has been developed. The results of the experiment studies indicate that the proposed approach has good ability to diagnose the rolling bearing faults. Furthermore, the comparison experiments also show that the proposed quantitative wavelet selection method improves the classification accuracy, which confirm the performance for the enhanced rolling bearing fault diagnosis approach.

## Conflict of Interests

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

TABLE 5: The classification results under different wavelet functions.

Wavelet function	Classification accuracy [%]				Classification rate [%]
	Normal	0.18 mm hole	0.36 mm hole	0.53 mm hole	
Bior5.5	100	95	85	100	95.00
Db10	100	90	85	100	93.75
rBio1.3	100	90	95	100	100.00

TABLE 6: Classification results of rolling bearing faults at different locations.

Target classes	Sample size	Classification results				Classification rate [%]	Total [%]
		Normal	Inner raceway fault	Ball fault	Outer raceway fault		
Normal	20	20	0	0	0	100	
Inner raceway fault	20	0	19	0	1	95	98.75
Ball fault	20	0	0	20	0	100	
Outer raceway fault	20	0	0	0	20	100	

TABLE 7: Comparison for classification results using different classifiers.

Classifier types	Classification rate [%]
SVM classifier	98.75
Bayes classifier	88.75
HMM classifier	86.25
BP-NN classifier	98.75

TABLE 8: The classification results under different wavelet functions.

Wavelet function	Classification accuracy [%]				Classification rate [%]
	Normal	Inner raceway fault	Ball fault	Outer raceway fault	
Bior5.5	100	80	100	100	95.00
Db10	100	65	95	100	90.00
rBio1.3	100	95	100	100	98.75

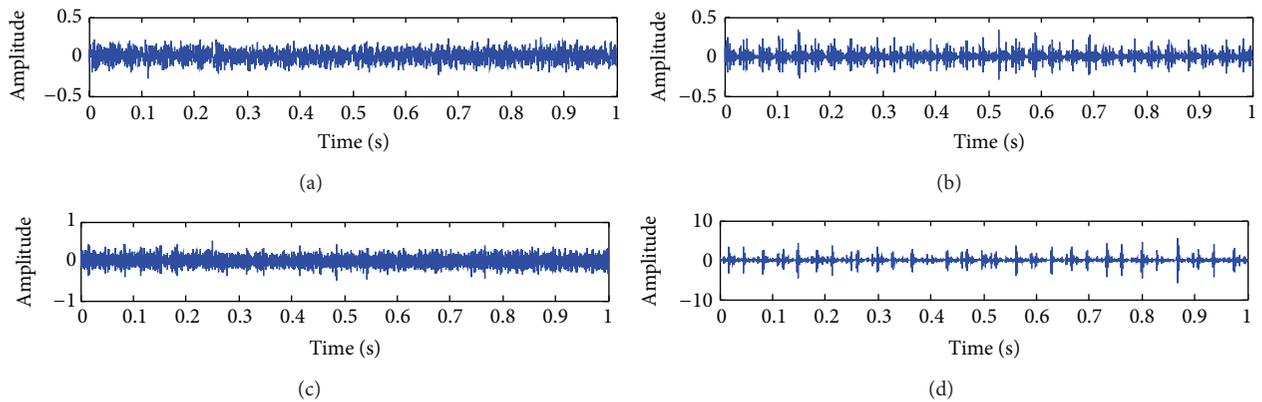


FIGURE 7: Vibration signals of bearings with different fault locations under varying load and rotating speed: (a) normal, (b) inner raceway fault, (c) ball fault, and (d) outer raceway fault (fault diameter: 0.53 mm).

## Acknowledgments

This work has been supported in part by the National Natural Science Foundation of China under Grant 51175080, the Aeronautical Science Foundation of China under Grant 20122269015, and the Science and Technology Support Program of Jiangsu Province under Grant BE2012740.

## References

- [1] B. Rao, *Handbook of Condition Monitoring*, Elsevier Advanced Technology, Oxford, UK, 1996.
- [2] Z. K. Peng and F. L. Chu, "Application of the wavelet transform in machine condition monitoring and fault diagnostics: a review with bibliography," *Mechanical Systems and Signal Processing*, vol. 18, no. 2, pp. 199–221, 2004.

- [3] R. Yan, R. X. Gao, and X. Chen, "Wavelets for fault diagnosis of rotary machines: a review with applications," *Signal Processing*, vol. 96, pp. 1–15, 2014.
- [4] C. Shen, D. Wang, F. Kong, and P. W. Tse, "Fault diagnosis of rotating machinery based on the statistical parameters of wavelet packet paving and a generic support vector regressive classifier," *Measurement*, vol. 46, no. 4, pp. 1551–1564, 2013.
- [5] Q. He, "Vibration signal classification by wavelet packet energy flow manifold learning," *Journal of Sound and Vibration*, vol. 332, no. 7, pp. 1881–1894, 2013.
- [6] P. Bošković and D. Juričić, "Fault detection of mechanical drives under variable operating conditions based on wavelet packet Rényi entropy signatures," *Mechanical Systems and Signal Processing*, vol. 31, pp. 369–381, 2012.
- [7] J. Li, S. Li, X. Chen, and L. Wang, "The hybrid KICA-GDA-LSSVM method research on rolling bearing fault feature extraction and classification," *Shock and Vibration*, vol. 2015, 9 pages, 2015.
- [8] A. Tabrizi, L. Garibaldi, A. Fasana, and S. Marchesiello, "Early damage detection of roller bearings using wavelet packet decomposition, ensemble empirical mode decomposition and support vector machine," *Meccanica*, vol. 50, no. 3, pp. 865–874, 2015.
- [9] J. Zhang, W. Ma, J. Lin, L. Ma, and X. Jia, "Fault diagnosis approach for rotating machinery based on dynamic model and computational intelligence," *Measurement*, vol. 59, pp. 73–87, 2015.
- [10] Y. Wang, G. Xu, L. Liang, and K. Jiang, "Detection of weak transient signals based on wavelet packet transform and manifold learning for rolling element bearing fault diagnosis," *Mechanical Systems and Signal Processing*, vol. 54–55, pp. 259–276, 2015.
- [11] S. Hong, Z. Zhou, E. Zio, and K. Hong, "Condition assessment for the performance degradation of bearing based on a combinatorial feature extraction method," *Digital Signal Processing*, vol. 27, no. 1, pp. 159–166, 2014.
- [12] J. Rafiee, P. W. Tse, A. Harifi, and M. H. Sadeghi, "A novel technique for selecting mother wavelet function using an intelligent fault diagnosis system," *Expert Systems with Applications*, vol. 36, no. 3, pp. 4862–4875, 2009.
- [13] J. Rafiee, M. A. Rafiee, and P. W. Tse, "Application of mother wavelet functions for automatic gear and bearing fault diagnosis," *Expert Systems with Applications*, vol. 37, no. 6, pp. 4568–4579, 2010.
- [14] D. Jiang and C. Liu, "Machine condition classification using deterioration feature extraction and anomaly determination," *IEEE Transactions on Reliability*, vol. 60, no. 1, pp. 41–48, 2011.
- [15] W.-X. Yang and X.-M. Ren, "Detecting impulses in mechanical signals by wavelets," *EURASIP Journal on Advances in Signal Processing*, vol. 8, pp. 1156–1162, 2004.
- [16] Z. X. Chen, "The optimal selection of base wavelet in fault diagnosis," *Journal of Mechanical Science and Technology*, vol. 24, no. 2, pp. 172–175, 2005.
- [17] E. L. Schukin, R. U. Zamaraev, and L. I. Schukin, "The optimisation of wavelet transform for the impulse analysis in vibration signals," *Mechanical Systems and Signal Processing*, vol. 18, no. 6, pp. 1315–1333, 2004.
- [18] F. Li, G. Meng, and L. Ye, "Optimal mother wavelet-based Lamb wave analyses and damage detection for composite structures," *Chinese Journal of Scientific Instrument*, vol. 28, no. 10, pp. 1729–1735, 2007.
- [19] F. Li, G. Meng, L. Ye, Y. Lu, and K. Kageyama, "Dispersion analysis of Lamb waves and damage detection for aluminum structures using ridge in the time-scale domain," *Measurement Science and Technology*, vol. 20, no. 9, Article ID 095704, pp. 695–704, 2009.
- [20] F. Li, G. Meng, K. Kageyama, Z. Su, and L. Ye, "Optimal mother wavelet selection for lamb wave analyses," *Journal of Intelligent Material Systems and Structures*, vol. 20, no. 10, pp. 1147–1161, 2009.
- [21] R. Yan and R. X. Gao, "Energy-based feature extraction for defect diagnosis in rotary machines," *IEEE Transactions on Instrumentation and Measurement*, vol. 58, no. 9, pp. 3130–3139, 2009.
- [22] R. Yan and R. X. Gao, "Base wavelet selection for bearing vibration signal analysis," *International Journal of Wavelets, Multiresolution and Information Processing*, vol. 7, no. 4, pp. 411–426, 2009.
- [23] P. K. Kankar, S. C. Sharma, and S. P. Harsha, "Fault diagnosis of ball bearings using continuous wavelet transform," *Applied Soft Computing Journal*, vol. 11, no. 2, pp. 2300–2312, 2011.
- [24] C. Wu, T. Chen, R. Jiang, L. Ning, and Z. Jiang, "A novel approach to wavelet selection and tree kernel construction for diagnosis of rolling element bearing fault," *Journal of Intelligent Manufacturing*, 2015.
- [25] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, San Diego, Calif, USA, 1999.
- [26] N. Saito and R. R. Coifman, "Local discriminant bases and their applications," *Journal of Mathematical Imaging and Vision*, vol. 5, no. 4, pp. 337–358, 1995.
- [27] K. Umamathy, S. Krishnan, and R. K. Rao, "Audio signal feature extraction and classification using local discriminant bases," *IEEE Transactions on Audio, Speech & Language Processing*, vol. 15, no. 4, pp. 1236–1246, 2007.
- [28] A. R. Harris, K. Schwerdtfeger, and D. J. Strauss, "Optimized shift-invariant wavelet packet feature extraction for electroencephalographic evoked responses," in *Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS '08)*, pp. 2685–2688, August 2008.
- [29] Q. He, R. Yan, and R. X. Gao, "Wavelet packet base selection for gearbox defect severity classification," in *Proceedings of the Prognostics and System Health Management Conference (PHM '10)*, 5, p. 1, IEEE, Macao, China, January 2010.
- [30] J. Cui, M. Shan, R. Yan, and Y. Wu, "Aero-engine fault diagnosis using improved local discriminant bases and support vector machine," *Mathematical Problems in Engineering*, vol. 2014, Article ID 283718, 9 pages, 2014.
- [31] <http://csegroups.case.edu/bearingdatacenter>.

