Research Article

Rolling Bearing Fault Diagnosis Based on ELCD Permutation Entropy and RVM

Jiang Xingmeng,[1] Wu Li,[2] Pan Liwu,[3] Ge Mingtao,[2] and Hu Daidi[2]

[1]*Zhengzhou Railway Vocational & Technical College, No. 9 Qiancheng Road, Zhengdong New District, Zhengzhou, Henan 451460, China*
[2]*College of Electronics and Information Engineering, SIAS International University, No. 168 Renmin Road, Xinzheng 451150, China*
[3]*Department of Automation & Control, Henan University of Animal Husbandry & Economy, Zhengzhou, Henan 451150, China*

Correspondence should be addressed to Pan Liwu; panliwu@163.com

Aiming at the nonstationary characteristic of a gear fault vibration signal, a recognition method based on permutation entropy of ensemble local characteristic-scale decomposition (ELCD) and relevance vector machine (RVM) is proposed. First, the vibration signal was decomposed by ELCD; then a series of intrinsic scale components (ISCs) were obtained. Second, according to the kurtosis of ISCs, principal ISCs were selected and then the permutation entropy of principal ISCs was calculated and they were combined into a feature vector. Finally, the feature vectors were input in RVM classifier to train and test and identify the type of rolling bearing faults. Experimental results show that this method can effectively diagnose four kinds of working condition, and the effect is better than local characteristic-scale decomposition (LCD) method.

1. Introduction

Rolling bearing is one of the most important elements of mechanical equipment. Its vibration signal gathered in case of default tends to be nonstationary [1–3]. Therefore, its fault diagnosis with appropriate signal processing is of significance.

As a new nonstationary signal processing method, local characteristic-scale decomposition (LCD) method can self-adaptively decompose a complex signal into several intrinsic scale components (ISCs). The decomposed ISC has its instantaneous frequency of physical significance. Any two ISCs are mutually independent [4, 5]. Compared with empirical mode decomposition (EMD), LCD is widely used in mechanical fault diagnosis due to shorter decomposition time, lower end effect, and less iterations. However, as with the EMD method, LCD method takes cubic spline interpolation for signal processing, so it causes a mode mixing effect, failing to obtain accurate ISCs [6, 7]. Characterized by even frequency distribution and zero-mean value, the method through which the white noise is added into LCD method to process

vibration signal, namely, ensemble local characteristic-scale decomposition (ELCD), can help eliminate mode mixing. Relevance vector machine (RVM) is a new classifier proposed on the basis support vector machine (SVM) and Bayesian theory. This classifier can create core functions without undue parameters. Characterized by shorter test time, high precision, and small sample classification availability, it is widely used in gear fault diagnosis, fault diagnosis of hydropower units, and medical diagnostics [8–10]. Therefore, aiming at coping with various working conditions, RVM classifier is adopted to realize the pattern recognition of rolling bearing faults.

Firstly, ELCD is used to decompose the vibration signal into ISCs; secondly, based on their kurtosis feature, the principal ISC components are selected; the permutation entropy of principal ISC is extracted. Thirdly, the permutation entropy obtained is taken as the inputs to RVM for fault diagnosis. Finally, the proposed method is applied to experimental data, and the analysis results show that the proposed method performs effectively for the rolling bearing fault diagnosis.

2. LCD and ELCD Principles

2.1. LCD Fundamentals. As a new self-adaptive signal decomposition method, LCD can decompose a nonstationary signal into several ISCs. Any two ISCs are mutually independent. With its instantaneous frequency of physical significance [5], ISC needs to meet the following two conditions [6].

(1) Signal $X(t)$ has its maxima as positive and minima as negative, and any adjacent maxima and minima witness monotonic relationship.

(2) Among the data, let all the maximal points be denoted as (τ_k, X_k). The line formed by any two adjacent extreme points, l_k, at τ_k as A_{k+1}, is specified as follows: $l_k = (X_{k+2} - X_k)(t - \tau_k)/(\tau_{k+2} - \tau_k) + X_k$. Then, the relation $A_{k+1} + X_{k+1} = 0$ should be true, where $A_{k+1} = (\tau_{k+1} - \tau_k)(X_{k+2} - X_k)/(\tau_{k+2} - \tau_k) + X_k$,

$$A_{k+1} + X_{k+1} = 0,$$
$$A_{k+1} = X_k + \frac{\tau_{k+1} - \tau_k}{\tau_{k+2} - \tau_k}(X_{k+2} - X_k), \qquad (1)$$
$$k = 1, 2, \ldots, M - 2.$$

Any complex signal $X(t)$ may have its LCD results written as follows:

$$X(t) = \sum_{p=1}^{n} \mathrm{ISC}_p(t) + r_n(t), \qquad (2)$$

where $r_n(t)$ is the residual component.

2.2. ELCD Principles. Like EMD, LCD takes cubic spline interpolation for signal processing and witnesses mode mixing [7]. Characterized by even frequency distribution, the white noise is used to eliminate mode mixing [11]. This study proposes ELCD for signal processing. ELCD repeatedly adds different white noises to the vibration signal for multiple cases of LCD. Based on white noise zero-mean value, this study conducted ensemble mean of multiple sets of ISCs and ultimately obtained a set of ISCs, eliminating mode mixing. ELCD specific algorithm flow is shown in Figure 1.

2.3. Algorithm Simulation and Analysis. To contrast the advantages and disadvantages of LCD and ELCD, this study counted up the impact signal and sine signal and formed simulation signal, shown in Figure 2.

This study conducted simulation signal LCD and ELCD. In LCD, this study added the noise signal amplitude 0.01 times the signal standard deviation and had the total mean as 120 times. The results are shown in Figures 3 and 4.

Simulation comparison showed that, after LCD, ISC1 and ISC2 had partials and ISC had mode mixing. After ELCD, it was possible to obtain the real signal component and properly decompose the simulation signal. Therefore, ELCD can effectively remove mode mixing.
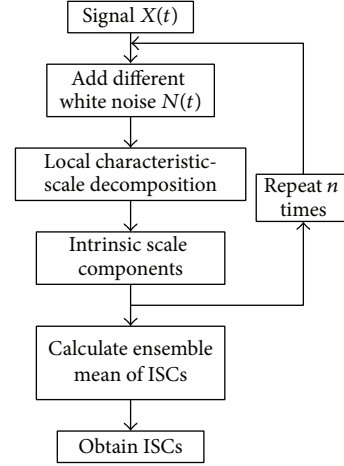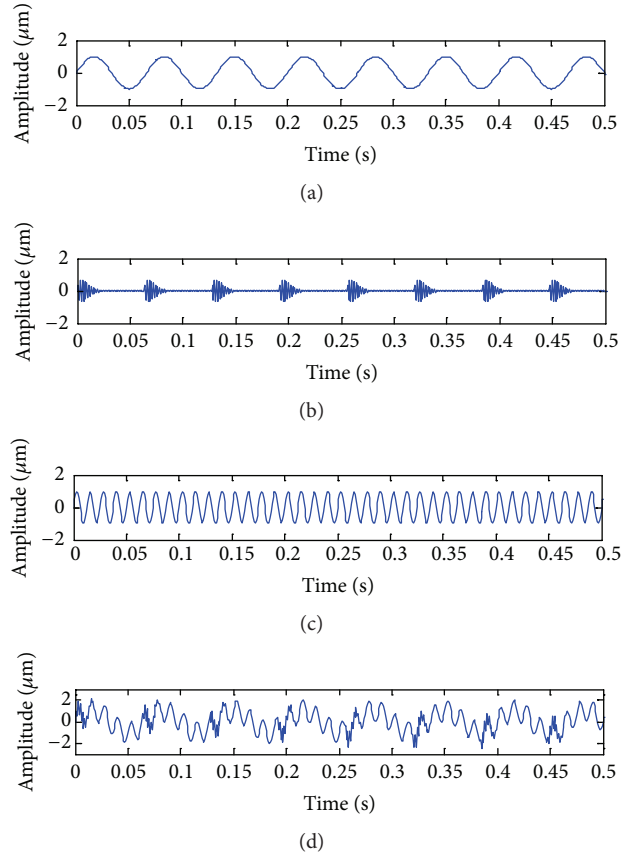

FIGURE 1: ELCD flow chart.

FIGURE 2: The simulation signal and its constituents.

3. Permutation Entropy

Permutation entropy is a nonlinear dynamics indicator to measure signal complexity. Compared with the Lyapunov exponent, permutation entropy is characterized by simple calculation and strong noise immunity [12]. This study took permutation entropy to present different vibration signal characteristics. The calculation of a time series $\{x(i), i = 1, 2, \ldots, N\}$ is as follows.
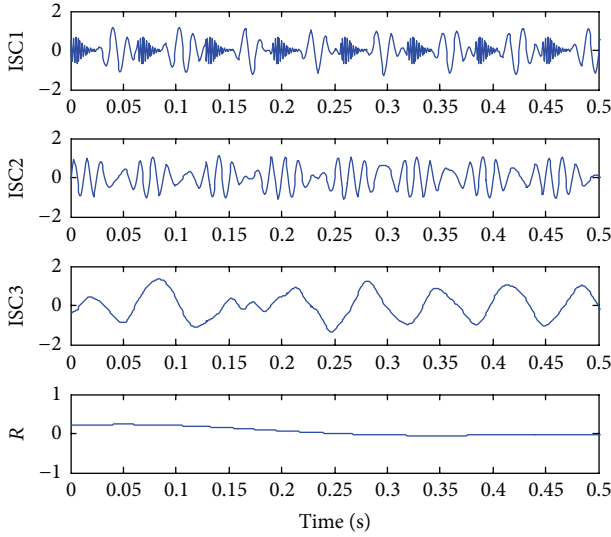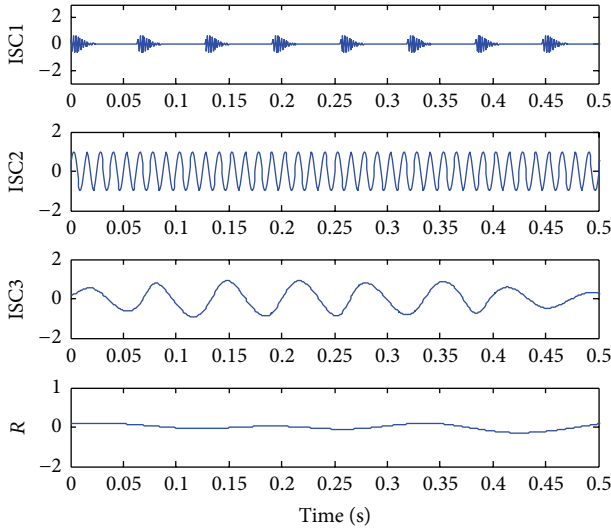
FIGURE 3: Simulation signal LCD results.

FIGURE 4: Simulation signal ELCD results.

(1) Time series phase-space reconstruction is as follows:

$$X(1) = \{x(1), x(1+\tau), \ldots, x(1+(m-1)\tau)\}$$

$$\vdots$$

$$X(i) = \{x(i), x(i+\tau), \ldots, x(i+(m-1)\tau)\}$$

$$\vdots \tag{3}$$

$$X(N-(m-1)\tau)$$

$$= \{x(N-(m-1)\tau), x(N-(m-2)\tau), \ldots, x(N)\},$$

where m is embedding dimension and τ is time delay.

(2) Sort the series. This study rearranged $X(i) = \{x(i), x(i+\tau), \ldots, x(i+(m-1)\tau)\}$ in ascending order and obtained a new series. If $x(i+(j_{i1}-1)\tau) = x(i+(j_{i2}-1)\tau)$ appeared in

$\{x(i+(j_1-1)\tau) \le x(i+(j_2-1)\tau) \le \cdots \le x(i+(j_m-1)\tau)\}$, this study would sort according to the value of j.

(3) It is obvious that, based on the sorting theory, any vector $X(i)$ can obtain a set of symbol series:

$$S(l) = [j_1, j_2, \ldots, j_m], \tag{4}$$

where $l = 1, 2, \ldots, k$, $k \le m!$; m different symbols $[j_1, j_2, \ldots, j_m]$ have $m!$ different arrangements, corresponding to $m!$ different symbols.

(4) To calculate the probability of different symbol series, $p_1, p_2, \ldots, p_k$, $\sum_{l=1}^{k} p_l = 1$; the time series permutation entropy is defined as follows:

$$H_p(m) = -\sum_{l=1}^{k} p_l \ln p_l. \tag{5}$$

In formula (5), in case $p_l = 1/m!$, the maximum of $H_p(m)$ is $\ln(m!)$ which is standardized with the mathematical expression as the following formula:

$$H_p = \frac{H_p(m)}{\ln(m!)}. \tag{6}$$

Entropy value range is $0 \le H_p \le 1$.

Entropy indicates the complexity of the time series. The more complex the time series is, the greater the entropy will be and vice versa.

4. Relevance Vector Machine

4.1. Fundamentals. RVM (relevance vector machine) is a learning machine based on Bayesian framework. Characterized by fewer parameters, it can conduct small sample data training and testing [8, 9]. With any training sample set $\{x_i, t_i\}_{i=1}^{N}$, RVM algorithm steps are as follows.

(1) Calculate the output function:

$$y(x) = \sum_{i=1}^{N} w_i \phi_i(x), \tag{7}$$

where w is weight vector and $\phi(x)$ is nonlinear basis function.

(2) In second classification, RVM uses logistic sigmoid function to map $y(x)$ $[0, 1]$ interval to obtain the decision probability distribution function:

$$P\{t = 1 \mid x\} = 1 + \frac{1}{e^{-y(x)}}. \tag{8}$$

In case of $P\{t = 1 \mid x\} < 0.5$, the results obtained are of Class I, and in case of $P\{t = 1 \mid x\} \ge 0.5$, the results obtained are of Class II.

(3) In case $P\{t = 1 \mid x\}$ complies with Bernoulli distribution and independent distribution, the likelihood function of the sample set is calculated:

$$P\{t \mid w\} = \prod_{i=1}^{N} p(t_i \mid x_i)^{t_i} (1 - p(t_i \mid x_i))^{1-t_i}. \tag{9}$$
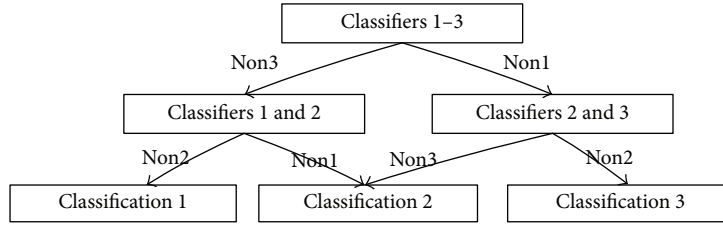
FIGURE 5: RVM classification graph.

(4) Maximize treatment according to the above formula and define the weight Gaussian prior probability distribution, thus reducing the sample size and improving learning and generalization:

$$P\{w \mid \alpha\} = \prod_{i=0}^{N} N\left(w_i \mid 0, \alpha_i^{-1}\right)$$
$$= \prod_{i=1}^{N} \sqrt{\frac{\alpha_i}{2\pi}} \exp\left(-\frac{\alpha_i w_i^2}{2}\right). \tag{10}$$

(5) RVM model can self-adaptively obtain its α_i through training. In order to ensure RVM model optimal performance, a fast PSO intelligent optimization algorithm is adopted to optimize the parameters. Fitness function is as the following formula:

$$\text{fitness} = \frac{y_t}{y} \times 100\%, \tag{11}$$

where y are total samples and y_t are correctly classified samples.

4.2. RVM Classification Principles. Against rolling bearing different working conditions, multiple classifiers are combined for classification in accordance with Directed Acyclic Graph (DAG). DAG classification is efficient and free from rejecting or misclassification areas. It starts from the root node, left or right classification based on the root node results [9, 10]. Assume that three classifiers classify three samples; DAG classification principles are shown in Figure 5.

5. Experimental Analysis

5.1. Experiments. This study adopted rolling bearing data of US Case Western Reserve University for processing. The experiment adopted 6205-2RS JEM SKF deep groove ball bearings, with rotating motor load power off 735.5 W, rolling bearing speed of 1797 r/min, adopting EDM technology to process the bearing into one with fault diameter of 0.3556 mm, and fault depth of 0.2794 mm. This study adopted the sensor sampling frequency of 12 KHZ, collecting four working state vibration signals, respectively, referring to the normal state, the rolling element fault, the inner ring fault, and outer ring fault, with each data sample's length N as 2500 points. The four signals gathered are shown in Figure 6.

The ISCs of the decomposition by ELCD of an Inner race fault signal are shown for illustration in Figure 7.
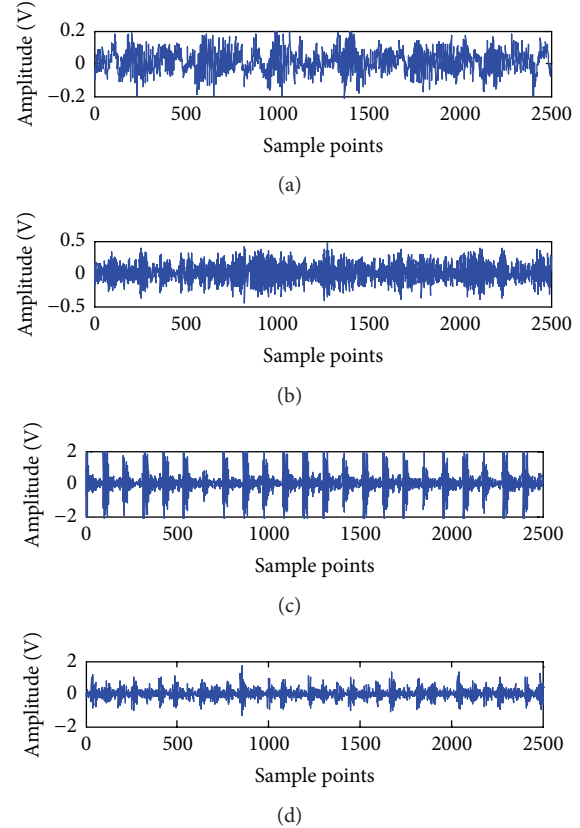

FIGURE 6: Vibration signal. (a) Normal. (b) Rolling element fault. (c) Outer race fault. (d) Inner race fault.

It can be obtained from Figure 7 that, after ELCD, the inner race fault signal obtained a residual component and 6 ISCs, some of which contain fault information and others contain fake components. Kurtosis can be used to measure signal impact. The more impact components the signal has, the greater the kurtosis will be [13]. So, this study took kurtosis for ISC processing and principal ISC extracting. ISC kurtosis after decomposition of different vibration signals was calculated, with the results shown in Figure 8.

Figure 8 shows that after ELCD different vibration signals had ISC kurtosis focusing on the first two components. The first two ISCs contained more fault information. Therefore, the first two ISCs were selected as the principal ISC, which were reconstructed. The mean square error and energy of the original signal and the reconstructed signal were, respectively, calculated, with the results shown in Table 1.

TABLE 1: The mean square error and energy of the original signal and the reconstructed signal.

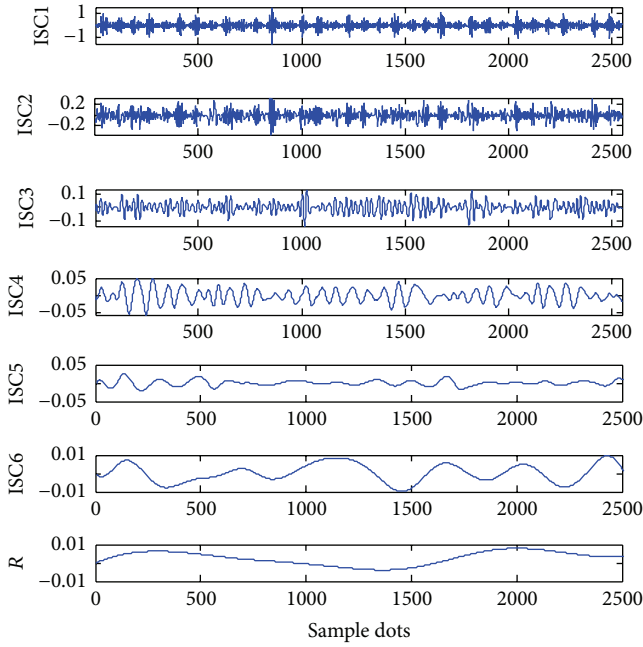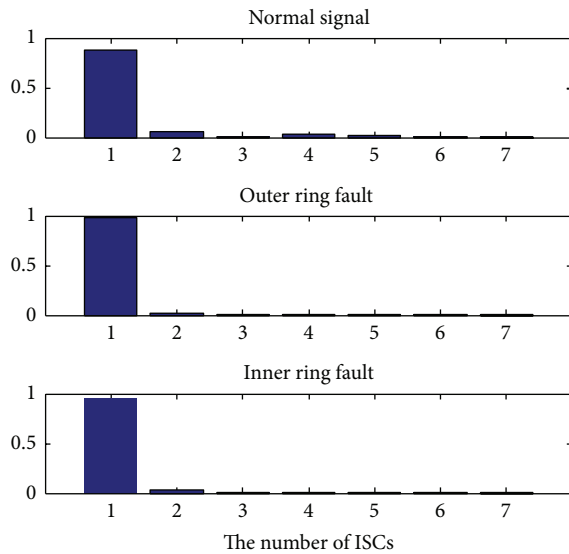| Signal type	Original signal		Reconstructed signal	
	Mean square error	Energy	Mean square error	Energy
Normal signal	0.0723	11.0282	0.0688	10.6203
Rolling element signal	0.1406	50.8750	0.1410	51.0342
Outer ring signal	0.6963	994.3297	0.6965	993.6019
Inner ring signal	0.2937	177.1054	0.2926	175.3812

FIGURE 7: Inner ring fault signal ELCD results.

FIGURE 8: ISC kurtosis after decomposition of different vibration signals.
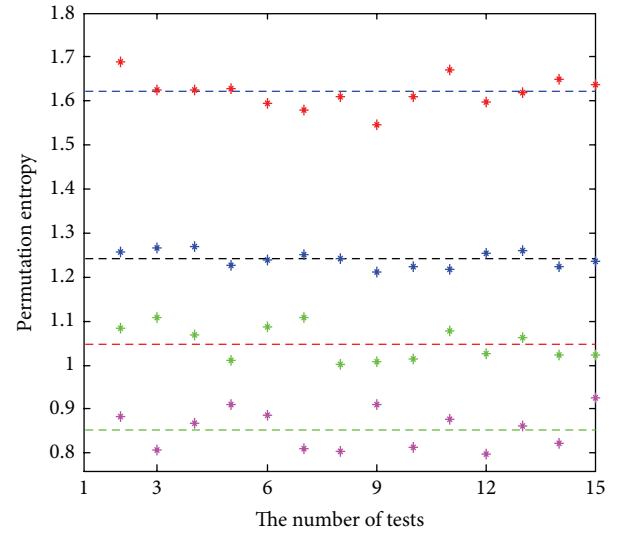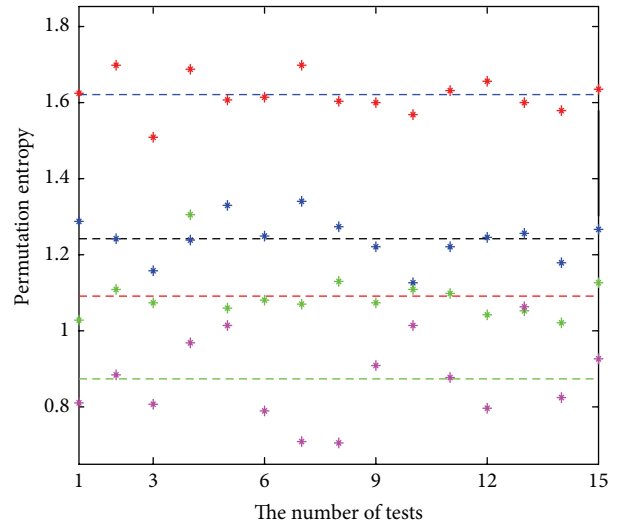
FIGURE 9: (a) Permutation entropy of LCD. (b) Permutation entropy of ELCD.
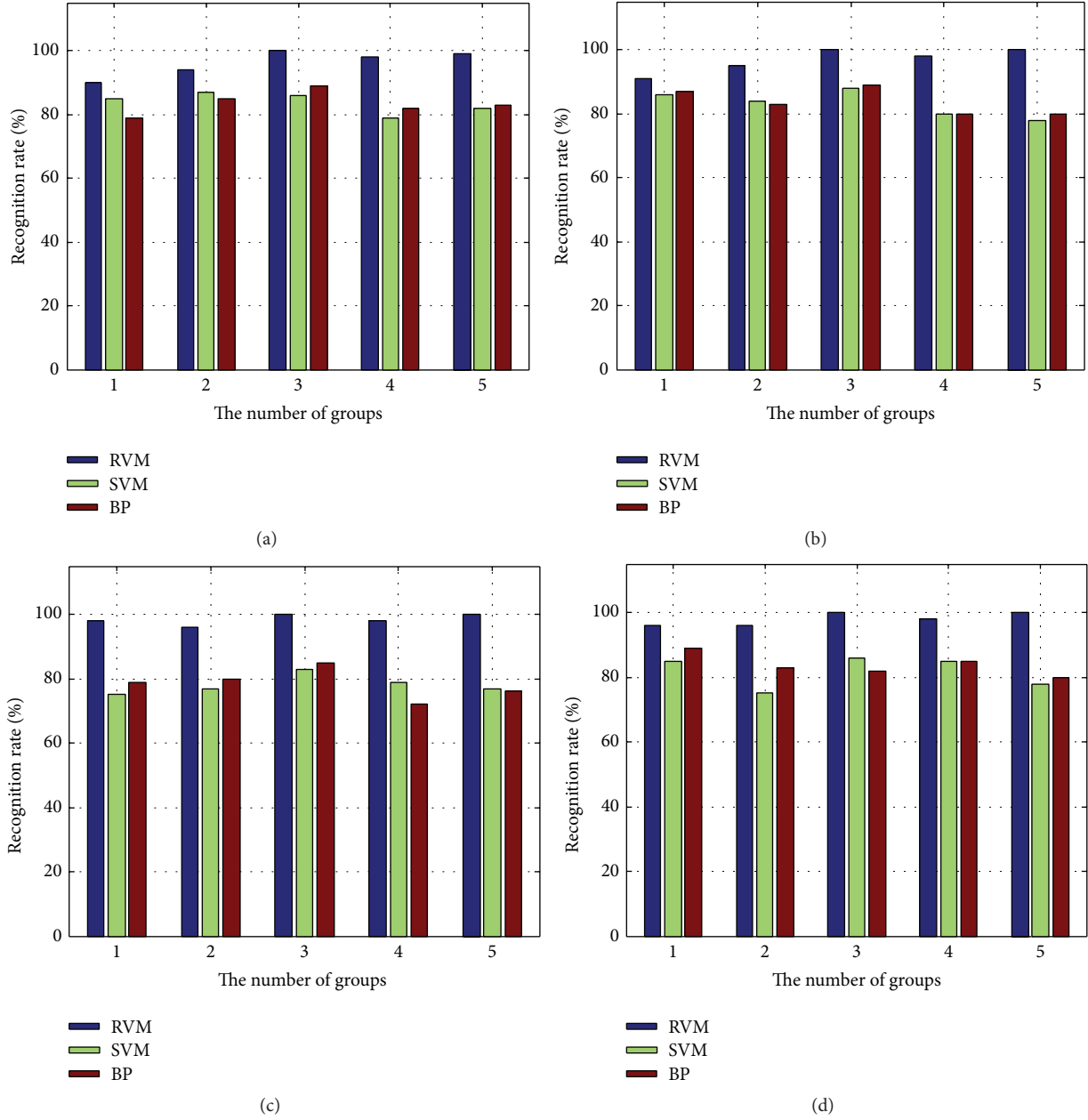
(a)

(b)

(c)

(d)

FIGURE 10: (a) Recognition rate of normal state. (b) Recognition rate of rolling element. (c) Recognition rate of inner ring fault. (d) Recognition rate of outer ring fault.

It can be obtained from Table 1 that characteristic statistics before and after the reconstruction are similar, indicating that kurtosis can effectively select the principal ISCs.

5.2. Calculation of Principal ISC Permutation Entropy. This study collected multiple sets of vibration signals, extracting principal ISCs against rolling bearing different working conditions and calculating permutation entropy, conducting multiple tests, and obtaining the permutation entropy. In case parameter $m = 6$ and parameter $t = 1$, entropy distinction was obvious. The mean of multiple sets of permutation entropy was calculated, with the results shown in Table 2.

TABLE 2: Permutation entropy of the first two ISCs of the three signals.

Principal ISC	Normal state	Rolling element fault	Outer ring fault	Inner ring fault
ISC1	0.8724	1.0125	1.2734	1.6346
ISC2	0.4356	0.6025	1.1804	1.3152

Table 2 shows that the rolling bearing different working conditions correspond to different permutation entropies. The normal working state signal is stable and less complex,

and the permutation entropy is the minimum and the signal gathered by the inner ring fault is far and noisy, and the permutation entropy is the maximum. In order to distinguish the advantages and disadvantages of ELCD and LCD, taking ISC1, for example, rolling bearing permutation entropies in different working conditions were calculated, with the results shown in Figure 9.

Figure 9 shows that ELCD is added with white noise of different amplitudes and compared with LCD; ISC1 is characterized by even permutation entropy distribution and obvious differences between different working conditions, overcoming mode mixing. Against different working conditions, in order to achieve accurate identification, it is necessary to adopt classifier for further processing. Therefore, RVM classifier is used to identify different working conditions.

5.3. Fault Recognition. As indicated in the above, permutation entropy of different vibration signals was calculated and feature vectors were composed. The feature vectors in different working conditions were input into RVM for training and testing. At the same time, BP and SVM classifiers were adopted to have feature training and testing. The recognition results of the three classifiers are shown in Figure 10.

Figure 10 shows that RVM classifier can effectively identify different states of rolling bearing, with a recognition rate higher than that of SVM classifier or BP classifier.

6. Summary

The rolling bearing is widely used in mechanical equipment; in order to keep the rolling bearings working in good state, condition monitoring and fault diagnosis are required. A rolling bearing fault diagnosis method based on permutation entropy of ELCD and RVM is proposed in this study.

(1) Targeting the disadvantage of mode mixing of LCD, the proposed ELCD in this study can have accurate ISC of the vibration signal, eliminating mode mixing.

(2) This study adopted RVM classifier to identify rolling bearing under different working conditions, with the results showing that this method can effectively identify rolling bearing fault, and the recognition rate is higher than that of SVM or BP.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

References

[1] Y. G. Lei and M. J. Zuo, "Fault diagnosis of rotating machinery using an improved HHT based on EEMD and sensitive IMFs," *Measurement Science and Technology*, vol. 20, no. 12, Article ID 125701, pp. 1–12, 2009.

[2] F. Zhang, Y. Liu, C. Chen, Y.-F. Li, and H.-Z. Huang, "Fault diagnosis of rotating machinery based on kernel density estimation and Kullback-Leibler divergence," *Journal of Mechanical Science and Technology*, vol. 28, no. 11, pp. 4441–4454, 2014.

[3] S. Wensheng, W. Fengtao, Z. Hong et al., "Feature extraction of rolling bearing fault based on wavelet packet sample entropy," *Journal of Vibration, Measurement & Diagnosis*, vol. 31, no. 2, pp. 162–166, 2011.

[4] J. Cheng, J. Zheng, and Y. Yang, "A non-stationary signal analysis approach-LCD," *Journal of Vibration Engineering*, vol. 25, no. 2, pp. 215–220, 2012.

[5] J.-D. Zheng, J.-S. Cheng, and Y. Yang, "A rolling bearing fault diagnosis method based on LCD and permutation entropy," *Journal of Vibration, Measurement and Diagnosis*, vol. 34, no. 5, pp. 802–806, 2014.

[6] C. Junsheng, Z. Jinde, and Y. Yu, "LCD-based Empirical envelope demodulation approach and its applications in mechanical fault diagnosis," *Journal of Mechanical Engineering*, vol. 48, no. 19, pp. 87–94, 2012.

[7] Y. Yang, M. Zeng, and J. Cheng, "Research on LCD and its capacities," *Journal of Vibration Engineering*, vol. 25, no. 5, pp. 602–609, 2012.

[8] C.-W. Wang and W.-H. You, "Boosting-SVM: effective learning with reduced data dimension," *Applied Intelligence*, vol. 39, no. 3, pp. 465–474, 2013.

[9] S. Qian, Z. Zhoumo, and L. Jian, "Application of RVM in pattern recognition of optical fiber pre-warning system," *Journal of Tianjin University (Science and Technology)*, vol. 47, no. 12, pp. 1115–1120, 2014.

[10] M. V. Shapovalov and R. L. Dunbrack Jr., "A smoothed backbone-dependent rotamer library for proteins derived from adaptive kernel density estimates and regressions," *Structure*, vol. 19, no. 6, pp. 844–858, 2011.

[11] Z. H. Wu and N. E. Huang, "A study of the characteristics of white noise using the empirical mode decomposition method," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 460, no. 2046, pp. 1597–1611, 2004.

[12] C. Bandt and B. Pompe, "Permutation entropy: a natural complexity measure for time series," *Physical Review Letters*, vol. 88, no. 17, Article ID 174102, 2002.

[13] J. Sun, Q. Xiao, J. Wen, and F. Wang, "Natural gas pipeline small leakage feature extraction and recognition based on LMD envelope spectrum entropy and SVM," *Measurement*, vol. 55, no. 9, pp. 434–443, 2014.