

Research Article

Rolling Bearing Degradation State Identification Based on LPP Optimized by GA

He Yu, Hong-ru Li, Zai-ke Tian, and Wei-guo Wang

Mechanical Engineering College, No. 97, Heping West Road, Shijiazhuang 050003, China

Correspondence should be addressed to He Yu; 13832329446@163.com

Received 5 May 2016; Accepted 12 July 2016

Academic Editor: Hyeon Joon Ahn

Copyright © 2016 He Yu 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.

In view of the problem that the actual degradation status of rolling bearing has a poor distinguishing characteristic and strong fuzziness, a rolling bearing degradation state identification method based on multidomain feature fusion and dimension reduction of manifold learning combined with GG clustering is proposed. Firstly, the rolling bearing all-life data is preprocessed by local characteristic-scale decomposition (LCD) and six typical features including relative energy spectrum entropy (LREE), relative singular spectrum entropy (LRSE), two-element multiscale entropy (TMSE), standard deviation (STD), RMS, and root-square amplitude (XR) are extracted and compose the original multidomain feature set. And then, locally preserving projection (LPP) is utilized to reduce dimension of original fusion feature set and genetic algorithm is applied to optimize the process of feature fusion. Finally, fuzzy recognition of rolling bearing degradation state is carried out by GG clustering and the principle of maximum membership degree and excellent performance of the proposed method is validated by comparing the recognition accuracy of LPP and GA-LPP.

1. Introduction

Rolling element bearings are one of the most important components for carrying heavy loads and providing constant rotational speed in rotating machines [1]. With continuous operation of rotating machines for a long time, rolling bearings' performance condition is changing all the time and that affects performance stability of the whole machine directly. Therefore, there is a practical significance for improving rotating machines' service life by rolling bearing degradation state identification in real time.

The fault feature extracted from vibration signals is analyzed to determine the bearing state [2]. And fault feature extraction is the basis of realizing rolling bearing degradation state recognition. Scientific degradation features can characterize the degradation degree of rolling elements accurately and stably. Degradation features are mainly selected from time domain, frequency domain, time-frequency analysis, and signal complexity aspects. Considering that the actual vibration signal of rolling bearings is nonlinear and nonstationary, the ability of time and frequency domain statistics to characterize different degradation states of the same bearings is relatively poor. For instance, kurtosis is insensitive

to initial damage [3] and it can hardly characterize the slight degradation state exactly. These years, the information entropy theory is widely used in signal processing and fault diagnosis and it develops into different forms of entropy with different properties such as approximate entropy (ApEn), sample entropy (SampEn), multiscale entropy (MSE), spatial information entropy (SIE), and fuzzy entropy (FuzzyEn) [4, 5]. These entropy features apply nonlinear dynamics theory which is different from traditional time-domain indexes in health monitoring and fault identification and have made achievements. Compared with the preset fault pattern recognition, rolling bearing degradation state recognition in its whole life is more ambiguous and complex. However, a single feature of vibration signals can only reflect fault characteristics of rotating machines at a certain fault degree and this can result in problems such as recognition inaccuracy, system instability, and ambiguous recognition results [6]. To address these problems, multidomain feature fusion is widely used in degradation state recognition and fault prediction of rotating machines [7, 8].

However, high-dimensional feature vector composed by multidomain features inevitably has the problems of information redundancy and characteristic conflict and the

effective information is easy to be submerged by high-dimensional data [9]. Moreover, the use of high-dimensional data leads to a sharp increase in the amount of calculation that is not conducive to the real-time identification of rolling bearing degradation state. Manifold learning theory has the ability to identify low-dimensional nonlinear structure which is hidden in high-dimensional data and thus, in recent years, those manifold learning algorithms including locally linear embedding (LLE), locally preserving projection (LPP), isometric feature mapping (IsoMap), and Laplacian eigenmaps (LE) [10]. By the neighbor graphs obtained from high-dimensional features, LPP algorithm can gain its projection in the low-dimensional space. In this way, fusion and reduction of high-dimensional data are achieved. Compared with IsoMap and LLE, LPP has an advantage of simple calculation and fast processing speed. The results of LPP are closely related to nearest neighbor parameters that have no definite criterion. Therefore, the optimized parameters are obtained by repeated experiments. Reference [11] proposed a modified kernel distance measure sensitivity factor to measure the ability that fault features characterize different fault patterns. In view of this, LPP algorithm can be optimized by taking the sensitivity factor as object function. When the factor reaches its maximum, the effect of LPP feature fusion is best.

Considering that the actual rolling bearing degradation states perform strong fuzziness and the boundaries of different degradation states are difficult to determine, Fuzzy C-means (FCM) clustering [12] and Gustafson-Kessel (GK) clustering [13] are widely used in fault diagnosis. And Gath-Geva (GG) clustering improves FCM and GK algorithm by fuzzy maximum likelihood estimation distance norm and the clustering effect is better [14].

Based on the above analysis, a rolling bearing degradation state identification method based on fusion and dimension reduction of multidomain features and GG clustering is put forward in this paper. Six features computed from information entropy and time domain are fused by LPP optimized by genetic algorithm (GA-LPP) in order to separate the training points of different degradation degrees as clear as possible. Finally, degradation state recognition is realized by GG clustering and the principle of maximum membership.

2. Multidomain Feature Extraction

2.1. Time-Domain Features. In order to fully characterize different degradation states of rolling bearings, multi-time-domain features are needed to analyze. Common time-domain indexes include mean, standard deviation (STD), root mean square (RMS), root-square amplitude, skewness, peak to peak, waveform index, pulse index, margin index, partial slope index, and kurtosis. These features are examined from three aspects of ability to follow degradation trend, monotonicity, and data smoothing. Three features including STD x_{std} , RMS x_{rms} , and root-square amplitude x_r are selected to compose a three-dimensional feature matrix as follows:

$$X_1 = [x_{std}; x_{rms}; x_r]. \quad (1)$$

2.2. Entropy Features. Combined with LCD theory [4], relative entropy theory, and multivariate multiscale entropy theory, the entropy features including LREE, LRSE, and TMSE are constructed below.

2.2.1. LREE and LRSE. (1) According to the LCD noise reduction criterion guided by the mutual correlation coefficient [15], the vibration signal is decomposed and reconstructed. Suppose that M samples in degradation state $x_i(t)$ ($i = 1, 2, \dots, M$) and a single sample $x_b(t)$ in normal state are acquired from the reconstructed signal.

(2) With the development of rolling bearing degradation state, the energy at characteristic frequency and its multiplications will become larger in the frequency spectrum which is obtained by LCD and Hilbert transform. For rolling bearings, different fault modes have different vibration characteristics. For a certain rolling bearing fault mode just as inner ring pitting, its vibration characteristic frequency and its frequency multiplication can be calculated by the following formula:

$$f_g = \frac{Z}{2} \left(1 + \frac{d}{D} \cos \alpha \right) f_r \times g, \quad (g = 1, 2, \dots, G), \quad (2)$$

where Z is roller number of the bearings, d is roller diameter, D represents pitch diameter, α denotes contact angle, and f_r is the rotor frequency.

(3) The sum energy of all samples at the characteristic frequency f_g ($g = 1, 2, \dots, G$) is computed as follows:

$$E_L^g = \sum_{i=1}^M E_i^g + E_b^g. \quad (3)$$

(4) At the characteristic frequency f_g , the energy proportion of $x_b(t)$ and $x_i(t)$ in the sum energy is p_b^g and p_i^g :

$$p_b^g = \frac{E_b^g}{E_L^g}, \quad (4)$$

$$p_i^g = \frac{E_i^g}{E_L^g}.$$

(5) The LREE between normal state and degradation state is defined as follows:

$$\text{LREE}_b^i = \sum_{g=1}^G \left| p_i^g \log \frac{p_i^g}{p_b^g} \right|. \quad (5)$$

(6) The singular value spectrum of normal samples and degradation samples can be obtained by singular value decomposition (SVD):

$$\sigma_b = [\sigma_b^1, \sigma_b^2, \dots, \sigma_b^G], \quad (6)$$

$$\sigma_i = [\sigma_i^1, \sigma_i^2, \dots, \sigma_i^G].$$

(7) Combined with the relative entropy theory, the related probabilities are defined as qp_b^g and qp_i^g :

$$\begin{aligned} qp_b^g &= \frac{\sigma_b^g}{\sigma_L^g}, \\ qp_i^g &= \frac{\sigma_i^g}{\sigma_L^g}, \end{aligned} \quad (7)$$

where $\sigma_L^g = \sum_{i=1}^M \sigma_i^g + \sigma_b^g$.

(8) The LRSE between normal state and degradation state is defined as follows:

$$\text{LRSE}_b^i = \sum_{g=1}^G \left| qp_i^g \log \frac{qp_i^g}{qp_b^g} \right|. \quad (8)$$

2.2.2. TMSE. Through LCD, there is enough degradation state information in the first two signal components whose cross-correlation coefficient is higher than others. For the two components whose sequence length is N , after coarse grain, two-element embedding reconstruction, composite delay vectors, and thresholds setting, assuming that the two composite delay vectors' embedding dimensions are m and $m+1$, the conditional probabilities are, respectively, $P^m(r)$ and $P^{m+1}(r)$ when similar capacity limit is r . TMSE can be expressed as the natural logarithm of the conditional probabilities' ratio:

$$\text{TMSE}(M, \lambda, r, N) = \ln \left[\frac{P^m(r)}{P^{m+1}(r)} \right], \quad (9)$$

where M is embedding vector and λ is delay vector.

The above three kinds of entropy features constitute another three-dimensional feature matrix, which can be expressed as

$$X_2 = [x_{\text{LREE}}; x_{\text{LRSE}}; x_{\text{TMSE}}]. \quad (10)$$

Above all, entropy features and time-domain features constitute a six-dimensional multidomain feature matrix:

$$X = [X_1; X_2]. \quad (11)$$

3. Optimized LPP Based on GA

3.1. The Principle of LPP. LPP algorithm can retain the nonlinear structure and local characteristics inside the data when it is applied for high-dimensional data reduction. The algorithm principle can be shown as below [16].

For n data samples with D dimensional space $X = \{x_1, x_2, \dots, x_n\}$, the matrix $Z = \{z_1, z_2, \dots, z_n\}$ is its low-dimensional space samples, where z_i ($i = 1, 2, \dots, d$) is a d dimensional vector ($d \ll D$). The similarity matrix can be defined by the following formula:

$$Q_{ij} = \begin{cases} \exp\left(-\frac{\|x_i - x_j\|^2}{t}\right), & -\|x_i - x_j\|^2 < \varepsilon \\ 0, & \text{others,} \end{cases} \quad (12)$$

where x_i and x_j are the nearest neighbor points and t is a constant.

LPP algorithm can be achieved by solving the following optimization problem:

$$\begin{aligned} \arg \min_W \sum_{ij} (W^T x_i - W^T x_j) Q_{ij} \\ = \arg \min_W W^T X L X^T W \end{aligned} \quad (13)$$

which needs to satisfy $W^T X D X^T W = 1$ and $L = D - Q$ is Laplace operator. The matrix $D_{ii} = \sum_j Q_{ij}$ reflects the density of the data distribution. Then, the transform matrix can be calculated by solving the generalized eigenvalue decomposition problem:

$$X L X^T W = \lambda X D X^T W. \quad (14)$$

In the above formula, the matrix $X D X^T$ is sometimes a singular case. For this problem, the feature set is usually projected onto a PCA subspace and in this way the singularity can be eliminated. And then the following linear mapping can be obtained:

$$\begin{aligned} x \longrightarrow z = W^T x, \\ W = W_{\text{PCA}} W_{\text{LPP}}. \end{aligned} \quad (15)$$

3.2. Kernel Space Measure Sensitivity Factor. In order to evaluate the distinction effect of different degradation states by training samples after fusion and dimension reduction, Zheyuan et al. [17] propose that distance between different types of samples in kernel space is taken as the basis of feature evaluation. However, in clustering analysis, clustering center selection not only depends on the distance, but also depends on the degree of aggregation of the same type of points. Therefore, reference [11] takes the ratio of different types' distance and divergence of the same type as the measure factor in kernel space. And this factor is regarded as the distinguishing criterion for high accuracy. The Gaussian radial basis kernel function is selected to calculate the distance between p_1 and p_2 . The form is as follows:

$$K(x, y) = \exp\left(-\frac{\|x - y\|^2}{2\sigma^2}\right). \quad (16)$$

Then, the distance between two points can be expressed as

$$D_k(p_1, p_2) = \sqrt{2 - 2K(p_1, p_2)}. \quad (17)$$

On this basis, the average distance between training samples of type a and type b can be calculated as follows:

$$D_k^{ab} = \frac{1}{N_a N_b} \sum_{i=1}^{N_a} \sum_{j=1}^{N_b} D(x_i^a, x_j^b), \quad (18)$$

where $a = 1, 2, \dots, C$; $b = 1, 2, \dots, C$. C is the number of sample categories. N_a and N_b are the number of samples of type a and type b .

The average distance between different sample categories is

$$J_k = \frac{1}{C(C-1)} \sum_{a=1}^C \sum_{b=1}^C D_k^{ab}. \quad (19)$$

The divergence of the same sample category can be expressed as

$$S_k = \frac{1}{C} \sum_{a=1}^C \left(\frac{1}{N_a} \sum_{i=1}^{N_a} D_k(x_i^a, \bar{x}_a) \right), \quad (20)$$

where \bar{x}_a is average of training samples of category a .

According to the definition, the kernel space measure sensitivity factor is

$$\varepsilon = \frac{J_k}{S_k}. \quad (21)$$

3.3. Optimization Based on GA. In order to make the fusion features gained from LPP dimension reduction distinguish different degradation states better, genetic algorithm (GA) is applied to optimize the kernel space where there are kinds of training samples. GA is a newly developing algorithm to search an optimal solution. The process of GA algorithm mainly includes population initialization, crossover, mutation, fitness calculation (individual evaluation), and selection (population replacement). The kernel space measure sensitivity factor is taken as the fitness function for optimization and the optimal individual is the case where the discrimination of different degradation states is highest.

Studies have shown that the clustering effect of LPP fusion features will change along with the changing kernel space. In the interest of finding the optimal kernel space, all training samples need to do affine transformation. Take 3D fusion features as an example, one training point is set as $P_1(x_0, y_0, z_0)$ and affine transform angles are set as $\theta_1 \in [0, 2\pi]$ and $\theta_2 \in [0, 2\pi]$. So the affine transformation matrix is

$$A = \begin{bmatrix} \sin \theta_1 \cos \theta_2 & \sin \theta_1 \sin \theta_2 & \cos \theta_1 & 0 \\ 1 & 0 & 0 & -\sin \theta_1 \cos \theta_2 \\ 0 & 1 & 0 & -\sin \theta_1 \sin \theta_2 \\ 0 & 0 & 1 & -\cos \theta_1 \end{bmatrix}. \quad (22)$$

The new sample feature points after kernel space transformation can be computed by the following equation:

$$X = A^{-1}b. \quad (23)$$

The two affine transform angles are used as the training entity and the individuals are randomly generated to complete initialization. By the optimization process of GA, the training sample clustering effect is found to be the best.

4. GG Clustering Algorithm

For the training sample set $X = \{X_1, X_2, \dots, X_N\}$, it is assumed that each sample is made up of d characteristics: $X_k = \{x_{k1}, x_{k2}, \dots, x_{kd}\}$. After initialization, all samples are divided into C categories; namely, the number of clustering classifications is C ($2 \leq C \leq N$). The clustering centers of all categories are $V_i = \{v_1, v_2, \dots, v_i\}$ and the membership matrix is $U = \{u_{ik}\}_{C \times N}$. The element $u_{ik} \in [0, 1]$ represents the membership degree of the k training sample to the i degradation state ($1 \leq i \leq C$). In GG algorithm, the following objective function can reach the minimum value with the iterative adjustment of U and V :

$$J_m(U, V) = \sum_{i=1}^C \sum_{k=1}^N (\mu_{ik})^m D_{ik}^2, \quad (24)$$

where m is the weighted index generally taken to 2.

Different from FCM clustering, D_{ik} indicates the distance measure calculated by the covariance matrix in GG clustering. In that way, the data samples of different directions and shapes can be reflected effectively.

5. The Process of Degradation State Identification

The original vibration signal is preprocessed by LCD. The time-domain features of STD, RMS, and root-square amplitude and the entropy features of LREE, LRSE, and TMSE are extracted from the selected signal components to compose the original characteristic set. The degradation state recognition processes are as shown in Figure 1.

The degradation state recognition algorithm mainly contains the following key steps:

- (i) *LCD Pretreatment.* According to the cross-correlation coefficient between the LCD components and the original signal, the useful components can be chosen. Considering the amount of information existing in components and the time of computation, the first two components whose coefficient is higher than others are selected for further analysis after many tests.
- (ii) *Feature Extraction and Fusion.* Six-dimensional multiple domain features are fused by LPP algorithm and the intrinsic dimension is three according to the maximum likelihood estimation. Therefore, the target dimension of feature fusion is set as three. On the basis of the maximum sensitive factor principle, the fusion features are optimized by GA to find the best kernel space for clustering analysis.
- (iii) The clustering centers are determined by GG algorithm and the rolling bearing degradation identification is achieved by the principle of maximum membership degree.

6. Instance Verification

6.1. Experimental Platform and Data Preprocessing. The bearing full-life data used in this paper comes from Hangzhou

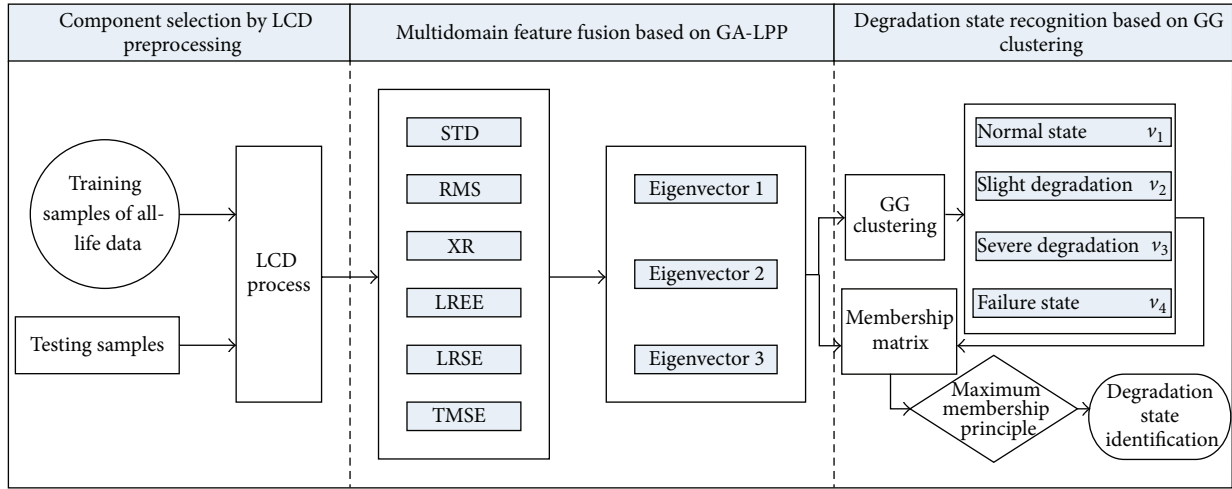


FIGURE 1: Flow chart of the identification method.

TABLE 1: The experimental parameters.

Motor speed	Sampling interval	Sampling time	Sampling frequency
1500 r/min	10 min	1 s	25.6 kHz

bearing test and research center [18]. As is shown in Figure 2(a), the test platform mainly consists of a ABLT-1A bearing test machine, a signal acquisition module, and state monitoring equipment. As Figure 2(b) shows, four CA-YD-139 acceleration sensors are, respectively, fixed up on four bearing test stations and connected to DH-5920 dynamic signal acquisition instrument. Four sets of rolling bearings can be intensively tested and multiple sets of full-life vibration data can be stored simultaneously. What is more, four thermal resistors and a YD-1 acceleration sensor are connected with a signal amplifier to monitor the operating parameters. When the index exceeds the alarm threshold, the test machine will stop working.

Deep groove ball bearings are widely used in rotating machinery. There is practical significance in engineering taking typical type of 6204 bearing as testing object. The real bearing in normal state is shown in Figure 3(a). The specific parameters are set as shown in Table 1.

When the test bench running time reaches 9600 minutes, the machine is shut down. Inner ring pitting occurs in the bearing at number 4 station and that result in bearing failure (as shown in Figure 3(b)).

The collected 960 groups of vibration data record the whole process of rolling bearing from normal state to failure state. Figure 4 shows the real-time monitoring curves of average amplitude versus time which reflect different degradation states of rolling bearing clearly. According to the change of curve amplitude and curvature, the rolling bearing performance variation can be initially divided into four states: normal state, slight degradation, severe degradation, and failure state. The details are presented in Table 2.

The original signal is preprocessed by LCD to get 10 intrinsic scale components (ISCs) and the first 5 ISCs are

TABLE 2: The division of degradation state of rolling bearing performance.

Degradation state	Normal state	Slight degradation	Severe degradation	Failure state
Sample group	1–682	683–802	803–900	901–960

shown in Figure 5. Further the cross-correlation coefficient between each component and the original signal is calculated and the value relation is as follows:

$$ISC_1 > ISC_3 > ISC_2 > ISC_4 > ISC_5. \quad (25)$$

What is more, there are only the first and the third ISC whose coefficient is more than 0.5, respectively, 0.6487 and 0.5395. Therefore, the two components are taken as signal source for degradation feature extraction.

6.2. Degradation Feature Fusion and Optimization. According to the degradation state division in Table 2, 100 groups of normal data, 100 groups of slight degradation data, 60 groups of severe degradation data, and 30 groups of failure data are selected as training samples. The characteristic indexes of different degradation states are extracted and normalized, respectively. The 3D time-domain feature points are shown in Figure 6. In the bearing degradation process from normal state to failure state, these three features are monotonically increasing and the effect of failure state distinguishing is obvious. However, the points of the other three degradation states are mixing severely and cannot be distinguished clearly. Although the time-domain features such as RMS are easy to get and have good stability to characterize degradation states, literature [19] indicates that these time-domain features are not sensitive to early bearing fault including slight degradation and severe degradation until bearing failure occurs. What is more, reference [20] points out that rolling bearings' vibration signals present nonlinear characteristics, and these three traditional time-domain features are similar and can

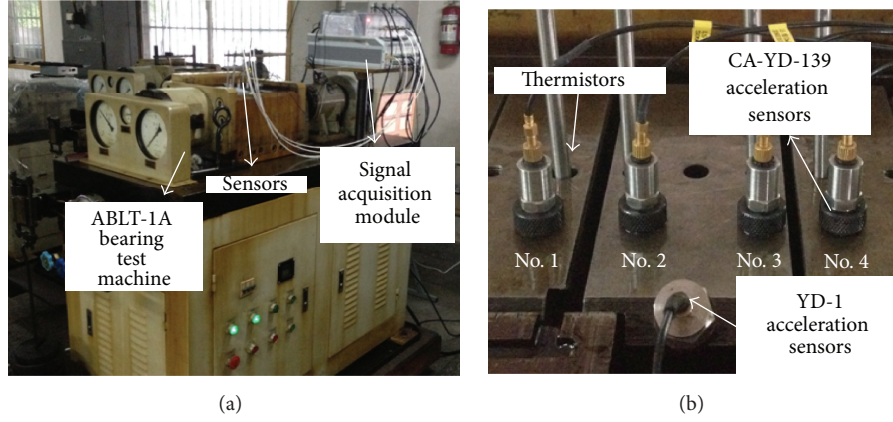


FIGURE 2: Bearing life experiment layout: (a) complete machine and (b) sensors.



FIGURE 3: 6204 Bearing: (a) normal state and (b) inner ring pitting.

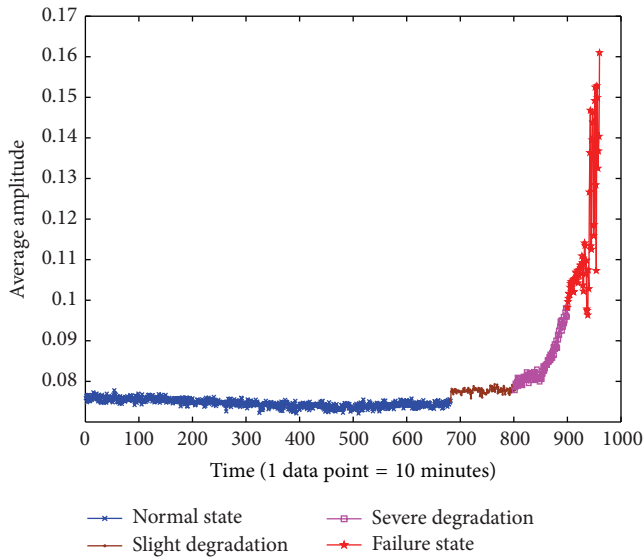


FIGURE 4: The preliminary division of degradation state based on average amplitude.

hardly make an accurate evaluation of the early degradation states of the bearings. These arguments explain clearly why the other three degradation states except for the failure one

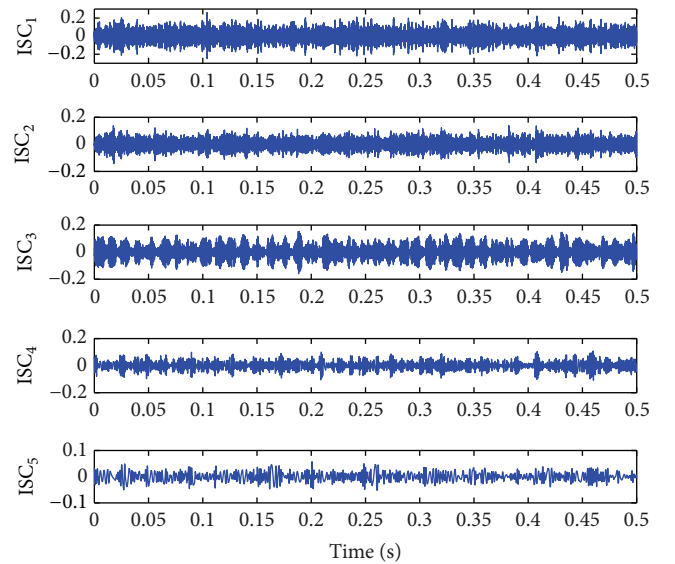


FIGURE 5: The first five-order components of LCD.

are mixed severely and cannot be distinguished by 3D time-domain features.

Similarly, the 3D complexity feature points made up of entropy indexes of LREE, LRSE, and TMSE (scale factor is

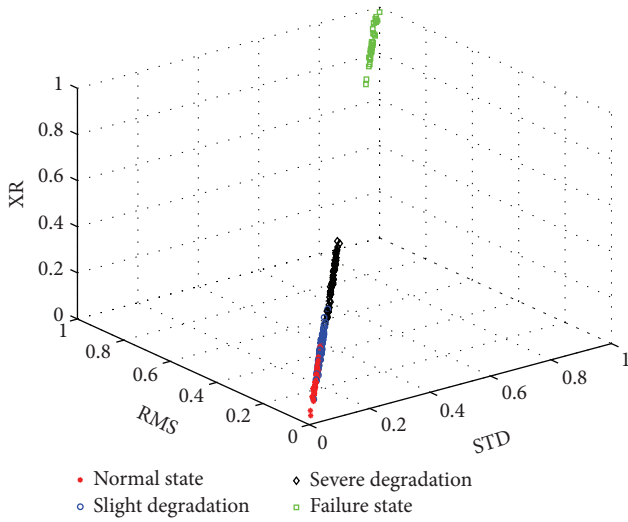


FIGURE 6: Space distribution of time-domain features.

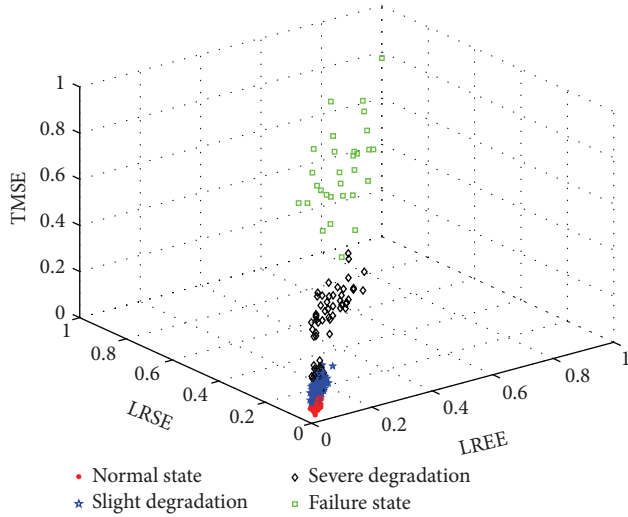


FIGURE 7: Space distribution of entropy features.

15) are shown in Figure 7. The entropy vector can distinguish normal state, slight degradation, and severe degradation on the whole. Nevertheless, in the failure state, the training samples' clustering effect is unsatisfying. Reference [21] demonstrates that entropy indexes are sole dependent on the probability distribution of the event occurrence in bearing fault signals. They are sensitive to the degradation state changing but are more susceptible to spurious vibrations. When the bearing comes to failure state, the violent condition changing will make the vibration signals mixed with a lot of spurious components and the entropy features cannot stably characterize the failure state of bearings. Therefore, the 3D entropy features at failure state show strong discreteness in Figure 7.

In order to improve the discrimination effect of different degradation states, the above time-domain features and entropy features need to be fused. Therefore, the six-dimensional multidomain feature vectors are input to the

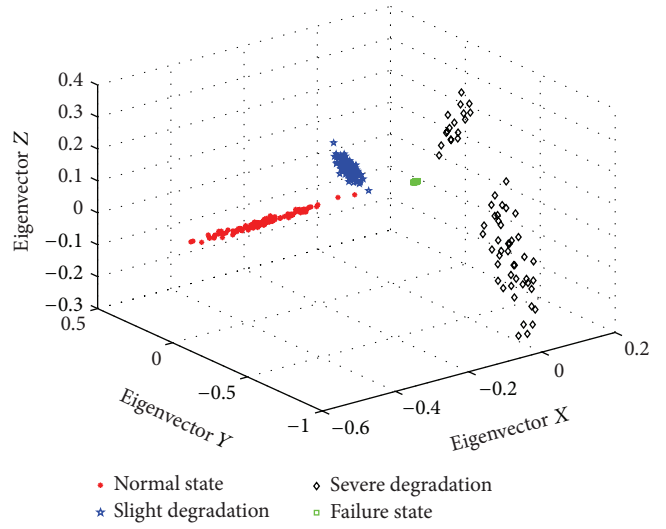


FIGURE 8: Space distribution of LPP fusion features.

LPP for feature fusion and dimension reduction. In order to ensure the information exchanging among the neighborhoods, the neighborhood number k should not be too small; yet if k is too large, the local features can be incomplete. Generally analyzed, the size of k should be between d and N where d is the intrinsic dimension and N is the number of training samples in each category. In this paper, $d = 3$ and $N = 30$. Thus, $3 < k < 30$.

The clustering effect is better when $k = 7$ that is presented in Figure 8. Compared with the time-domain features and the entropy features, the degradation state distinguishing ability of the LPP fusion features is better and the clustering effects of normal state, slight degradation, and failure state are satisfying. But the robustness of fusion features in severe degradation state is relatively poor and this results in the fact that the same severe degradation state is divided into two sample parts. Meanwhile, the sample class spacing is relatively small and the clustering effect is not good. So the process of feature fusion needs to be optimized.

The kernel space measure sensitive factor is taken as the objective function. According to formula (22) and formula (23), the kernel space is optimized by GA so that the factor has a maximum value. In order to improve the convergence speed and ensure the search quality, the population size is set as $N = 20 \sim 200$. After several experiments, $N = 30$. The larger the crossover probability is, the higher the loss rate of excellent results is. But when the probability is too small, the search will be blocked. In general, crossover probability $P_c = 0.6 \sim 1.0$ and here it is 0.8. Mutation probability generally should not be too large; otherwise GA will become a random search method and the precision and speed of convergence will be influenced. Therefore, the mutation probability $P_m = 0.03$.

As shown in Figure 9, after 26 iterations, the kernel space measure sensitivity factor tends to be stable and the maximum value is achieved. And the optimized affine transformation angles are $\theta_1 = 1.4910$ and $\theta_2 = 3.8532$. Figure 10 presents the space distribution of optimized fusion

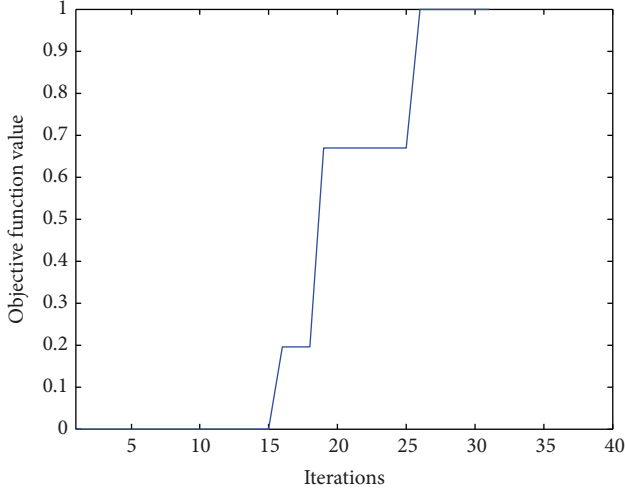


FIGURE 9: The curve of objective function with iterations.

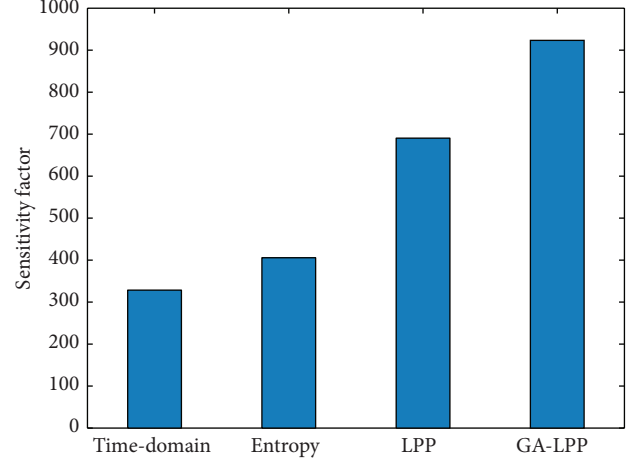


FIGURE 11: Clustering effects of different combinations of features.

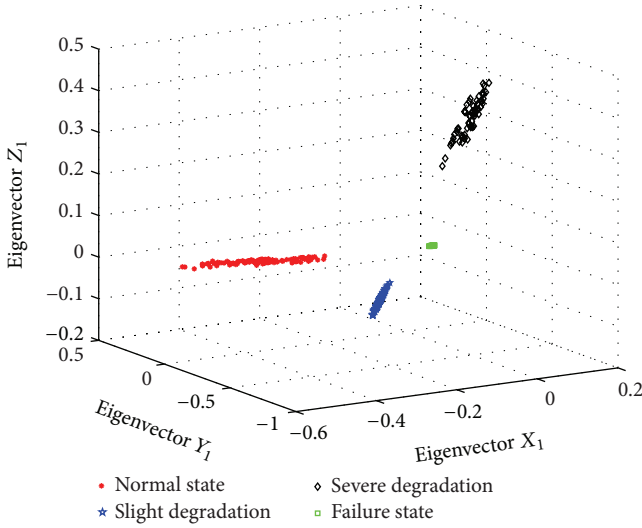


FIGURE 10: Space distribution of GA-LPP fusion features.

feature points. In comparison with Figure 8, the optimized fusion features distinguish different degradation states better than the original features and especially the clustering effect of training samples in severe degradation state improves a lot. What is more, the different class distinctions are further widening. Thus, the optimization effect is obvious.

In order to furtherly illustrate the excellent performance of the proposed method, the sensitivity factors of time-domain features, entropy features, LPP fusion features, and GA-LPP fusion features are calculated, respectively, and the result is just as Figure 11 shows. The kernel space measure sensitivity factor of GA-LPP fusion features is the maximum one and it indicates that the fusion features have a strong ability to characterize different bearing degradation states after GA optimization.

6.3. Degradation State Recognition Based on GG Method. According to the number of bearing degradation states, the number of clustering centers is determined as $c = 4$. The

weighted factor is $m = 2$ and the iterative stopping threshold value is 10^{-5} . The 290×3 matrix composed by GA-LPP fusion features is computed by GG clustering and the clustering center matrix is

$$V = \begin{bmatrix} -0.2388 & 0.0078 & 0.0097 \\ -0.0076 & 0.1597 & -0.1367 \\ 0.0237 & -0.4320 & 0.3697 \\ 0.0595 & -0.0016 & -0.0027 \end{bmatrix}. \quad (26)$$

In accordance with Table 1, every 5 groups of data are chosen randomly as testing samples from each degradation state. The selected 20 groups of data's multidomain features are optimized by GA-LPP at the same affine transformation angles. The fusion feature space distribution is shown in Figure 12 where the testing feature points are well distributed around the clustering centers and the testing sample spacing is large enough. This method can effectively avoid identification misjudgment and improve the recognition accuracy.

The membership matrix U is established based on grey correlation analysis. Based on this, bearing degradation state recognition is realized guided by the principle of maximum membership value. Table 3 is the membership matrix between the testing samples and each standard degradation state. By comparing the membership value of the same sample point and different degradation states, the recognition result is the degradation state whose membership value is maximum. Here are two LPP results before and after GA optimization. Without GA optimization, LPP fusion features judge slight degradation state as normal state and severe degradation state is mistaken as failure state. The accuracy of degradation state recognition is only 85%. In comparison, GA-LPP fusion features have a better distinguishing ability. 20 groups of identification results are in complete agreement with the real degradation states and the excellent performance of the proposed method is verified.

TABLE 3: Result of the rolling bearing degradation state identification.

Real state	Sample number	Normal state		Slight degradation		Severe degradation		Failure state		Recognition result	
		LPP	GA-LPP	LPP	GA-LPP	LPP	GA-LPP	LPP	GA-LPP	LPP	GA-LPP
Normal state	1	0.8267	0.8743	0.4836	0.3241	0.1256	0.2520	0.3545	0.2147	✓	✓
	2	0.8580	0.8975	0.3348	0.2334	0.2053	0.2683	0.2770	0.2091	✓	✓
	3	0.9181	0.9290	0.3580	0.2181	0.2784	0.1682	0.2901	0.2412	✓	✓
	4	0.6187	0.8452	0.7820	0.2452	0.3029	0.1314	0.3432	0.3455	×	✓
	5	0.9036	0.9342	0.2028	0.1465	0.2113	0.2101	0.1982	0.1136	✓	✓
Slight degradation	6	0.7578	0.1750	0.5978	0.9012	0.2214	0.1018	0.3010	0.2071	×	✓
	7	0.2677	0.2205	0.8888	0.9001	0.2036	0.1815	0.2350	0.2274	✓	✓
	8	0.2660	0.1827	0.8422	0.9126	0.2151	0.1312	0.2855	0.2562	✓	✓
	9	0.2790	0.1921	0.8241	0.9075	0.2382	0.1833	0.2287	0.1821	✓	✓
	10	0.2395	0.2275	0.8461	0.8925	0.2108	0.2063	0.2956	0.2632	✓	✓
Severe degradation	11	0.1506	0.1231	0.2487	0.2196	0.8675	0.9196	0.3252	0.1753	✓	✓
	12	0.2216	0.2100	0.3150	0.2473	0.6015	0.8825	0.7757	0.2104	×	✓
	13	0.2067	0.1143	0.2821	0.2235	0.8226	0.8333	0.2862	0.2690	✓	✓
	14	0.1922	0.1813	0.2421	0.2193	0.8483	0.8777	0.2988	0.2734	✓	✓
	15	0.1232	0.1112	0.1726	0.1853	0.8793	0.9169	0.2276	0.2112	✓	✓
Failure state	16	0.1384	0.1028	0.2005	0.1714	0.2308	0.2158	0.8903	0.9072	✓	✓
	17	0.1766	0.1433	0.1938	0.1526	0.2457	0.2264	0.8900	0.9157	✓	✓
	18	0.1845	0.1664	0.1584	0.1687	0.2283	0.2041	0.8775	0.8909	✓	✓
	19	0.1870	0.1659	0.1426	0.1585	0.2012	0.1993	0.9063	0.9305	✓	✓
	20	0.1985	0.1753	0.2232	0.2054	0.2106	0.1980	0.8959	0.9231	✓	✓

Note: “✓” represents the right recognition result and “×” represents the wrong one.

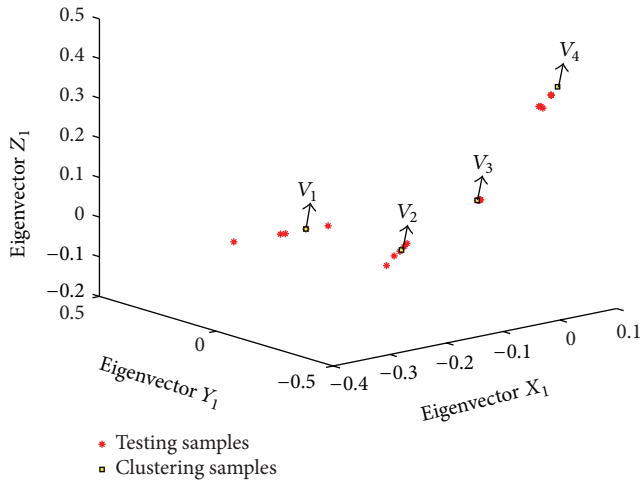


FIGURE 12: Space distribution of test samples and clustering centers.

7. Conclusion

In order to improve degradation state recognition accuracy in rolling bearing all-life cycle, this paper proposes a new degradation state identification method based on GA-LPP and GG clustering. Through the actual signal processing and analysis, the following conclusions can be obtained.

- (1) Compared with preset fault degrees, it is difficult to distinguish different degradation states in the bearing cycle life. Single domain features usually measure

degradation states from only one perspective, so the ability of single domain features to characterize complex and fuzzy degradation states can be insufficient. In manifold learning theory, LPP algorithm can fuse multidomain features and reduce dimension to improve distinguishing effects of different degradation states.

- (2) The kernel space measure sensitivity factor is taken as the optimization criterion. GA algorithm based on kernel space transformation is applied to optimize the LPP feature fusion process which can separate different degradation samples better. In this way, the clustering effect of the same degradation state is more satisfactory and the accuracy is higher.
- (3) There is some engineering value combining GA-LPP multidomain feature fusion and GG clustering in the field of degradation state recognition.
- (4) The GA parameter setting has a certain effect on the convergence speed and the calculation precision. Even if the parameters are the same for repeated experiments, the results can fluctuate. Therefore, the following work is to improve the proposed method applicability by parameter optimization and enhancing GA searching stability.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

This project is supported by National Natural Science Foundation of China (Grant no. 51541506).

References

- [1] P. Nguyen, M. Kang, J.-M. Kim, B.-H. Ahn, J.-M. Ha, and B.-K. Choi, "Robust condition monitoring of rolling element bearings using de-noising and envelope analysis with signal decomposition techniques," *Expert Systems with Applications*, vol. 42, no. 22, pp. 9024–9032, 2015.
- [2] S. Zhang, Y. Zhang, and J. Zhu, "Rolling element-bearing feature extraction based on combined wavelets and quantum-behaved particle swarm optimization," *Indian Journal of Thoracic and Cardiovascular Surgery*, vol. 31, no. 1, pp. 605–610, 2015.
- [3] T. Xiao, B.-P. Tang, Y. Qin, and C. Chen, "Degradation trend prediction of rolling bearing based on manifold learning and least squares support vector machine," *Journal of Vibration and Shock*, vol. 34, no. 9, pp. 149–153, 2015.
- [4] J. Zheng, J. Cheng, and Y. Yang, "A rolling bearing fault diagnosis approach based on LCD and fuzzy entropy," *Mechanism and Machine Theory*, vol. 70, pp. 441–453, 2013.
- [5] Y.-K. Wang, H.-R. Li, B. Wang, and B.-H. Xu, "Spatial information entropy and its application in the degradation state identification of hydraulic pump," *Mathematical Problems in Engineering*, vol. 2015, Article ID 532684, 7 pages, 2015.
- [6] L. Yun, Z. Lifeng, and Z. Shujun, "A hand gesture recognition method based on multi-feature fusion and template matching," *Procedia Engineering*, vol. 29, pp. 1678–1684, 2012.
- [7] X. Li, P. Li, L. Jiang, and Y. Cao, "Fault diagnosis method of asynchronous motor based on heterogeneous information feature fusion," *Chinese Journal of Scientific Instrument*, vol. 34, no. 1, pp. 227–233, 2013.
- [8] X. Ding, Q. He, and N. Luo, "A fusion feature and its improvement based on locality preserving projections for rolling element bearing fault classification," *Journal of Sound and Vibration*, vol. 335, pp. 367–383, 2015.
- [9] H.-Y. Pan, Y. Yang, Y.-G. Li, and J. Cheng, "The rolling bearings fault diagnosis method based on manifold learning and improved VPMCD," *Journal of Vibration Engineering*, vol. 27, no. 6, pp. 934–941, 2014.
- [10] 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.
- [11] Z.-Q. Su, B.-P. Tang, and J.-B. Yao, "Fault diagnosis method based on sensitive feature selection and manifold learning dimension reduction," *Journal of Vibration and Shock*, vol. 33, no. 3, pp. 70–75, 2014.
- [12] M. Zarinbal, M. H. Fazel Zarandi, and I. B. Turksen, "Relative entropy fuzzy c-means clustering," *Information Sciences*, vol. 260, pp. 74–97, 2014.
- [13] Y.-I. Kim, D.-W. Kim, D. Lee, and K. H. Lee, "A cluster validation index for GK cluster analysis based on relative degree of sharing," *Information Sciences*, vol. 168, no. 1–4, pp. 225–242, 2004.
- [14] L. Zhang, P. Li, M. Li, S. Zhang, and Z. Zhang, "Fault diagnosis of rolling bearing based on ITD fuzzy entropy and GG clustering," *Chinese Journal of Scientific Instrument*, vol. 35, no. 11, pp. 2624–2632, 2014.
- [15] Y. Yang, H. Pan, and J. Cheng, "A rolling bearing fault diagnosis method based on LCD de-noising and VPMCD," *China Mechanical Engineering*, vol. 24, no. 24, pp. 3338–3344, 2013.
- [16] X.-X. Ding and Q.-B. He, "Machine fault diagnosis based on WPD and LPP," *Journal of Vibration and Shock*, vol. 33, no. 3, pp. 89–93, 2014.
- [17] C. Zheyuan, Y. Jianguo, L. Xianpeng et al., "Feature selection algorithm based on kernel distance measure," *PR & AI*, vol. 23, no. 2, pp. 235–240, 2010.
- [18] B. Wang, H.-R. Li, Q.-H. Chen, and B.-H. Xu, "Rolling bearing performance degradative state recognition based on mathematical morphological fractal dimension and fuzzy center means," *Acta Armamentarii*, vol. 36, no. 10, pp. 1982–1990, 2015.
- [19] Y. N. Pan, J. Chen, and X. L. Li, "Spectral entropy: a complementary index for rolling element bearing performance degradation assessment," *Proceedings of the Institution of Mechanical Engineers, Part C*, vol. 223, no. 5, pp. 1223–1231, 2009.
- [20] K. Zhu, X. Song, and D. Xue, "A roller bearing fault diagnosis method based on hierarchical entropy and support vector machine with particle swarm optimization algorithm," *Measurement*, vol. 47, no. 1, pp. 669–675, 2014.
- [21] B. Tao, L. Zhu, H. Ding, and Y. Xiong, "An alternative time-domain index for condition monitoring of rolling element bearings—a comparison study," *Reliability Engineering and System Safety*, vol. 92, no. 5, pp. 660–670, 2007.

