

Research Article

A Fault Diagnosis Method for Rotating Machinery Based on PCA and Morlet Kernel SVM

Shaojiang Dong,^{1,2} Dihua Sun,¹ Baoping Tang,³
Zhenyuan Gao,⁴ Wentao Yu,⁵ and Ming Xia⁶

¹ School of Mechatronics and Automotive Engineering, Chongqing Jiaotong University, Chongqing 400074, China

² School of Automation, Chongqing University, Chongqing 400044, China

³ The State Key Laboratory of Mechanical Transmission, Chongqing University, Chongqing 400030, China

⁴ Chongqing Academy of Metrology and Quality Inspection, Chongqing 401121, China

⁵ School of Mechanical and Electronic Engineering, Zhongyuan University of Technology, Zhengzhou 450007, China

⁶ Mechanical and Electrical Engineering Department, Chongqing Vocational Institute of Safety & Technology, Wanzhou, Chongqing 404020, China

Correspondence should be addressed to Shaojiang Dong; dongshaojiang100@163.com

Received 23 April 2014; Revised 8 June 2014; Accepted 15 June 2014; Published 8 July 2014

Academic Editor: Ruqiang Yan

Copyright © 2014 Shaojiang Dong 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.

A novel method to solve the rotating machinery fault diagnosis problem is proposed, which is based on principal components analysis (PCA) to extract the characteristic features and the Morlet kernel support vector machine (MSVM) to achieve the fault classification. Firstly, the gathered vibration signals were decomposed by the empirical mode decomposition (EMD) to obtain the corresponding intrinsic mode function (IMF). The EMD energy entropy that includes dominant fault information is defined as the characteristic features. However, the extracted features remained high-dimensional, and excessive redundant information still existed. So, the PCA is introduced to extract the characteristic features and reduce the dimension. The characteristic features are input into the MSVM to train and construct the running state identification model; the rotating machinery running state identification is realized. The running states of a bearing normal inner race and several inner races with different degree of fault were recognized; the results validate the effectiveness of the proposed algorithm.

1. Introduction

Rotating machinery is widely used in the modern factory. Unexpected mechanical faults could cause unscheduled downtime and loss. So, it is very important to diagnose the fault of the rotating machinery, to achieve effective fault diagnosis of the rotating machinery; firstly, the features should be extracted from the collected vibration data. Then, based on the extracted features an effective diagnosis model should be selected [1]. Feature extraction is the process of transforming the raw vibration data collected from running equipment to relevant information of health condition. There are three types of methods to deal with the raw vibration data: time domain analysis, frequency domain analysis, and time-frequency domain analysis. The three types of methods are often chosen to extract the feature. For example,

Yan et al. [2] introduce that the time-frequency domain transform method wavelet is often used to describe the characteristics of the vibration signals. Gebraeel et al. [3] chose the average of the amplitudes of the defective frequency and its first six harmonics over time as the features. Yan et al. [4] chose the short-time Fourier transform to extract the features. Ocak et al. [5] chose the wavelet packet transform to extract the feature of bearing wear information. Because the time domain analysis and the frequency features from FFT analysis results often tend to average out transient vibrations and thus do not provide a wholesome measure of the bearing health status, in this paper, the time-frequency EMD is used to decompose the vibration signal and the EMD Shannon entropy is used to extract the original features from the signal.

Although the original features can be extracted, they are still with high-dimensional and include superfluous

information. So, the original features fusion and dimensional reduction method should be used to deal with the original features so as to select the typical features. The most commonly used features fusion and dimensional reduction method is principal component analysis (PCA). Sun et al. [6] used PCA to extract features from the run-to-failure test of vibration signals of bearings. Dong and Luo [7] proposed a PCA-based multivariate analysis method for bearing degradation process prediction. In this paper, the PCA is used to achieve the extraction of the most sensitive features.

After selecting the typical features, another challenge is how to achieve effective fault diagnosis of the rotating machinery based on the extracted features. The existing machinery fault diagnosis methods can be roughly classified into model-based (or physics-based models) and data-driven methods. The model-based methods diagnosis the equipment fault using two models, the physical models based on the components and the damage propagation models based on damage mechanics. However, equipment dynamic response and damage propagation processes are typically very complex, and authentic physics-based models are very difficult to build [8]. Data-driven methods, known as artificial intelligent approaches, are derived directly from routine condition monitoring data of the monitored system, which achieve the fault diagnosis based on the learning or training process. The more prior the data used for the training process, the more accurate the model obtained. Artificial intelligent techniques have been increasingly applied to rotating machinery fault diagnosis recently. There have been some methods which are usually used for machinery fault diagnosis such as neural network and support vector machine (SVM) [9, 10]. However, the neural networks have the drawbacks of slow convergence; difficulty in escaping from local minima; and uncertain network structure, especially when doing the bearing fault diagnosis with large data. Those problems will be more troublesome. The SVM do not have those problems; however, the traditional SVM is not sensitive to the nonlinear feature classification, and, in recent years, the combination of wavelet theories and SVM has drawn considerable attention owing to its high classification ability for a wide range of applications and better performance than other traditional learning machines. In this paper, the Morlet kernel is used to construct the new SVM model, and the PSO method is used to select the parameters [11].

The paper is organized as follows. In Section 2, the concept of EMD energy entropy is proposed and the EMD energy entropies of different vibration signals are calculated; the PCA is used to achieve the extraction of the most sensitive features. In Section 3, the Morlet wavelet kernel SVM model is presented. In Section 4, the running state identification model for rotating machinery fault diagnosis is applied to roller bearing. The conclusion of this paper is given in Section 5.

The flowchart of the proposed method is shown in Figure 1.

2. Methods of Signal Processing for Feature Extraction

This section presents a brief discussion on feature extraction from EMD. EMD is developed to decompose a signal into IMF components and every IMF has a unique local frequency. The IMF should satisfy two conditions. (1) In the whole data set, the number of extreme and the number of zero crossings must be either equal or different at most by one and (2) at any point, the mean value of the upper envelope and lower envelope is zero [12].

Once the extreme is identified, the maxima are connected by using the cubic spline and used as the upper envelope. The minima are interpolated as well to form the lower envelope. The upper and the lower envelopes should cover all the data in the time series. The mean of the upper and the lower envelope, $m_1(t)$, is subtracted from the original signal to obtain the first component $h_1(t)$ of the sifting process:

$$h_1(t) = x(t) - m_1(t). \quad (1)$$

Ideally, if $h_1(t)$ is an intrinsic mode function, the sifting process will stop. So, it will shift the signal again in the same way to get another component $h_2(t)$:

$$h_2(t) = h_1(t) - m_2(t), \quad (2)$$

where $m_2(t)$ is the mean of the upper and lower envelopes of $h_1(t)$.

Repeat steps until the residue satisfies some stopping criterion. The signal can be expressed as

$$x(t) = \sum_{i=1}^n c_i(t) + r_n(t), \quad (3)$$

where n is the number of IMFs, $r_n(t)$ is the residue which is a constant, a monotonic, or a function with only maxima and one minimum from which no more IMF can be derived, and $c_i(t)$ denotes IMF.

Once the n IMFs and a residue $r_n(t)$ are obtained, where the energy of the n IMFs $E_1; E_2; \dots; E_n$ can be calculated, respectively, then, due to the orthogonality of the EMD decomposition, the sum of the energy of the n IMFs should be equal to the total energy of the original signal when the residue $r_n(t)$ is ignored. As the IMFs $c_1(t); c_2(t); \dots; c_n(t)$ include different frequency components, $E = \{E_1, E_2, \dots, E_n\}$ forms an energy distribution in the frequency domain of roller bearing vibration signal and then the corresponding EMD energy entropy is designated as

$$H_{\text{entropy}} = - \sum_{i=1}^n p_i \log p_i, \quad (4)$$

where $p_i = E_i/E$ is the percent of the energy of $c_i(t)$ in the whole signal energy ($E = \sum_{i=1}^n E_i$).

After the EMD energy entropy of the rotating machinery is calculated, the feature extraction method PCA is used to fuse the relevant useful features and extract the most sensitive features to work as the input of the proposed prediction model.

The procedure of feature extraction can be described as follows.

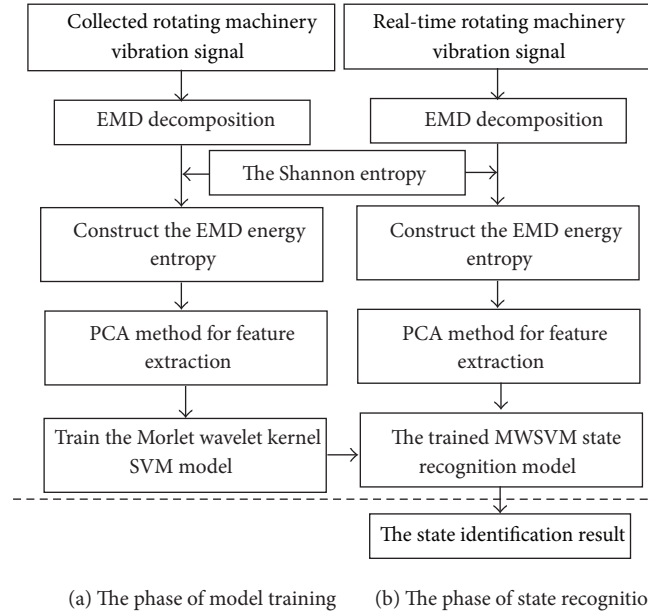


FIGURE 1: The flowchart of the proposed method.

- (1) Use the energy of the first five IMF components to get the features of the rotating machinery at each time.
- (2) Use the PCA to reduce the original features dimensions and get one set of typical features as follows.

(a) Compute the covariance matrix from the data as

$$C = (X - \bar{x})(X - \bar{x})^T, \quad (5)$$

where X is the data matrix of EMD IMFs, N is the total number of patterns, and \bar{x} represents mean vector of X .

(b) Compute the matrix of eigenvectors V and diagonal matrix of eigenvalues D as

$$V^{-1}CV = D. \quad (6)$$

(c) Sort the eigenvectors in V in descending order of eigenvalues in D and the data is projected on these eigenvector directions by taking the inner product in the data matrix sorted eigenvectors matrix as

$$\text{Projected data} = [V^T(X - \bar{x})]^T, \quad (7)$$

where V is of $n \times n$ dimension, and each row of it is an eigenvector. The features can be obtained.

- (3) Use the features as input of the MSVM for rotating machinery fault state identification.

3. The Morlet Wavelet Kernel SVM Model

The support vector's kernel function can be described as the horizontal floating function, such as $k(x, x') = k(\langle x \cdot x' \rangle)$.

In fact, if a function satisfies the condition of Mercer's theorem, it is the allowable support vector's kernel function. A specific Mercer's theorem description can be found in literature [13].

According to Mercer's theorem, the number of wavelet kernel functions which can be shown by the existent functions is few. Now, an existent wavelet kernel is given, the Morlet wavelet kernel. It can prove that this function can satisfy the condition of allowable support vector's kernel function. The Morlet wavelet function is defined as follows:

$$\psi(x) = \cos(w_0 x) e^{-x^2/2}. \quad (8)$$

The Morlet wavelet kernel function is defined as follows:

$$k(x, x') = k(x - x') = \prod_{i=1}^d \psi\left(\frac{x_i - x'_i}{a_i}\right) \quad (9)$$

$$= \prod_{i=1}^d \cos\left[w_0 \left(\frac{x_i - x'_i}{a_i}\right)\right] e^{[-(x_i - x'_i)^2 / 2a_i^2]}.$$

Then, the Morlet wavelet kernel function is being used as the support vector's kernel function, and the SVM is defined as

$$f(x) = \text{sgn} \left\{ \sum_{i=1}^n a_i y_i \prod_{j=1}^d \cos\left[w_0 \left(\frac{x_j - x'_j}{a_j}\right)\right] \times e^{[-(x_j - x'_j)^2 / 2a_j^2]} + b \right\}. \quad (10)$$

Through (9) and (10), the Morlet wavelet kernel SVM is constructed, and the new constructed SVM which is effective in classification is used to achieve bearing running state recognition.

3.1. The Morlet Kernel SVM Parameters Selection. The particle swarm optimization algorithm (PSO) is used to select the SVM parameters, and the PSO was first proposed in 1995. It is an optimization method based on a set of particles whose coordinates are potential solutions in the search space. Particles in PSO will change their coordinates (their solutions) by migration. During migration, each particle adjusts its own coordinates based on its own past experience and other particles' past experiences.

The PSO was chosen to optimize the Morlet kernel SVM parameters through the following formula:

$$v_{ij}(t+1) = wv_{ij}(t) + c_1r_{1j}(p_{ij}(t) - x_{ij}(t)) + c_2r_{2j}(p_{gj}(t) - x_{ij}(t)), \quad (11)$$

$$x_{ij}(t+1) = x_{ij}(t) + v_{ij}(t+1), \quad (12)$$

where the subscript “ i ” represents the i th particle. “ j ” represents the j -dimensional.

The subscript “ t ” represents the t generation. $v_{ij}(t)$ is the velocity of the i th particle in the t th iteration; $x_{ij}(t)$ is the position of the i th particle; $p_{ij}(t)$ is the pbest position of the i th particle; p_{gj} is the gbest position (pbest represents the local optimum of the particles; gbest represents the overall situation optimum of the particles). The w represents the inertia weight. c_1, c_2 are learning factors. $r_1 \sim U(0, 1)$, $r_2 \sim U(0, 1)$ represent two independent random functions.

The process of optimizing the parameters γ, σ based on the PSO is given as follows.

- (1) At the beginning of the optimization process, randomly initialize population sizes, c_1, c_2, w , rand(1), and rand(2), determine the termination condition, positions, and velocities of the particle, mapping the Morlet kernel SVM parameters w_0, a into a group of particles, and initialize the initial position of each particle, pbest, gbest.
- (2) When training the Morlet kernel SVM, use (11) as the PSO fitness function.
- (3) Use the target parameters w_0, a as the particles, use their initial values as the LS-SVM parameters in step (2), and use the corresponding value of (11) as the optimal solution of the w_0, a .
- (4) Use the initial error value of step (2) as the particle's initial fitness value and search the optimal value as the global fitness value among the initial fitness value and the corresponding particles as the current global optimal solution.
- (5) Update the velocity and position vector.
- (6) Resubstitute the updated parameters w_0, a into the Morlet kernel SVM model, retraining the Morlet kernel SVM model according to the step (2), save the output value, and calculate the fitness value of the particles again.
- (7) Compare the saved global fitness value gotten in step (6) with the current particle's fitness value, and if

the global fitness value is superior to the current particle's fitness value, update the current particle's fitness value according to step (5) and update the current particle's optimal value equal to the corresponding particle's optimal value gotten in step (6).

- (8) While the termination conditions are not met, return to step (5).
- (9) End the loop.

4. Validation

4.1. Case 1. In order to verify the effectiveness of the proposed method, bearing running state data sets of the normal state and several fault states were analyzed. The proposed method was applied to bearing fault signals obtained from the Case Western Reserve University [14]. The bearing type in the experiments is SKF 6205-2RS JEM. Experiments were conducted by using a 2 hp reliance electric motor. Bearings were seeded with faults by using electrodischarge machining. The test is to simulate the bearing normal running state and fault running states, with fault depths of 0.18 mm, 0.36 mm, 0.53 mm, and 0.71 mm at the inner raceway, outer raceway, and the ball to reflect the deteriorating state of the bearing; the inner-raceway fault signals were chosen in this case. Data was collected at the rate of 12,000 samples per second. 4096 data points were selected to analyze. 50 groups of test data of each fault state were selected, with 20 groups for training and the other 30 groups for testing. The collected vibration signals of normal state and inner-race four different fault depths are shown in Figure 2.

Next, the EMD decomposition was used to decompose each group of signals into IMFs, and Shannon entropy was used to extract the features. A group of inner-race entropy of Figure 2 is obtained, as shown in Table 1 (not normalized before).

Then, normalize the 20 groups of entropy values and input them into the PCA to reduce the dimension. In order to compare the dimension reduction and redundant treatment effect of PCA, the manifold learning method, local tangent space alignment (LTSA) [15], and the locality preserving projections (LPP) [16] method are used to reduce the dimension. The results are shown in Figures 3, 4, and 5. To be comparable, the dimensions of PCA, LTSA, and LPP are set to 3, so the input dimension of MWSVM is 3 and the neighborhood number is set to 10.

By comparing Figures 3, 4, and 5, the results show that the LTSA-based data dimension reduction method can not effectively separate the high-dimensional features, and there is still serious aliasing, which will affect the accuracy of the SVM state recognition effect. The LPP-based data dimension reduction method works better than the LTSA methods; however, there still have some features mixed together. The PCA-based data dimension reduction method can effectively separate the features of different running states with high calculation accuracy and a higher computational efficiency than the LPP and LTSA methods, which conform more to the actual project requirement. Thus, in the study, the PCA method is selected.

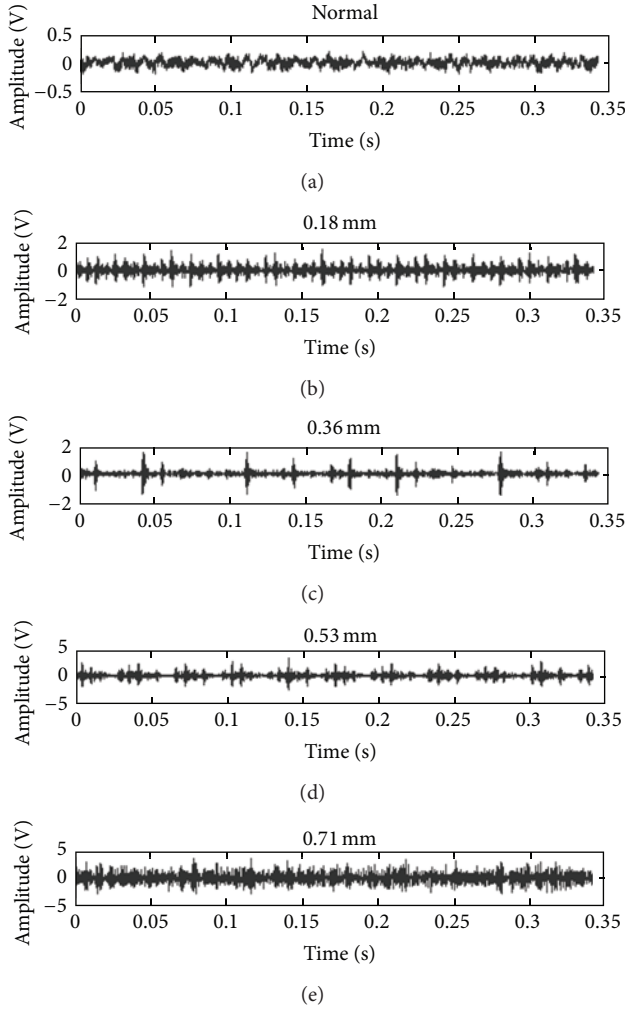


FIGURE 2: The collected vibration signals of normal state and inner-race four different fault depths.

After dimension reduction with the PCA, the extracted features are input into the SVM to train the model so as to recognize the states. And the PSO is used to obtain the main parameters of the model, the particle swarm population size is set to 100, and the number of the particles is set to 20. The fitness function is set to get the minimum prediction error with the optimized parameters. The prediction error is set to 0.0001. The PSO particle's dimension is set to 2, the w is set to 0.5, the c_1 is set to 1, and the c_2 is set to 1. The optimized obtained parameter w_0 is 5, and a is 0.3. Then, the two parameters are used to build the Morlet kernel SVM model to train and predict the value. In order to compare the identifying effect with and without manifold learning method, the following comparisons are done.

- (1) Use EMD Shannon entropy to extract the features and directly input the extract features into the MWSVM, without the PCA dimension reduction process.
- (2) Use EMD Shannon entropy to extract the features and process the extracted features by LTSA to reduce

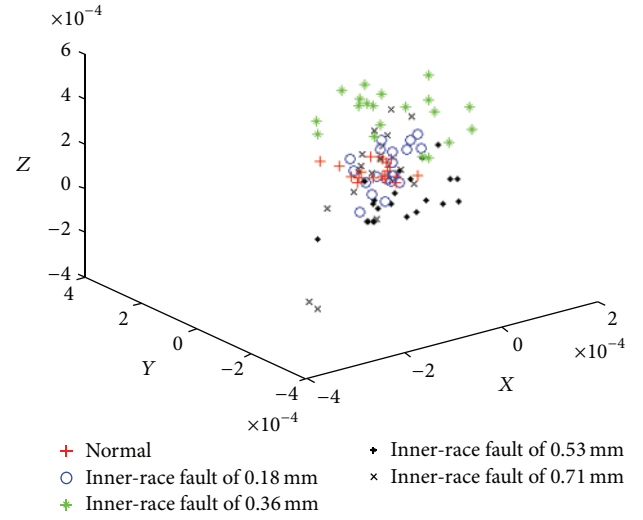


FIGURE 3: The dimension reduction and redundant treatment effect of LTSA.

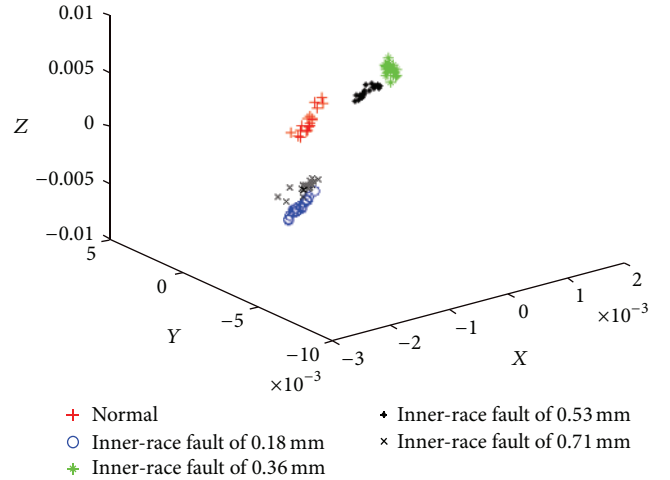


FIGURE 4: The dimension reduction and redundant treatment effect of LPP.

the dimension and then input the features into the MWSVM.

- (3) Use EMD Shannon entropy to extract the features and process the extracted features by LPP to reduce the dimension and then input the features into the MWSVM.
- (4) Use the method proposed in this paper.

The comparison results are shown in Table 2.

Table 2 shows that, after the PCA-based dimension reduction method and features extraction, the accuracy of states recognition improved significantly, much higher than the other algorithms. Therefore, the use of PCA for dimension reduction in this research is necessary and valuable.

In order to further verify the identification accuracy of the proposed method, the features extracted by PCA are input into the neural network, traditional RBF SVM (with penalty factor C set to 100 and nuclear parameter γ set to 0.1), the

TABLE 1: A group of inner-race EMD energy entropy of different running states.

Running states	H_1	H_2	H_3	H_4	H_5
Normal state	1.3989	1.3798	1.3876	1.3946	1.3866
0.18 mm fault depth	1.1342	1.1153	1.1016	1.1276	1.1213
0.36 mm fault depth	0.8766	0.8547	0.8987	0.8895	0.8673
0.53 mm fault depth	0.6449	0.6451	0.6689	0.6783	0.6884
0.71 mm fault depth	0.49814	0.5121	0.5565	0.5223	0.5127

TABLE 2: The states recognition rate of three different methods (recognition rate $\eta\%$).

States recognition methods	Normal state	0.18 mm fault depth	0.36 mm fault depth	0.53 mm fault depth	0.71 mm fault depth
Without PCA dimension reduction	91	61	76	83	81
Use the LTSA method dimension reduction	90	83	90	84	88
Use the LPP method dimension reduction	96	92	98	98	100
Use the PCA method dimension reduction	100	100	100	100	100

TABLE 3: The recognition rate of traditional RBF SVM and the MWSVM.

SVM type	Recognition rate $\eta\%$				
	Normal state	0.18 mm fault depth	0.36 mm fault depth	0.53 mm fault depth	0.71 mm fault depth
Neural network	90	96	94	93	90
RBF kernel	94	93	91	96	91
Symlet wavelet kernel	91	95	94	97	99
db wavelet kernel	90	98	96	92	90
Gaussian kernel	98	99	97	98	98
MWSVM	100	100	100	100	100

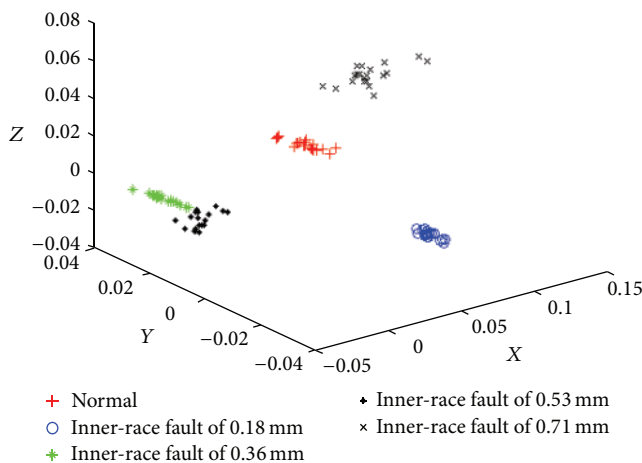


FIGURE 5: The dimension reduction and redundant treatment effect of PCA.

Symlet wavelet kernel SVM (with penalty factor C set to 100 and nuclear parameter γ set to 0.1), the db wavelet kernel SVM (with penalty factor C set to 100 and nuclear parameter γ set to 0.1), the Gaussian kernel SVM (with $a = 23.7$, penalty

factor C set to 100, and nuclear parameter γ set to 0.1), and the MWSVM (with w_0 set to 5 and a set to 0.3). The comparison results are shown in Table 3.

Table 3 shows that the MWSVM can better identify and approach the sensitive features because of Morlet wavelet kernel. Thus, the choice of MWSVM to determine the bearing running states can effectively improve recognition accuracy.

Next, a comparison about the training and test time loss of different methods is implemented.

- (1) The vibration data processed by EMD Shannon entropy and the extract features are directly input into the MWSVM, without the PCA dimension reduction.
- (2) The vibration data processed by EMD Shannon entropy and the features are processed by PCA to reduce the dimension. Then, the extracted features are input into the RBF kernel SVM.
- (3) The proposed method is in this research.

The comparison results are shown in Table 4.

In Table 4, after the dimension reduction, the recognition speed of SVM improved significantly. The time loss of the proposed method is the shortest. The reason is that the

TABLE 4: The time loss of three different methods.

Methods	Without PCA dimension reduction	By RBF kernel SVM	The proposed method
Time/s	133	79	58

TABLE 5: A group of EMD energy entropy of different running states of the actual signal.

Running states	H_1	H_2	H_3	H_4	H_5
Normal state	1.2657	1.2316	1.1091	1.1012	1.1316
Running for 2 months	1.0684	1.0201	1.2452	1.2566	1.3910
Running for 4 months	0.9506	0.9896	0.9597	1.1166	0.9259
Running for 6 months	0.8923	0.8369	0.8949	0.8283	0.8675
Running for 8 months	0.6031	0.5108	0.6392	0.6967	0.7042
Running for 10 months	0.7663	0.7095	0.7953	0.7325	0.8034
Running for 12 months	0.7509	0.7993	0.8331	0.7408	0.8296

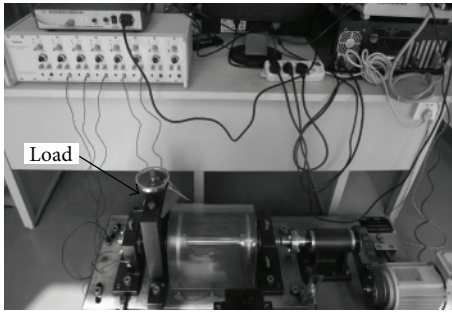


FIGURE 6: The test rig.

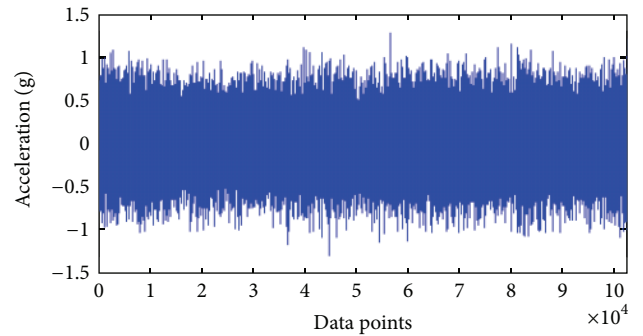


FIGURE 7: The collected vibration signal.

Morlet kernel is more sensitive to features classification and identification than the RBF kernel SVM. The result validates the proposed method and can effectively recognize the bearing running state.

4.2. Case 2. After validating the efficacy of the proposed method, the method is used on the actual application. The test rig is shown in Figure 6.

The bearings are hosted on the shaft; the shaft is driven by AC motor, the power is 0.55 KW, and the rotation speed is kept at 1000 rpm, with speed control and AC inverter controller. The brake maximum torque is 5 N·m, with a radial booster, using the magnetic clutch and brake. The rolling bearing is used, and a radial load of 29.4 N is added to the bearing. The data sampling rate is 25600 Hz and the data length is 102400 collected points, as shown in Figure 7. Every 2 hours, the vibration data is collected once. The bearing is run for one year. Then, a set of data from each of the 2 months is selected; the data sets are used to test whether or not the proposed method can identify the bearing running state. 4096 data points are selected to analyze, and 60 groups of collected data of different faults are obtained, with 30 groups for training and the other 30 groups for testing.

Next, each group of signals is decomposed by the EMD method, and the Shannon entropy is calculated. A group of features of different fault conditions are obtained, as shown in Table 5 (not normalized beforehand).

TABLE 6: The states recognition rate of different states based on the proposed method (recognition rate $\eta\%$).

Running states	Recognition rate $\eta\%$
Normal state	100
Running for 2 months	97
Running for 4 months	95
Running for 6 months	96
Running for 8 months	94
Running for 10 months	95
Running for 12 months	95

Then, the 30 groups' entropy values are normalized and input into the PCA in order to reduce the dimension and extract the typical features; the extracted features are input into the Morlet kernel SVM. The recognized results are shown in Table 6.

Table 6 shows that, although the actual bearing running state is very complex, the proposed method yields a high recognized accuracy. The results confirm that the proposed method can recognize the bearing running states effectively.

5. Conclusion

Firstly, this research used the EMD Shannon entropy method to extract the original features from the rotating machinery

vibration signals. The PCA was used to reduce the dimension and data redundancy of the entropy features. Through those methods, the typical features could be extracted effectively.

Then, in order to more accurately identify the bearing running state, the Morlet kernel was applied to construct the SVM recognition model; this can effectively improve the recognition accuracy of SVM.

Thirdly, through different comparisons, we can see that the proposed method makes good use of the advantage of all parts together to obtain better recognition accuracy and efficiency.

Finally, through the simulated signals and the tested signals in the research, the results show the significant efficacy of the proposed method in identifying the rotating machinery fault state.

Conflict of Interests

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

Acknowledgments

This research is supported by the Natural Science Foundation Project of CQ cstc2013jcyjA70896, Fundamental Research Funds for the Central Universities (Project no. CDJZR10118801). China Postdoctoral Science Foundation funded this research, Project no. 2014M552316. The authors are grateful to the anonymous reviewers for their helpful comments and constructive suggestions.

References

- [1] S. Dong, B. Tang, and Y. Zhang, "A repeated single-channel mechanical signal blind separation method based on morphological filtering and singular value decomposition," *Measurement*, vol. 45, no. 8, pp. 2052–2063, 2012.
- [2] 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, 2013.
- [3] N. Gebraeel, M. Lawley, R. Liu, and V. Parmeshwaran, "Residual life predictions from vibration-based degradation signals: a neural network approach," *IEEE Transactions on Industrial Electronics*, vol. 51, no. 3, pp. 694–700, 2004.
- [4] J. Yan, C. Guo, and X. Wang, "A dynamic multi-scale Markov model based methodology for remaining life prediction," *Mechanical Systems and Signal Processing*, vol. 25, no. 4, pp. 1364–1376, 2011.
- [5] H. Ocak, K. A. Loparo, and F. M. Discenzo, "Online tracking of bearing wear using wavelet packet decomposition and probabilistic modeling: a method for bearing prognostics," *Journal of Sound and Vibration*, vol. 302, no. 4-5, pp. 951–961, 2007.
- [6] C. Sun, Z. S. Zhang, and Z. J. He, "Research on bearing life prediction based on support vector machine and its application," *Journal of Physics*, vol. 305, Article ID 012028, 2011.
- [7] S. Dong and T. Luo, "Bearing degradation process prediction based on the PCA and optimized LS-SVM model," *Measurement*, vol. 46, no. 9, pp. 3143–3152, 2013.
- [8] Z. G. Tian, L. N. Wong, and N. M. Safaei, "A neural network approach for remaining useful life prediction utilizing both failure and suspension histories," *Mechanical Systems and Signal Processing*, vol. 24, no. 5, pp. 1542–1555, 2010.
- [9] J. Lee, J. Ni, D. Djurdjanovic, H. Qiu, and H. Liao, "Intelligent prognostics tools and e-maintenance," *Computers in Industry*, vol. 57, no. 6, pp. 476–489, 2006.
- [10] S. Dong, B. Tang, and R. Chen, "Bearing running state recognition based on non-extensive wavelet feature scale entropy and support vector machine," *Measurement*, vol. 46, no. 10, pp. 4189–4199, 2013.
- [11] K. C. Gryllias and I. A. Antoniadis, "A Support Vector Machine approach based on physical model training for rolling element bearing fault detection in industrial environments," *Engineering Applications of Artificial Intelligence*, vol. 25, no. 2, pp. 326–344, 2012.
- [12] Y. Gan, S. Lifen, and W. Jiangfei, "An EMD threshold de-noising method for inertial sensors," *Measurement*, vol. 49, pp. 34–41, 2014.
- [13] N. Cristianini and J. S. Taylor, *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*, Cambridge University Press, New York, NY, USA, 2000.
- [14] Case Western Reserve University Bearing Data Center, <http://csegroups.case.edu/bearingdatacenter/pages/welcome-case-western-reserve-university-bearing-data-center-website>.
- [15] Y. Zhan and J. Yin, "Robust local tangent space alignment via iterative weighted PCA," *Neurocomputing*, vol. 74, no. 11, pp. 1985–1993, 2011.
- [16] F. Dornaika and A. Assoum, "Enhanced and parameterless Locality Preserving Projections for face recognition," *Neurocomputing*, vol. 99, pp. 448–457, 2013.

