Mechanical Systems and Signal Processing 109 (2018) 330–351

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

journal homepage: www.elsevier.com/locate/ymssp

Hyper-spherical distance discrimination: A novel data description method for aero-engine rolling bearing fault detection

Tong Lin [a], Guo Chen [a,*], Wenli Ouyang [b], Quande Zhang [a], Hongwei Wang [c], Libo Chen [c]

[a] College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China
[b] Innovative Research Center, China Aero-Polytechnology Establishment, Beijing 100076, China
[c] Beijing Aeronautical Technology Research Center, Beijing 100076, China

ARTICLE INFO

Article history:
Received 9 July 2016
Received in revised form 16 September 2017
Accepted 9 January 2018
Available online 10 March 2018

Keywords:
Fault detection
Rolling bearing
Condition monitoring
Feature fusion
Degradation assessment
Feature transform
Aero-engine

ABSTRACT

A novel method called hyper-spherical distance discrimination (HDD) is proposed in order to meet the requirement of aero-engine rolling bearing on-line monitoring. In proposed method, original multi-dimensional features extracted from vibration acceleration signal are transformed to the same dimensional reconstructed features by de-correlation and normalization while the distribution of feature vectors is transformed from hyper-ellipsoid to hyper-sphere. Then, a simple model built up by distance discriminant analysis is used for rolling bearing fault detection and degradation assessment. HDD is compared with the support vector data description (SVDD) and the self-organizing map (SOM) in rolling bearing fault simulation experiments. The results show that the HDD method is superior to the SVDD and SOM in terms of recognition rate. Besides, HDD is applied to a run-to-failure test of aero-engine rolling bearing. It proves that the evaluating indicator obtained by HDD method is able to reflect the degradation tendency of rolling bearing, and it is also more sensitive to initial fault than the root mean square (RMS) of vibration acceleration signal. With the advantages of low computational complexity and no need to tuning parameters, HDD method can be applied to practical engineering effectively.

1. Introduction

Rolling bearing failure is one of the leading causes of aviation accidents. In order to maintain aero-engine uptime at the highest possible level and reduce maintenance costs, maintenance should be carried out in a proactive way. It means a transformation of maintenance strategy from the traditional fail-and-fix practices (diagnostics) to a predict-and-prevent methodology (prognostics) [1]. However, predict-and-prevent methodology is based on effective condition monitoring technology.

Condition monitoring data are very versatile, including vibration data, acoustic data, oil analysis data, etc. Vibration data collection is a widely used approach for fault detection [2–4]. However, the sensitivity of various original features that are characteristics of bearing performance may vary significantly under different working conditions [5]. Hence, it is critical to devise an evaluating indicator that provides a useful and automatic guidance on using the most effective features for bearing degradation assessment without human intervention.

* Corresponding author.
E-mail address: cgzyx@263.net (G. Chen).

https://doi.org/10.1016/j.ymssp.2018.01.009
0888-3270/

Generally, fault detection and condition monitoring of the aero-engine rolling bearings should be considered as a data domain description problem (also called one-class classification), in view of that the fault samples are hard to be acquired in engineering. That is, when nothing about the outlier distribution can be assumed, only a description of the boundary of the target class can be made [6]. Since the on-line monitoring of aero-engine rolling bearing can be regarded as a data domain description problem, it is necessary to study the distribution of multidimensional feature vectors in space in order to establish a more accurate model by greater using the prior knowledge. The distribution of feature vectors cannot be visualized because the dimensionality of feature vectors is usually more than three, but its two-dimensional (2-D) projection can be easily study. If the projection on each 2-D feature plane tends to an ellipse, we can infer that the distribution of feature vectors in high-dimension space tends to a hyper-ellipsoid. When different features are chosen, length and direction of the hyper-ellipsoid principal axis may change. Hence a strongly nonlinear algorithm is needed for describing such a complicated distribution. Different methods have been developed to solve this problem such as SVDD [7–9], SOM [10–12], gaussian mixture model (GMM) [5,13], etc, and have been proved effective in experiments. However, these methods have a deficiency of high computational complexity when training. The bearing detection model is envisioned to reside in the engine controller and operates on-board. The engine controller has to carefully prioritize and distribute computing resources among multiple processes to ensure the safety of the critical tasks such as flight and engine controls. Therefore, a simple fusion model is strongly preferable to a computationally complex one [14].

According to above analysis, the limited computing resources in engineering and the complexity of the model constitute a contradiction. The reason why the models of high computational complexity are chosen is because the described boundary is complicated. So if it is possible to improve the spatial distribution of feature vectors, then it is possible to greatly simplify the algorithm describing the boundary of the data domain. Based on this, a novel method called hyper-spherical distance discrimination (HDD) is proposed. Compared with some typical data domain description like SVDD, SOM etc., HDD has the advantages of low computational complexity and no need to tuning parameters during the training stage. In this study, we implemented HDD on aero-engine rolling bearing monitoring.

The remaining part of the paper is organized as follows. Section 2 introduces the multi-dimensional features used in following sections and briefly describes how to extract these features. In Section 3, two kinds of experiments (including experiment 1: rolling bearing fault simulation experiment and experiment 2: run-to-failure test) were carried out and the distribution of original feature vectors is discussed in detail. The discussion reveals a potential approach for simplifying the distribution. Section 4 proposes a novel method for bearing fault detection and degradation assessment. Section 5 shows the results of two experiments. The performance of proposed method under different operating conditions and different measurement points is compared with SVDD and SOM. Section 6 discusses some problems in detail. Finally, conclusions are made in Section 7.

2. Feature extraction

2.1. Time-domain features

Six dimensionless time-domain features used in this study are summarized in Table 1, including shape indicator T_{SI}, crest indicator T_{CI}, impulse indicator T_{MI}, clearance indicator T_{CLI}, kurtosis T_{KU} and skewness T_{SK}, where y_i is raw waveform data, y_{pi} is the maximum absolute value of each section where the raw waveform data is divided into 10 sections.

2.2. Frequency-domain features

Three dimensionless frequency-domain features used in this study are summarized in Table 2, including frequency center F_{FC}, mean square of frequency F_{MSF} and variance of frequency F_{VF}, where $S(f_i)$ is the spectral amplitude at frequency f_i.

Table 1
Dimensionless time-domain features.

Shape indicator	Crest indicator	Impulse indicator	Clearance indicator	Kurtosis	Skewness
$T_{SI}=\frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i^2)}}{\frac{1}{N}\sum_{i=1}^{N}\lvert y_i\rvert}$	$T_{CI}=\frac{\sum_{i=1}^{10}y_{pi}}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i)^2}}$	$T_{MI}=\frac{\sum_{i=1}^{10}y_{pi}}{\frac{1}{N}\sum_{i=1}^{N}\lvert y_i\rvert}$	$T_{CLI}=\frac{\sum_{i=1}^{10}y_{pi}}{\left[\frac{1}{N}\sum_{i=1}^{N}\sqrt{\lvert y_i\rvert}\right]^2}$	$T_{KU}=\frac{\frac{1}{N}\sum_{i=1}^{N}y_i^4}{\left(\frac{1}{N}\sum_{i=1}^{N}y_i^2\right)^2}$	$T_{SK}=\frac{\frac{1}{N}\sum_{i=1}^{N}y_i^3}{\left(\frac{1}{N}\sum_{i=1}^{N}y_i^2\right)^{\frac{3}{2}}}$

Table 2
Dimensionless time-domain features.

Frequency center	Mean square frequency	Variance of frequency
$F_{FC}=\frac{\sum_{i=0}^{n}f_iS(f_i)}{\sum_{i=0}^{n}S(f_i)}$	$F_{MSF}=\frac{\sum_{i=0}^{n}f_i^2S(f_i)}{\sum_{i=0}^{n}S(f_i)}$	$F_{VF}=\frac{\sum_{i=0}^{n}(f_i-F_{FC})^2S(f_i)}{\sum_{i=0}^{n}S(f_i)}$

There should be some differences between the healthy bearing and fault bearing both in the time-domain waveform and spectrum of vibration acceleration signal. This is why such frequently-used time-domain and frequency-domain features are calculated for bearing fault detection. However, for the surface damage fault of rolling bearing, the characteristic frequencies (or called ball-pass frequencies) are of great significance. Hence, the features corresponding to the characteristic frequencies are extracted in Section 2.3 by applying wavelet transform.

2.3. Wavelet-domain features

As the rolling elements strike a local fault on the outer or inner race, a shock is introduced that excites high frequency resonances of the whole structure between the bearing and the transducer [3]. The key to detecting bearing faults is to capture the low amplitude response caused by bearing defect excitation without including the high amplitude rotational vibration signals and system fundamental resonant frequency responses [15]. To accomplish this, a bandpass filter is used to isolate the signal.

The wavelet transform has the function of adaptive band-pass filtering. Original signal can be decomposed into several sub-signals including detailed signals and approximate signal after wavelet transform. These sub-signals can be treated as the results from the original signal passing through different band-pass filters whose center frequencies and bandwidths are adaptively determined by original signal. And the original signal is the sum of these sub-signals. For example, suppose here is a bearing vibration acceleration original signal whose sampling frequency f_s = 10,240 Hz, then the frequency band of this original signal is 0–5120 Hz. Meanwhile, suppose the original signal is decomposed into 6 sub-signals after wavelet decomposition, including 5 detailed signals d1, d2, d3, d4, d5 and 1 approximate signal a5. Then, approximately, we can think that the frequency bands of such sub-signals are d1:2560–5120 Hz, d2: 1280–2560 Hz, d3: 640–1280 Hz, d4: 320–640 Hz, d5: 160–320 Hz and a5: 0–160 Hz. Generally, such detailed signals cover the rolling bearing resonance frequency bands. However, for different signal samples, the structural resonance frequencies caused by the local damage of rolling bearings are usually not fixed. So in order to find which "band-pass filter" is the best, we design a simple algorithm as following.

By taking the wavelet transform of the bearing signals and postprocessing it with the Hilbert transform and Fourier transform, the wavelet envelope spectrum is obtained. The wavelet envelope spectrum is able to reveal the frequency and amplitude uniquely associated with the damaged bearing component. In the wavelet envelope spectrum $W(f)$, suppose spectral peaks exist at characteristic frequency f_d, which is the ball-pass frequency of outer ring (BPFO), or the ball-pass frequency of inner ring (BPFI), or the ball-spin frequency (BSF) and their frequency multiplications, f_e is the analyzed bandwidth (generally satisfy $f_e > 3\max(f_d)$), and N_e is the number of spectrum lines. Then, the average of envelope spectrum S_{ea} can be calculated by

$$S_{ea} = \frac{1}{N_e}\sum_{i=0}^{N_e} W(f_i) \tag{1}$$

Let S_{ed} to be the average of envelope spectrum at f_d and its kth harmonic, n_e is the number of spectrum line corresponding to f_d and its kth harmonic, then

$$S_{ed} = \frac{1}{n_e}\sum_{k=1}^{n_e} W(kf_d) \tag{2}$$

and the dimensionless feature is given by

$$\Delta S_e = \frac{S_{ed}}{S_{ea}} \tag{3}$$

In general, the characteristic frequency calculated by bearing geometry and rotation speed is not exactly equal to the actual characteristic frequency. Hence a maximum spectral value at f_d and its neighborhoods is chosen as $W(f_d)$.

In specific calculation process, the signal is decomposed into 5 levels by db8 wavelet basis, including 5 detailed signals d1, d2, d3, d4, d5 and 1 approximate signal a5. For each characteristic frequency, the 5 detailed signals are used to calculate the ΔS_e according to Eq. (3), the maximum value of which is chosen as the dimensionless wavelet-domain feature (the feature of BPFO W_{BPFO}, the feature of BPFI W_{BPFI} or the feature of BSF W_{BSF}).

3. Experiments and data analysis

In this section, two kinds of experiments including experiment 1: rolling bearing fault simulation experiment and experiment 2: rolling bearing run-to-failure test were carried out. Taking the data acquired from experiment 1 as an example, the distribution of original feature vectors is discussed in detail. The discussion reveals a potential approach for simplifying the distribution.

3.1. Experiment 1: Fault simulation experiment of rolling bearing

3.1.1. Aero-engine rotor experimental rig

To simulate the rotor vibration of a real aero-engine, an aero-engine rotor experimental rig was designed and manufactured [16]. In comparison to a real aero-engine, the tester has the following features: (1) its size is one third that of a real aero-engine, and its shape is similar to the stator of an aero-engine; (2) its internal structure is simplified: the core machine is simplified to a 0-2-0 support structure, which is a rotor with two disks supported on two bearings, where the outsides of the two disks have no bearings s, and the support stiffness is adjustable to adjust the dynamic characteristics of the system; the multistage compressor is simplified to a single stage disk structure; the compressor blade is simplified to an inclined plane shape; (3) its shaft is solid and rigid and its maximal rotating speed is 7000 rpm; and (4) the rotor is driven by the motor, and the flame flask is canceled. Therefore, the aero-engine rotor tester is a single-rotor system model. A full-scale photo is shown in Fig. 1(a), and a section drawing is shown in Fig. 1(b).

In the aero-engine rotor experimental rig, the compressor part of the rotor is supported by a roller bearing, and the turbine part of the rotor is supported by a ball bearing. In this study, faults were introduced on the ball bearings, and fault simulation experiments were carried out.

3.1.2. Experiment introduction

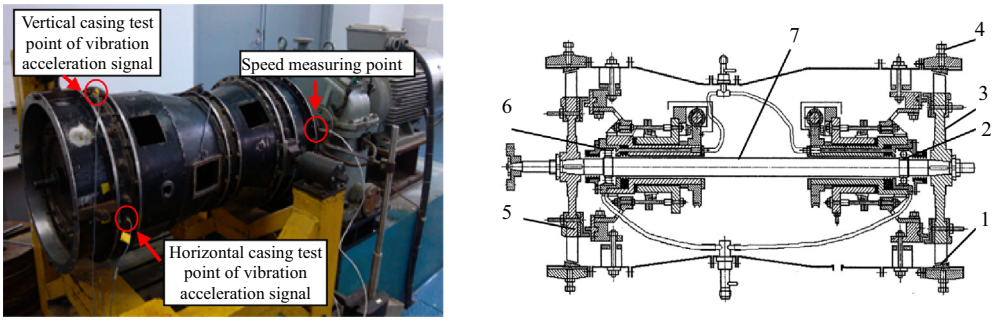
In online monitoring, the sensor usually cannot be placed on the bearing house of aero-engine. Hence, the vibration acceleration sensors were respectively placed on the vertical and horizontal direction of the turbine casing, as shown in Fig. 1(a). However, it also brings a challenge to fault detection. According to reference [16], with a low connect stiffness, there was great attenuation of the vibration response on the casing, while the impulse characteristic of the ball bearings fault was very weak, submerged in other signals.

In order to study the casing responses caused by ball bearing faults, defects were artificially seeded in 6206 ball bearings (as shown in Fig. 2) in the following three configurations: (1) a crack with about 0.6 mm width on outer ring; (2) a crack with about 0.6 mm width on inner ring; (3) a dent with about 1 mm diameter and 2 mm depth on ball. The 6206 ball bearing dimensions are listed in Table 3.

Two groups of tests were carried out, where the rotational speeds were 1500 rpm and 1800 rpm respectively. Vibration signals were collected by means of the USB9234 data acquisition card of the NI Company, the 4805 type ICP acceleration sensors of B&K Company were used to pick up the acceleration signals, and the eddy current sensors were used to measure the rotating speeds. The sampling rate was 10.24 kHz. Each sample contains 8192 points, so the frequency resolution is 1.25 Hz. For 6206 ball bearing, corresponding fault characteristic frequencies in 1500 rpm are BPFO = 89.27 Hz, BPFI = 135.73 Hz and BSF = 57.95 Hz, corresponding fault characteristic frequencies in 1800 rpm are BPFO = 107.12 Hz, BPFI = 162.88 Hz and BSF = 69.53 Hz. Hence, the frequency resolution is sufficient for detecting bearing fault characteristic frequencies.

3.2. Experiment 2: run-to-failure test of rolling bearing

A run-to-failure test was performed to acquire the signal of bearings over a whole lifetime. In this experiment, the faults of rolling bearing were naturally generated, as opposed to the experiment 1 in which the data is for artificially inserted faults, as well as the operating conditions were close to the practical condition.



(a)Photo (b)Section drawing

1-rubbing ring, 2-ball bearing, 3-turbine disk, 4-bolt for deforming rubbing ring, 5-disk of compressor, 6-roller bearing, 7-shaft

Fig. 1. Aero-engine rotor experimental rig.

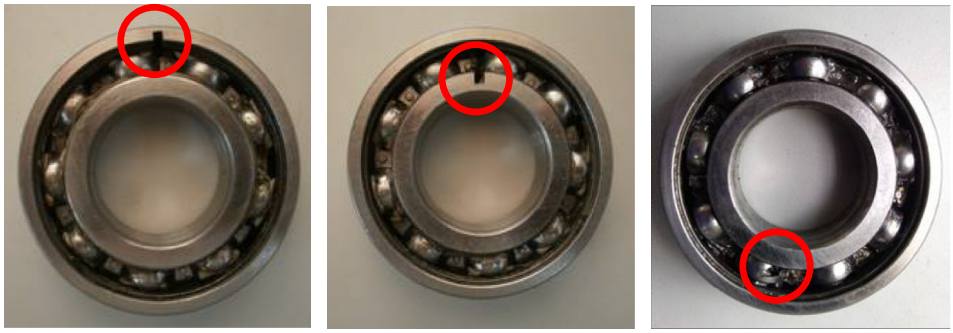


Fig. 2. Faults in 6206 ball bearing of aero-engine rotor experimental rig.

Table 3
6206 ball bearing dimensions (unit:mm).

Diameter of inner ring	Diameter of outer ring	Diameter of ball	Thick	Pitch diameter
30	62	9.5	16	46

3.2.1. Aero-engine ball bearing failure monitoring system

The experimental facilities called aero-engine ball bearing failure monitoring system is designed and developed by Luoyang Bearing Research Institute, composed of a rolling bearing test rig, a power-drive module, a lubrication system, hydraulic loading system, computer monitoring system, etc., as shown in Fig. 3.

Fig. 4 is a section diagram of the rolling bearing test rig. The tested bearing was mounted on the cantilever end of the spindle and was lubricated by 928 aviation lubricating oil. The radial and axial loads were applied through the hydraulic loading system. The whole experiment was controlled and monitored by the computer monitoring system.

3.2.2. Experiment introduction

The dimensions of tested bearing are listed in Table 3. In this test, the operating conditions of tested bearing were close to the practical condition in aero-engine: the rotating speed of the shaft was 12,000 rpm, the radial load was 2.5 kN and the axial load was 3.5 kN. Vibration signals were collected by means of the USB-9234 data acquisition card of the NI Company, the YD-3 type acceleration sensors of Far East Vibration (Beijing) System Engineering Technology Company were used to pick up the acceleration signals. The sampling rate was 10.24 kHz.

Due to the limitation of time and economy, the method of cutting off the oil supply of the tested ball bearing was used to accelerate the bearing degradation. When the test is carried out for a period of time under the condition of insufficient lubrication, the torque of electric motor will increase. Finally, when the motor torque increases over the threshold (set based on previous experiments), the system will automatically trigger protection to suspend the test. At such a time, an initial fault probably occurs in the tested bearing. Therefore, we took the test procedure as follows:

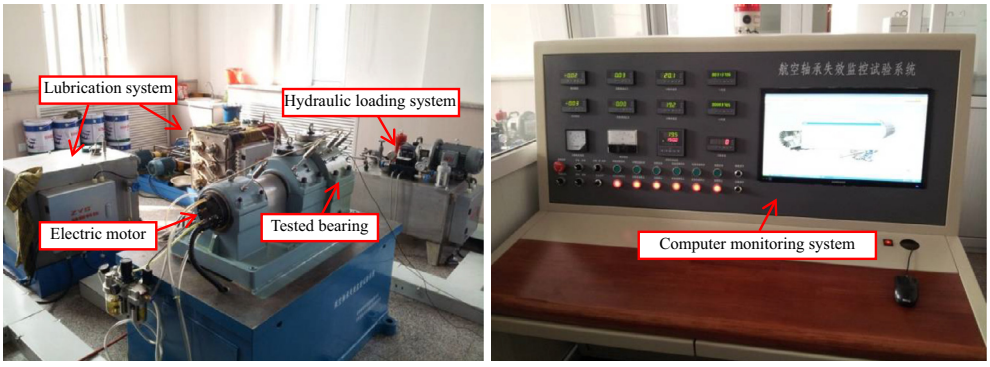


Fig. 3. Aero-engine ball bearing failure monitoring system.

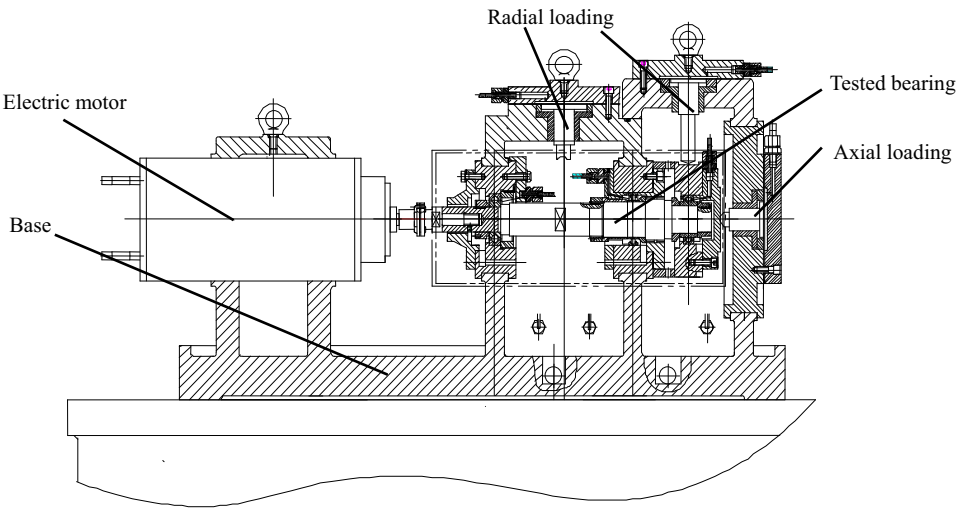


Fig. 4. Section diagram of the rolling bearing test rig.

Step 1. The test was carried out under rated condition (the rotating speed of the shaft was 12,000 rpm, the radial load was 2.5 kN and the axial load was 3.5 kN) for approximately an hour. It is obvious that the tested bearing is healthy during this period.

Step 2. Then, the oil supply was cut off and the tested bearing operated under the condition of insufficient lubrication. The test was suspended when the system automatically triggers protection.

Step 3. The test was resumed under the rated condition. The test was stopped until the monitored RMS value of vibration acceleration was 3 times of the average RMS value in Step 1.

In fact, in order to prove above test procedures are doable, a pre-experiment was conducted. This will be discussed in detail in Section 5.2.1.

3.3. Experimental data analysis

3.3.1. Original features

Take the vibration acceleration signal monitored at vertical casing measurement point in 1500 rpm of experiment 1 for example, the scatter plots of 12 original features (see Section 2 for details) are shown in Fig. 5, where the X-axis is the sequence number of samples, No. 1–110 are points of healthy sample, No. 111–220 are points of inner ring fault sample, No. 221–330 are points of outer ring fault sample and No. 331–440 are points of ball fault sample.

In Fig. 5, it can be seen that: (1) the sensitivity of various original features may vary significantly under different bearing fault type; (2) the fault of inner ring is relatively easy to be detected while ball fault is almost impossible to be detected by single original feature.

Therefore, it is very critical to devise an evaluating indicator (EI) that provides a useful and automatic guidance on using the most effective information in features for bearing fault detection and degradation assessment.

3.3.2. Distribution of original feature vectors

As mentioned in Section 1, it is necessary to study the distribution of the original feature vectors. The 12 dimensional features have been obtained. The distribution of feature vectors cannot be visualized, but its 2-D projection can be easily study. Fig. 6 shows the 2-D projection distribution of normalized feature vectors: Fig. 6(a) is the projection distribution in $T_{SI} - F_{FC}$ plane, Fig. 6(b) is the projection distribution in $F_{VF} - F_{FC}$ plane. T_{SI}, F_{FC} and F_{VF} were extracted from the vibration acceleration signal of healthy rolling bearing.

After normalization, every single feature satisfy $\mu = 0$ and $\sigma = 1$, where μ is the sample average and σ is the sample standard deviation. Therefore, the distribution can be described in the approximate scale. There are $C_{12}^2 = 66$ kinds of combinations of 2-D projection distributions, but all can be grouped into the two categories represented in Fig. 6. In Fig. 6(a), the boundary of the distribution can be approximately described as a circle (the radius of the circle is 2.5), whereas the boundary of the distribution in Fig. 6(b) tends to be an ellipse, whose principal axis are not parallel to the coordinate axes. It is desirable to describe the distribution by using a circle, however, if a circle is used for describing the distribution like Fig. 6(b), two types of errors (false positive and false negative) will be caused inevitably, as shown in Fig. 7.

Furthermore, it could be found in mathematical that the reason why the distribution in Fig. 6(b) is elliptical is because that there is a strong correlation between the features. Namely, when F_{FC} is larger, the F_{VF} tends to be larger (cause that

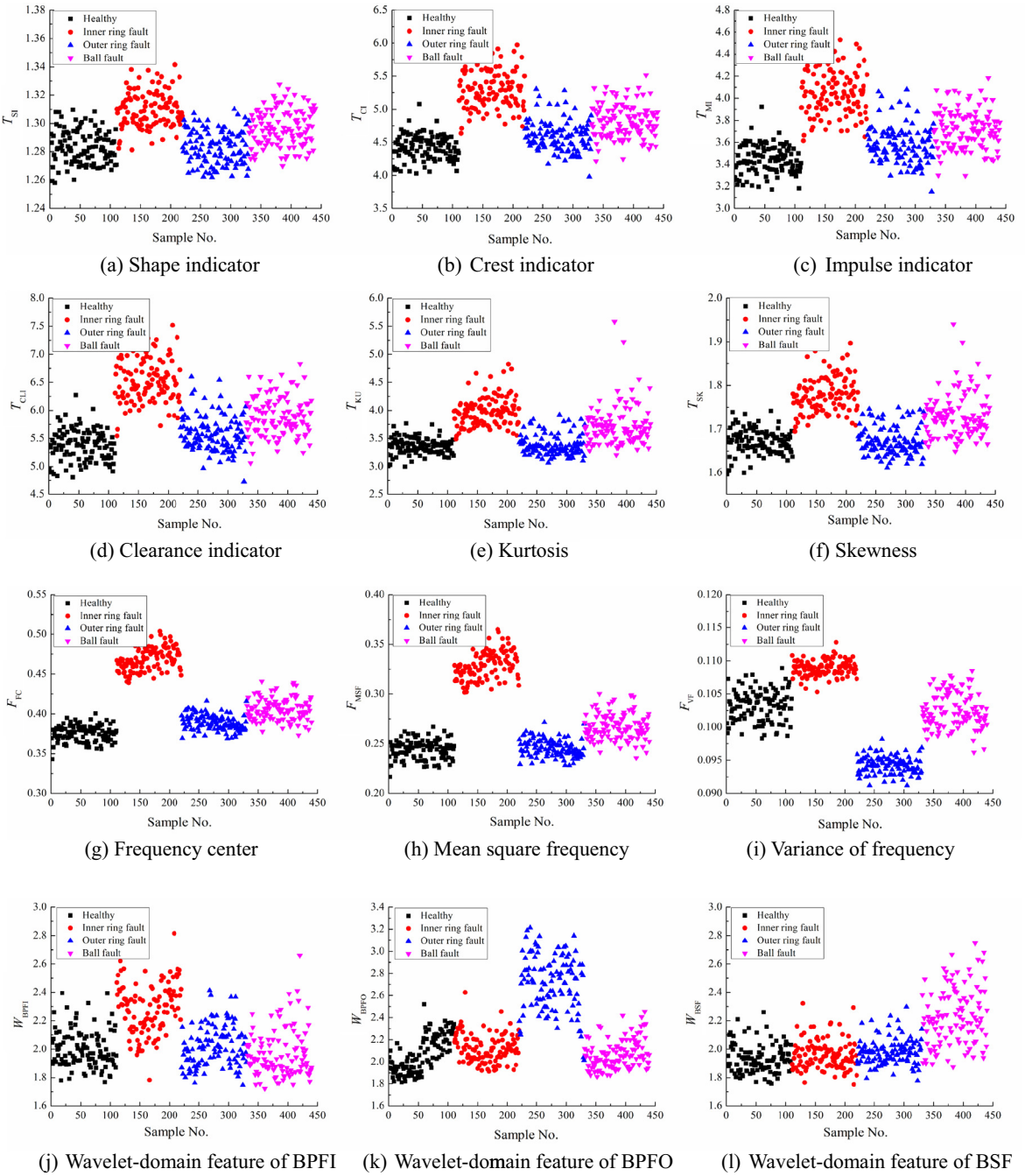


Fig. 5. Scatter plots of original features.

the calculating formula of F_{VF} contains F_{FC}, see Table 2). Generally, the correlation between the features can be represented by the correlation coefficient ρ, which is given by

$$\rho = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \cdot \sum_{i=1}^{n} y_i}{\sqrt{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \cdot \sqrt{n\sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i\right)^2}} \tag{4}$$

where x_i, y_i are the values of different features, n is sample size. The correlation matrix is shown in Table 4 and the correlation coefficient ρ of T_{SI} and F_{FC}, F_{VF} and F_{FC} are in bold.

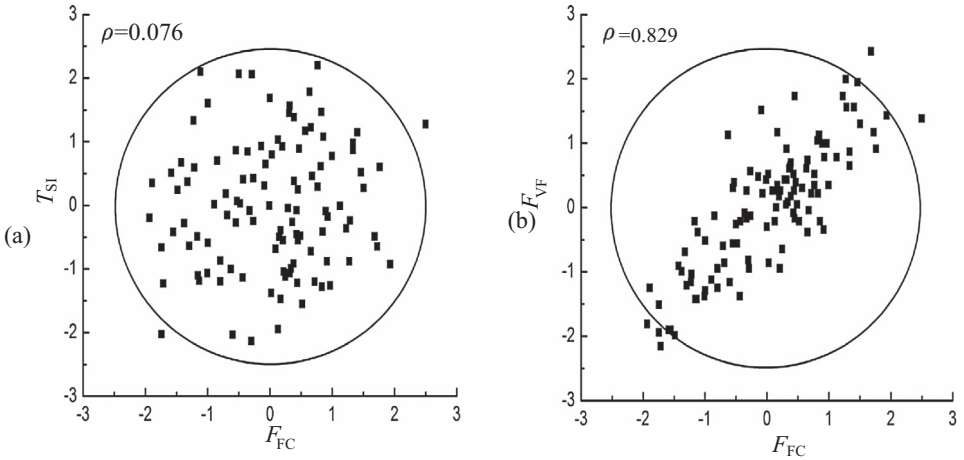


Fig. 6. The 2-D projection distribution of normalized feature vectors.

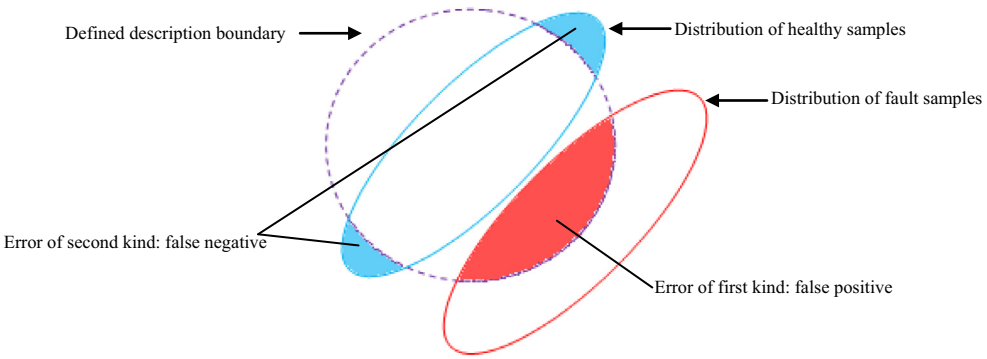


Fig. 7. Errors caused by inappropriate description of actual distribution.

Table 4
Correlation matrix of original features.

	T_{SI}	T_{CI}	T_{MI}	T_{CLI}	T_{KU}	T_{SK}	F_{FC}	F_{MSF}	F_{VF}	W_{BPFO}	W_{BPFI}	W_{BSF}
T_{SI}	1.00	0.62	0.46	0.62	0.80	0.87	**0.08**	0.09	0.12	−0.12	−0.14	0.19
T_{CI}		1.00	0.98	1.00	0.72	0.71	0.15	0.15	0.11	−0.11	−0.24	−0.03
T_{MI}			1.00	0.98	0.62	0.59	0.16	0.15	0.10	−0.09	−0.23	−0.09
T_{CLI}				1.00	0.72	0.71	0.15	0.15	0.11	−0.11	−0.24	−0.03
T_{KU}					1.00	0.98	0.13	0.12	0.08	−0.15	−0.2	0.12
T_{SK}						1.00	0.12	0.12	0.10	−0.13	−0.18	0.16
F_{FC}							1.00	0.99	**0.83**	−0.07	−0.07	−0.02
F_{MSF}								1.00	0.90	−0.06	−0.07	−0.03
F_{VF}				sym.					1.00	−0.04	−0.06	−0.08
W_{BPFO}										1.00	0.05	0.11
W_{BPFI}											1.00	0.14
W_{BSF}												1.00

There is a strong correlation between F_{FC} and F_{VF}, but the F_{FC} is more sensitive to the inner ring fault while F_{VF} is more sensitive to the outer ring fault (see Fig. 5). Namely, features with strong correlation may also contain complementary information that is helpful for fault detection.

The above conclusions of 2-D projection can be extended to the original feature space: if the correlation coefficient ρ of each pair of features is small, then the distribution of feature vectors in original feature space tends to a hypersphere, whose boundary can be described simply; otherwise, the distribution tends to a hyper-ellipsoid, whose boundary is hard to be described. However, the features with strong correlation may be chosen, because they may also contain information that is complementary for fault detection.

4. Hyper-spherical distance discrimination

In this section, the hyper-spherical distance discrimination method is proposed for rolling bearing fault detection and degradation assessment. Hyper-spherical distance discrimination consists of two parts, the hyper-spheroidization and distance discriminant analysis.

4.1. Hyper-spheroidization

According to above analysis, the distribution of feature vectors can be transformed from hyper-ellipsoid to hyper-sphere by de-correlation and normalization. This procedure is called hyper-spheroidization.

4.1.1. Eliminating the correlation

Let $\boldsymbol{x}_i$ be normalized feature vectors, then $\Sigma_i \boldsymbol{x}_i = 0$. Suppose the new coordinate system after projection transformation is $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_d\}$, where d is feature dimensionality, $\boldsymbol{w}_i$ are orthonormal basis vectors subject to $||\boldsymbol{w}_i||_2 = 1$, $\boldsymbol{w}_i^{\mathrm{T}}\boldsymbol{w}_j = 0$ $(i \neq j)$. Suppose projection vectors of $\boldsymbol{x}_i$ in new coordinate system are $\boldsymbol{z}_i = (z_{i1}, z_{i2}, \ldots, z_{id})$, where $z_{ij} = \boldsymbol{w}_j^{\mathrm{T}}\boldsymbol{x}_i$, z_{ij} is the j-th coordinate of $\boldsymbol{x}_i$ in new coordinate system. Reconstructing the $\boldsymbol{x}_i$ based on the $\boldsymbol{z}_i$, then

$$\hat{\boldsymbol{x}}_i = \sum_{j=1}^{d} z_{ij}\boldsymbol{w}_j \tag{5}$$

Taking the whole training set into consideration, the distance between $\boldsymbol{x}_i$ and $\hat{\boldsymbol{x}}_i$ is

$$\sum_{i=1}^{n}\left\|\sum_{j=1}^{d} z_{ij}\boldsymbol{w}_j - \boldsymbol{x}_i\right\|_2^2 = \sum_{i=1}^{n}\boldsymbol{z}_i^{\mathrm{T}}\boldsymbol{z}_i - 2\sum_{i=1}^{n}\boldsymbol{z}_i^{\mathrm{T}}\boldsymbol{W}^{\mathrm{T}}\boldsymbol{x}_i + \text{const} \propto -\mathrm{tr}\left(W^T\left(\sum_{i=1}^{n} x_i x_i^{\mathrm{T}}\right)W\right) \tag{6}$$

According to the criterion of minimum reconstruction error, Eq. (6) should be minimized. Considering that $\boldsymbol{w}_j$ are orthonormal basis vectors, $\Sigma_i \boldsymbol{x}_i \boldsymbol{x}_i^{\mathrm{T}}$ is covariance matrix, the objective function can be written as

$$\begin{aligned} &\min_{\boldsymbol{W}} \ -\mathrm{tr}(\boldsymbol{W}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}) \\ &\text{s.t.} \quad \boldsymbol{W}^{\mathrm{T}}\boldsymbol{W} = \boldsymbol{I} \end{aligned} \tag{7}$$

Eq. (7) can be solved by Lagrangian multiplier method.

The principle of the de-correlation method stated above is essentially the same as the principle of principal component analysis (PCA), which is well known as a feature dimensionality reduction method and have been widely used in fault diagnosis of rolling bearing [17,18]. However, the former just make a linear transformation and the whole information is preserved, while principle of principal component analysis usually pays attention to feature dimensionality reduction, which may cause information loss. The loss of information may be noise while it can also be the helpful information. Namely, reducing the dimensionality may not be conducive to classification [19]. In this study, we would like to focus on eliminating the correlation of original features instead of reducing feature dimensionality.

4.1.2. Normalization

The distribution of $\boldsymbol{z}_i$ also tends to be an ellipse, but the principal axes of ellipse are parallel to coordinate axes. This means that the distribution of $\boldsymbol{z}_i$ can be transformed to a hypersphere by normalization. The normalized $\boldsymbol{z}_i^*$ is given by

$$\boldsymbol{z}_i^* = (z_{i1}^*, z_{i2}^*, \ldots, z_{id}^*), \ z_{ij}^* = \frac{z_{ij} - \hat{\mu}_j}{\hat{\sigma}_j} \tag{8}$$

where $\hat{\mu}_j$ is the sample average of the j-th features used as an estimate of the population mean; $\hat{\sigma}_j$ is the sample standard deviation used as an estimate of the population standard deviation:

$$\hat{\mu}_j = \frac{1}{n}\sum_{i=1}^{n} z_{ij}, \ \hat{\sigma}_j^2 = \frac{1}{n-1}\sum_{i=1}^{n}(z_{ij} - \hat{\mu}_j)^2 \tag{9}$$

4.2. Boundary description

Suppose the distribution of restructured feature vectors is a hyper-sphere, then we can use a hyper-spherical boundary to describe it. What we need to do is to determine the center and radius. In this section, distance discriminant analysis method is applied for solving this problem.

Distance discriminant analysis is used to classify sample points by calculating the distance between the sample point and the population. In d-dimensional feature space, consider using a d-dimensional vector $\boldsymbol{x}_0$ to represent a sample set whose

size is n. It is hope that the quadratic sum of distances between $\boldsymbol{x}_0$ and each sample i ($i = 1, \ldots, n$) is as small as possible. Square error criterion function $J_0(\boldsymbol{x}_0)$ is defined as

$$J_0(\boldsymbol{x}_0) = \sum_{i=1}^{n} \|\boldsymbol{x}_0 - \boldsymbol{x}_i\|^2 \tag{10}$$

then

$$\boldsymbol{x}_0 = \hat{\boldsymbol{\mu}} = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{x}_i \tag{11}$$

This conclusion can be proved as follows:

$$\begin{aligned} J_0(\boldsymbol{x}_0) &= \sum_{i=1}^{n} \|(\boldsymbol{x}_0 - \hat{\boldsymbol{\mu}}) - (\boldsymbol{x}_i - \hat{\boldsymbol{\mu}})\|^2 = \sum_{i=1}^{n} \|\boldsymbol{x}_0 - \hat{\boldsymbol{\mu}}\|^2 - 2(\boldsymbol{x}_0 - \hat{\boldsymbol{\mu}})^{\mathrm{T}} \sum_{i=1}^{n} (\boldsymbol{x}_i - \hat{\boldsymbol{\mu}}) + \sum_{i=1}^{n} \|\boldsymbol{x}_i - \hat{\boldsymbol{\mu}}\|^2 \\ &= \sum_{i=1}^{n} \|\boldsymbol{x}_0 - \hat{\boldsymbol{\mu}}\|^2 + \sum_{i=1}^{n} \|\boldsymbol{x}_i - \hat{\boldsymbol{\mu}}\|^2 \end{aligned} \tag{12}$$

The second item on the right of Eq. (12) does not depend on $\boldsymbol{x}_0$, that is, the Eq. (10) reach a minimum under the condition of Eq. (11). Therefore, the sample average can represent the center of distribution. Note that the sample average of normalized feature is zero, so the distribution of $\boldsymbol{z}_i^*$ centers at the origin.

In view of that the fault samples are hard to be acquired in engineering, only the sample of healthy bearing can be used for training. According to some kind of measure criteria, the distance D_i between the feature vector and the mean vector can be calculated. Meanwhile, the threshold $D_{\max}$ is set for classification. If $D_i \leq D_{\max}$, then the i-th sample point is classified as healthy, otherwise the abnormal. The boundary described by distance discriminant analysis is a hypersphere centering at the origin of radius $D_{\max}$, as shown in Fig. 8. If the Euclidean distance is used as a measure criterion, then

$$D_i = \|\boldsymbol{z}_i^*\|_2 \tag{13}$$

According to Eq. (13), if feature values are monotonic, D_i have a specific meaning: it can reflect the degeneration degree of rolling bearing, hence D_i can be used as the evaluating indicator of rolling bearing. As shown in Fig. 9, the more serious the damage of bearing, the greater the D_i. Furthermore, multiple thresholds such as warning, abnormal, etc, can be set according to appropriate criteria. We would like to provide some perspectives in Section 6 to explain how this would be done in practice. For simplicity, only one threshold $D_{\max}$ is set in experiments to determine whether the bearing is normal.

4.3. Algorithm flow

The flow of aero-engine rolling bearing fault detection and condition monitoring is shown in Fig. 10, consisting of *model training* and *condition assessment*.

Assuming that evaluating indicator (EI) obtained from the training set follow the normal distribution, then define

$$D_{\max} = \hat{\mu}_{\mathrm{EI}} + 3\hat{\sigma}_{\mathrm{EI}} \tag{14}$$

where $\hat{\mu}_{\mathrm{EI}}$, $\hat{\sigma}_{\mathrm{EI}}$ are respectively the sample average and the sample standard deviation of EI.

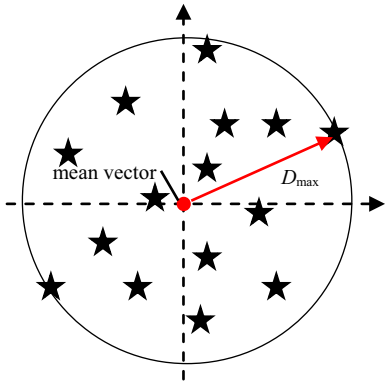


Fig. 8. Diagrammatic sketch of distance discriminant analysis.

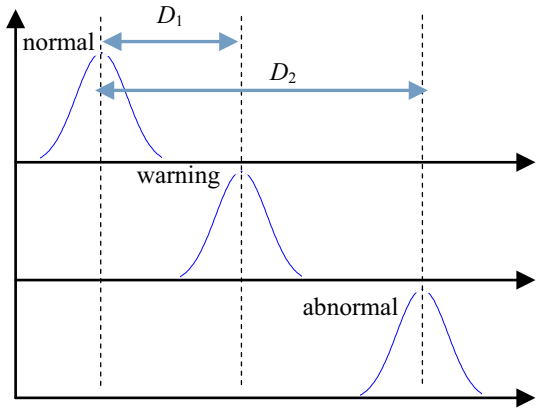


Fig. 9. Diagrammatic of bearing condition assessment.

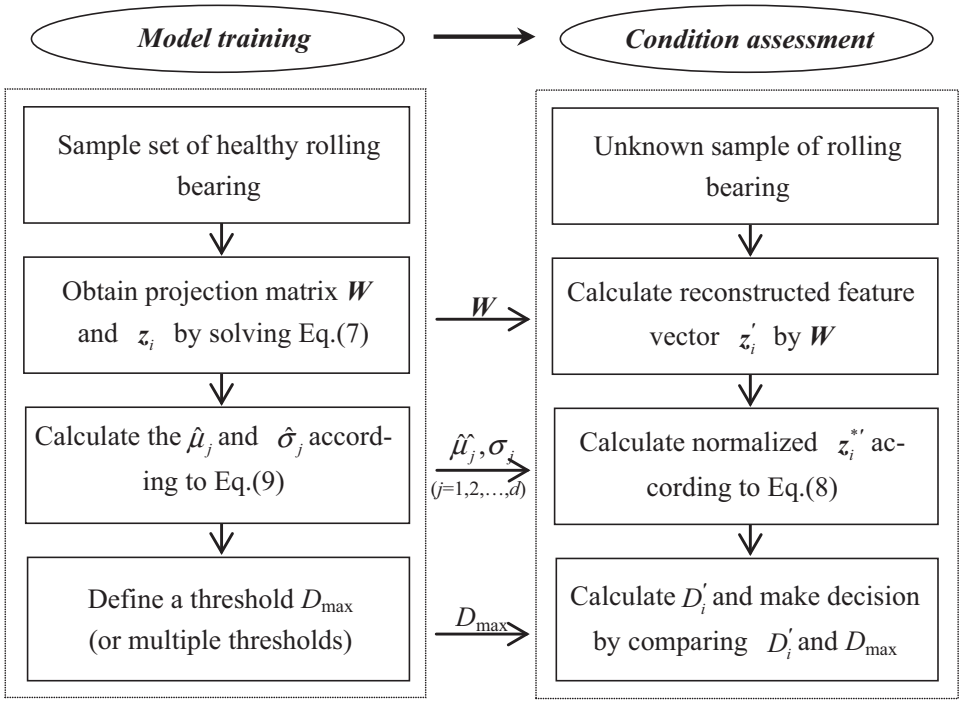


Fig. 10. Flow of aero-engine rolling bearing fault detection and condition monitoring.

5. Results and method validation

5.1. Results of experiment 1

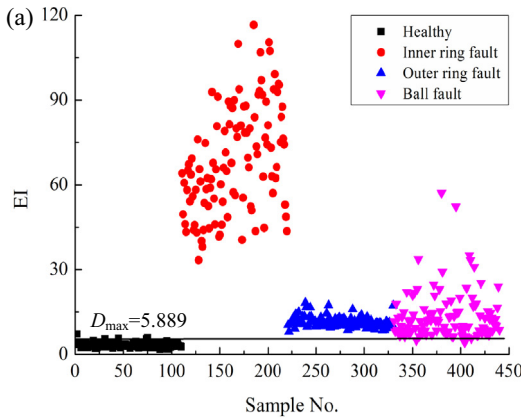
In experiment 1, HDD is compared with SVDD and SOM. For all methods, the randomly selected 50% samples of healthy bearing were used for training and all the samples (including healthy, inner ring fault, outer ring fault, ball fault) were used for testing.

For proposed method, D_i is regarded as an EI of rolling bearing. For SVDD method, decision value is regarded as the EI. For SOM method, the minimum matching distance between the feature vector and the weight vector of the nearest neurons is regarded as the EI. The SVDD is achieved by applying the LibSVM-3.18 [20] and the SOM is achieved by secondary development of the SOM-toolbox [21]. All parameters of both methods were set by 10-fold cross validation. The threshold $D_{\max}$ of each method is given by Eq. (14).

The validity of HDD is tested on data from different operating conditions and measurement points including 1500 rpm at vertical measurement point, 1500 rpm at horizontal casing measurement point and 1800 rpm at vertical measurement

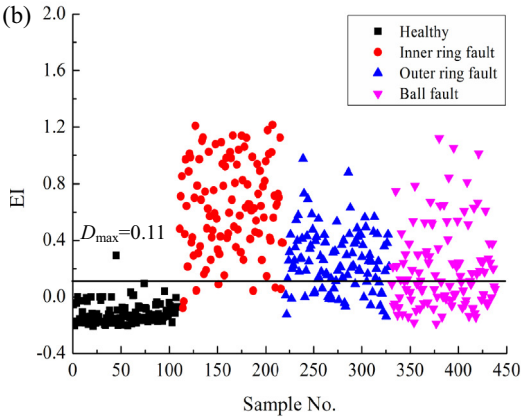
point. The results and corresponding confusion matrices of different methods are shown respectively in Figs. 11–13, where the X-axis is the sequence numbers of samples, No. 1–110 are points of healthy sample, No. 111–220 are points of inner ring fault sample, No. 221–330 are points of outer ring fault sample and No. 331–440 are points of ball fault sample. Note that the No. 1–55 samples were the randomly selected 50% samples of healthy bearing used for training. The recognition rates are shown in Tables 5–7 respectively.

Compare Figs. 11–13(a) to Fig. 5: The evaluating indicator obtained by proposed method provides a useful and automatic (sensitive to all types of faults) guidance on using the most effective information in original features. According to Figs. 11–13 and Tables 5–7, although the speeds and measurement points are changed, proposed method is always superior to SVDD and SOM.



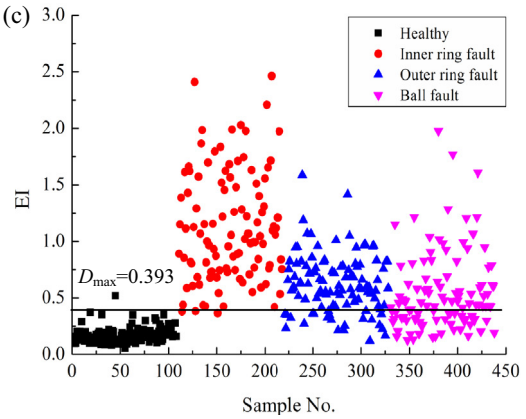
Confusion matrix

		Predicted class	
		Healthy	Abnormal
Actual class	Healthy	108 TP	2 FP
	Inner ring fault	0 FG	110 TN
	Outer ring fault	0 FG	110 TN
	Ball fault	3 FG	107 TN



Confusion matrix

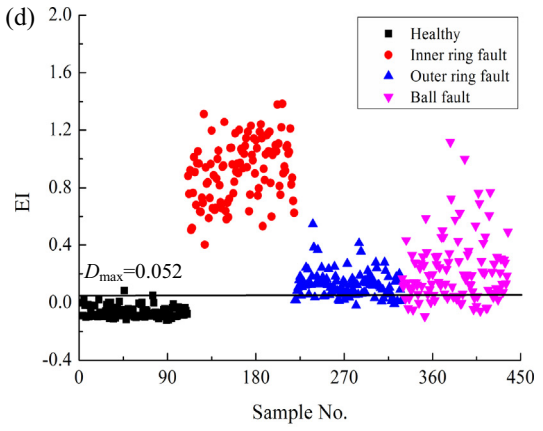
		Predicted class	
		Healthy	Abnormal
Actual class	Healthy	109 TP	1 FP
	Inner ring fault	7 FG	103 TN
	Outer ring fault	27 FG	83 TN
	Ball fault	53 FG	57 TN



Confusion matrix

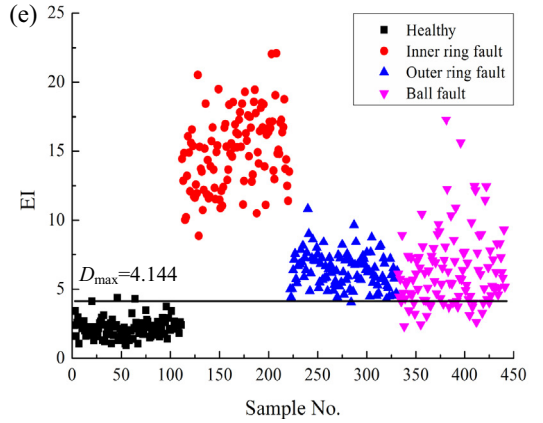
		Predicted class	
		Healthy	Abnormal
Actual class	Healthy	109 TP	1 FP
	Inner ring fault	3 FG	107 TN
	Outer ring fault	22 FG	88 TN
	Ball fault	43 FG	67 TN

Fig. 11. Results of different methods in1500 rpm at vertical measurement point: (a) HDD; (b) SVDD on original features; (c) SOM on original features; (d) SVDD on reconstructed features; (e) SOM on reconstructed features.



Confusion matrix

		Predicted class	
		Healthy	Abnormal
Actual class	Healthy	108 TP	2 FP
	Inner ring fault	0 FG	110 TN
	Outer ring fault	21 FG	89 TN
	Ball fault	25 FG	85 TN



Confusion matrix

		Predicted class	
		Healthy	Abnormal
Actual class	Healthy	108 TP	2 FP
	Inner ring fault	0 FG	110 TN
	Outer ring fault	1 FG	109 TN
	Ball fault	22 FG	88 TN

Fig. 11 (*continued*)

Furthermore, if assuming that simpler distribution is easier to be described, then when using $\boldsymbol{z}_i^*$ as input vector, the performance of SVDD and SOM should be improved. This assumption has been confirmed, as shown in Figs. 11–13 and Tables 5–7, when reconstructed features vectors are used as input vectors, the recognition rate of SVDD and SOM are improved significantly. It means that classifier performance is able to be improved by pre-processing of appropriate feature transformation.

5.2. Results of experiment 2

As mentioned in Section 3.2.2, a pre-experiment was conducted before the formal experiment to prove the stated test procedures are doable.

5.2.1. Pre-experiment

The test procedures have been stated in Section 3.2.2. In the pre-experiment, only the **Step 1** and **Step 2** are taken. The tested bearing has been operating approximately 1 h in Step 2 before the system automatically triggers protection. When the system automatically triggered protection, tested bearing was removed and decomposed.

As shown in Fig. 14, the inner ring, outer ring and rolling elements of tested bearing have a slight spall. Then the scanning electron microscope was used for further inspection, shown in Fig. 15.

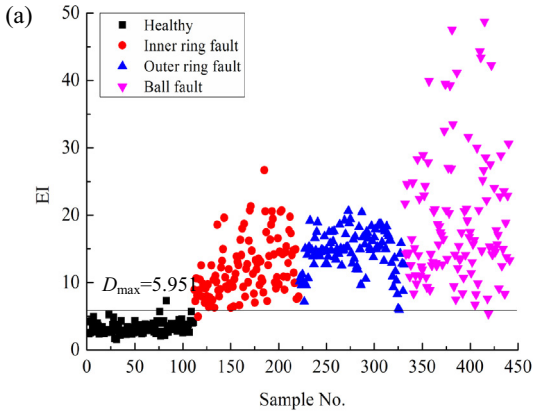
There is a slight spall on the surface of outer ring. In inner ring, the characteristics of metal adhesion can be seen. The size of small pieces of metal adhesion is approximately 20 μm. Besides, as shown in Fig. 15(f), some pits corresponding to the spall can be seen. Therefore, it can draw the conclusion that tested bearing have a slight spall after **Step 2**. Note that the spall is a common failure mode of aero-engine ball bearing.

5.2.2. Formal experiment

HDD is applied in a run-to-failure test of aero-engine ball bearing. The first 50% data collected from **Step 1** is used for training while 100% data collected from **Step 1** as well as 100% data collected from **Step 3** is used for testing. Considering that RMS of vibration acceleration signal is a widely used feature which is sensitive to the spall fault of ball bearing. The

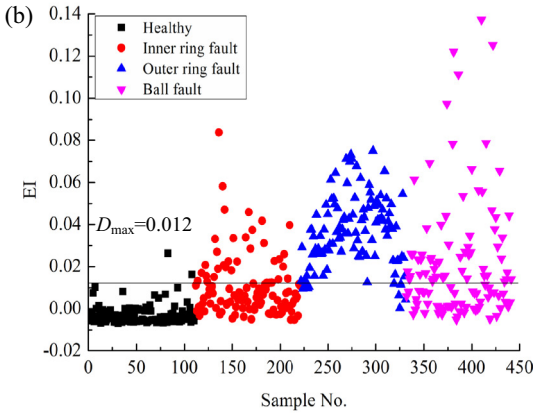
RMS is also chosen as one of the input features in this experiment, in addition to the features represented in Section 2. Meanwhile, the EI obtained from HDD is compared with the RMS, as shown in Fig. 16. The threshold D_{max} is given by Eq. (14).

As we can see in the Fig. 16, both EI and RMS are below the D_{max} during the test of **Step 1**. During the test of **Step 3**, the tendency of EI and RMS are similar, however, there is a difference between EI and RMS in the early stage of **Step 3**. The magnification figures of this stage are shown in Fig. 17. There is the volatility for both EI and RMS. However, the first time EI exceeds the threshold D_{max} is at t_1 = 8s while the RMS is at t_3 = 360 s; the last time EI below the threshold D_{max} is at t_2 = 1248 s while the RMS is at t_4 = 1600 s. In Section 5.2.1, it has been proved that there is a slight spall fault in tested bearing at the beginning of **Step 3**. Therefore, the EI obtained by proposed method is able to reflect the degradation tendency of rolling bearing, and it is also more sensitive to initial fault than the RMS.



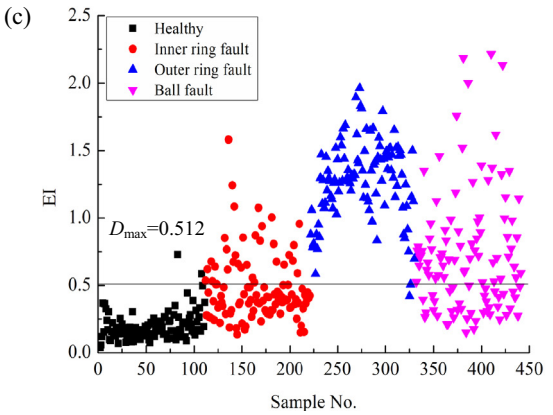
Confusion matrix

		Predicted class	
		Healthy	Abnormal
Actual class	Healthy	109 TP	1 FP
	Inner ring fault	1 FG	109 TN
	Outer ring fault	0 FG	110 TN
	Ball fault	1 FG	109 TN



Confusion matrix

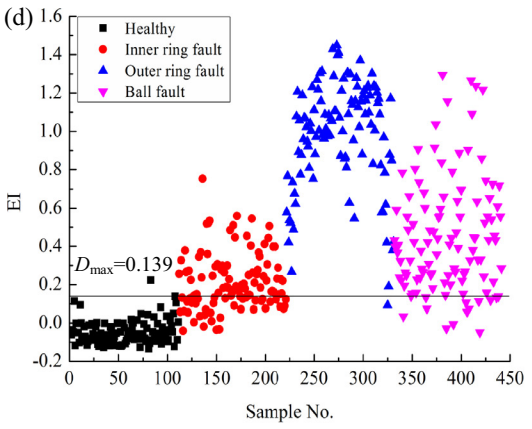
		Predicted class	
		Healthy	Abnormal
Actual class	Healthy	108 TP	2 FP
	Inner ring fault	79 FG	31 TN
	Outer ring fault	9 FG	101 TN
	Ball fault	50 FG	60 TN



Confusion matrix

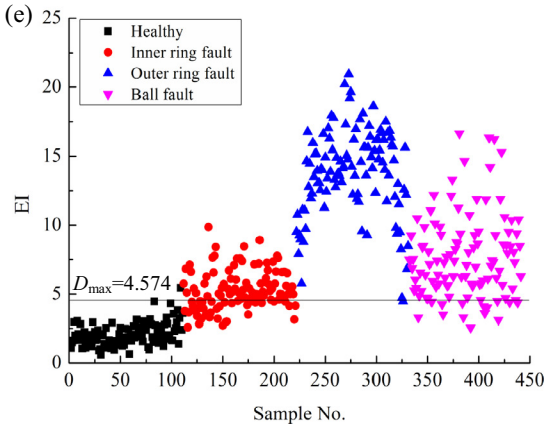
		Predicted class	
		Healthy	Abnormal
Actual class	Healthy	108 TP	2 FP
	Inner ring fault	78 FG	32 TN
	Outer ring fault	2 FG	108 TN
	Ball fault	40 FG	70 TN

Fig. 12. Results of different methods in1500 rpm at horizontal measurement point: (a) HDD; (b) SVDD on original features; (c) SOM on original features; (d) SVDD on reconstructed features; (e) SOM on reconstructed features.



Confusion matrix

		Predicted class	
		Healthy	Abnormal
Actual class	Healthy	109 TP	1 FP
	Inner ring fault	40FG	70 TN
	Outer ring fault	9 FG	101 TN
	Ball fault	13 FG	97 TN



Confusion matrix

		Predicted class	
		Healthy	Abnormal
Actual class	Healthy	109 TP	1 FP
	Inner ring fault	36 FG	74 TN
	Outer ring fault	1 FG	109 TN
	Ball fault	14 FG	96 TN

Fig. 12 (*continued*)

6. Discussion

The results of experiments show the superiority of hyper-spherical distance discrimination method. A question is that why proposed method performs well. The explanation is as follow. Firstly, it is easier to describe a hypersphere than to describe a hyperellipsoid. This is also the reason why SVDD and SOM perform better on reconstructed features than on original features. Secondly, distance discriminant analysis has an assumption that the described distribution is hypersphere while the distribution after feature transformation is indeed hyperspherical. Therefore, two kinds of errors shown in Fig. 7 can be avoided.

Hyper-spherical distance discrimination can be considered as a feature fusion method, too. Reviewing the conclusion draw in experiment 2 that the evaluating indicator obtained by proposed method is able to reflect the degradation tendency of rolling bearing, and it is also more sensitive to initial fault than RMS. An explanation is that the evaluating indicator is dominated by some features (like kurtosis, skewness, etc.) more sensitive to initial fault than RMS at the beginning. Accompanying the development of bearing fault, theses features sensitive to initial fault tend to be saturation and the evaluating indicator is dominated by RMS. Therefore, the evaluating indicator obtained by proposed method can adaptively use the most effective features for bearing degradation assessment without human intervention.

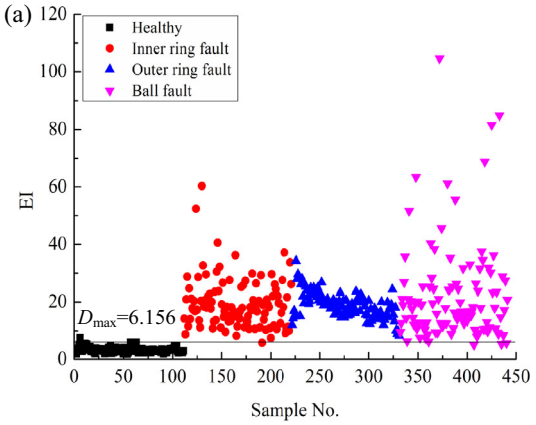
In addition, there are some interesting questions we want to make a further discussion.

6.1. Threshold

For proving the validity of proposed method, a threshold D_{max} is set according to Eq. (14). A question is that whether this threshold is appropriate, especially the thresholds of SVDD and SOM are also set by Eq. (14).

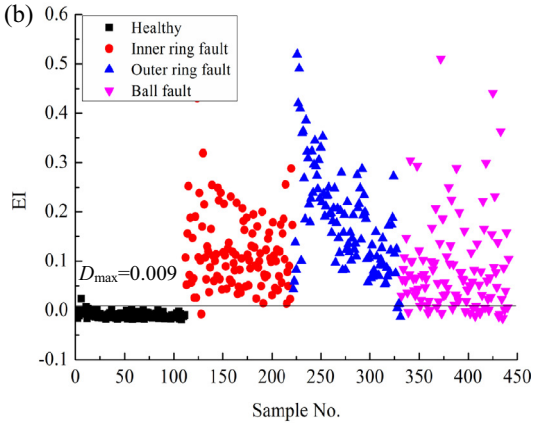
On the one hand, the reason why such a threshold setting criteria is introduced into this paper is for the convenience that proposed method can be compared to different methods quantitatively. If the probability distributions of the evaluating indicator like Fig. 18(a), no matter what the threshold is, there is always a large error. By comparison, the Fig. 18(b) is better. What should make a point is that different methods lead to the different probability distributions of the evaluating indicator, a better result like Fig. 18(b) is wanted. On the other hand, threshold indeed affects the result of the certain method a lot. If

we want to get a minimum error ratio, a good threshold like D_1 is desirable compare to D_2. However, in view of that the fault data of aero-engine rolling bearing are hard to be acquired in engineering, only the healthy bearing data can be used for setting the threshold. If the probability distributions subject to normal distribution, then $D_{max} = \mu + 3\sigma$ means approximately 99.73% samples of healthy bearing are within the range of threshold. In other word, the threshold D_{max} corresponds to a confidence of 99.73%. Similarly, $\mu + 2\sigma$ corresponds to a confidence of 95.44%, etc. Hence, multiple thresholds such as warning, abnormal, etc, can be set according to appropriate confidence. But there is also a problem that the probability distributions are unknown in practical engineering. An alternative solution is to set the thresholds based on the methods of probability density estimation, but this will increase the computation. How to set thresholds appropriately still need further study.



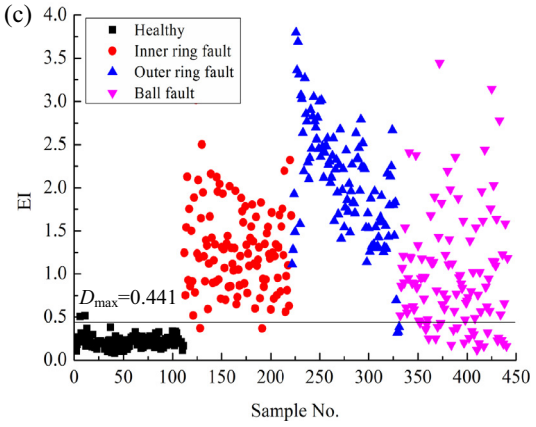
Confusion matrix

		Predicted class	
		Healthy	Abnormal
Actual class	Healthy	109 TP	1 FP
	Inner ring fault	1 FG	109 TN
	Outer ring fault	0 FG	110 TN
	Ball fault	3 FG	107 TN



Confusion matrix

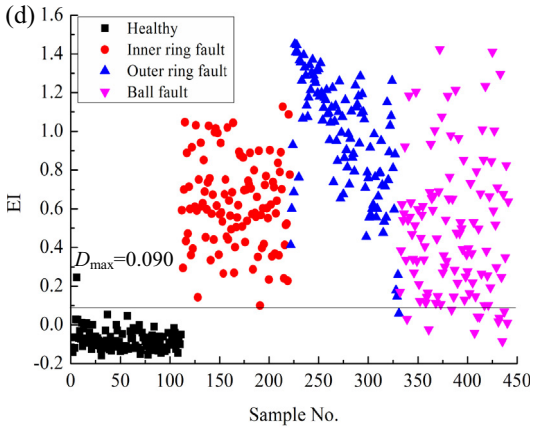
		Predicted class	
		Healthy	Abnormal
Actual class	Healthy	108 TP	2 FP
	Inner ring fault	1 FG	109 TN
	Outer ring fault	2 FG	108 TN
	Ball fault	21 FG	89 TN



Confusion matrix

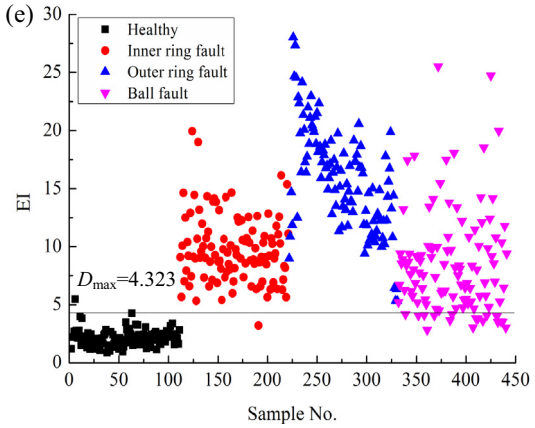
		Predicted class	
		Healthy	Abnormal
Actual class	Healthy	108 TP	2 FP
	Inner ring fault	2 FG	108 TN
	Outer ring fault	3 FG	107 TN
	Ball fault	23 FG	87 TN

Fig. 13. Results of different methods in1800 rpm at vertical measurement point: (a) HDD; (b) SVDD on original features; (c) SOM on original features; (d) SVDD on reconstructed features; (e) SOM on reconstructed features.



Confusion matrix

		Predicted class	
		Healthy	Abnormal
Actual class	Healthy	109 TP	1 FP
	Inner ring fault	0 FG	110 TN
	Outer ring fault	1 FG	109 TN
	Ball fault	11 FG	99 TN



Confusion matrix

		Predicted class	
		Healthy	Abnormal
Actual class	Healthy	109 TP	1 FP
	Inner ring fault	1 FG	109 TN
	Outer ring fault	0 FG	110 TN
	Ball fault	17 FG	93 TN

Fig. 13 (*continued*)

Table 5
Recognition rate of different methods in1500rpm at vertical measurement point.

Method	Recognition rate			
	Healthy sample	Inner fault sample	Outer fault sample	Ball fault sample
HDD	98.18%	**100%**	**100%**	**97.27%**
SVDD on original features	**99.09%**	93.64%	75.45%	51.82%
SOM on original features	**99.09%**	97.27%	80.00%	60.91%
SVDD on reconstructed features	98.18%	**100%**	80.91%	77.27%
SOM on reconstructed features	98.18%	**100%**	99.09%	80.00%

Table 6
Recognition rate of different methods in1500rpm at horizontal measurement point.

Method	Recognition rate			
	Healthy sample	Inner fault sample	Outer fault sample	Ball fault sample
HDD	**99.09%**	**99.09%**	**100%**	**99.09%**
SVDD on original features	98.18%	28.18%	91.81%	54.55%
SOM on original features	98.18%	29.09%	98.18%	63.64%
SVDD on reconstructed features	**99.09%**	63.64%	99.09%	88.18%
SOM on reconstructed features	**99.09%**	67.27%	99.09%	87.27%

6.2. Comparison with different feature preprocessing methods

Hyperspherical distance discrimination consists of two parts, the procedure of hyper-spheroidization and distance discrimination analysis. The hyper-spheroidization procedure is essentially a feature transformation method. So it is interesting

Table 7
Recognition rate of different methods in1500rpm at vertical measurement point.

Method	Recognition rate			
	Healthy sample	Inner fault sample	Outer fault sample	Ball fault sample
HDD	**99.09%**	99.09%	**100%**	**97.27%**
SVDD on original features	**99.09%**	99.09%	98.18%	80.91%
SOM on original features	98.18%	98.18%	97.27%	79.09%
SVDD on reconstructed features	**99.09%**	**100%**	99.09%	90.00%
SOM on reconstructed features	**99.09%**	99.09%	**100%**	84.55%

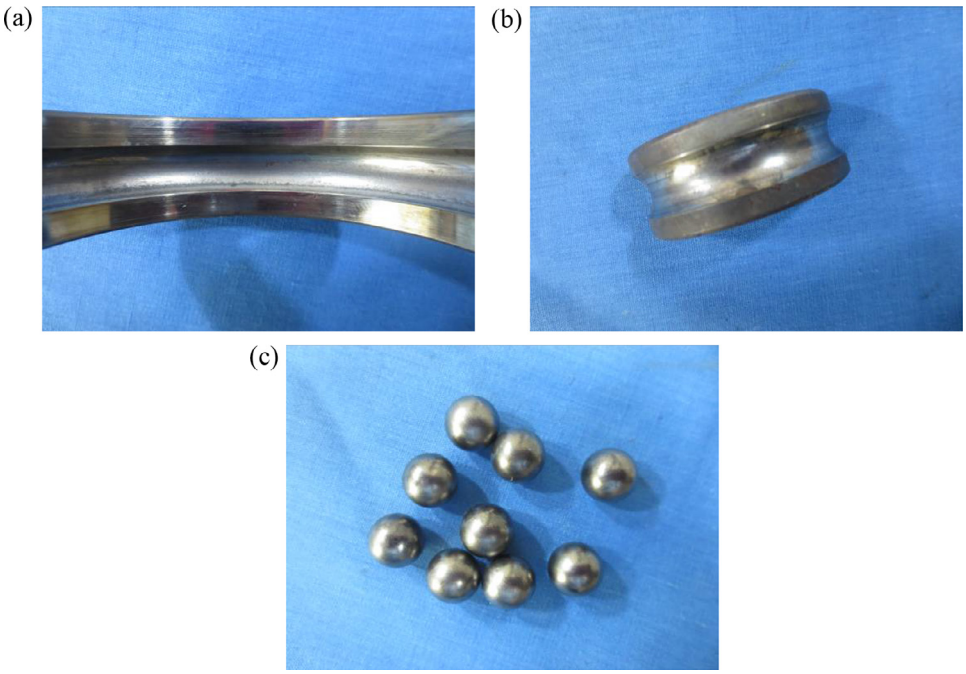


Fig. 14. Appearance of tested bearing: (a) outer ring; (b) inner ring; (c) rolling elements.

to compare hyper-spheroidization with other feature preprocessing methods such as principle component analysis and normalization.

Here is an example. There are 110 healthy samples and 110 ball fault samples from experiment 1 in 1500 rpm at vertical measurement point. Such samples are chosen because the recognition rate of ball fault is the lowest, as shown in Table 5. The randomly selected 50% healthy samples are used for training and all the samples are used for testing. Fig. 19 shows the receiver operating characteristic (ROC) curves of RBF kernel SVDD results (achieved by incremental SVDD method from dd_tools MATLAB toolboxes [22] for obtaining better performance) and SOM results under different feature preprocessing methods. All parameters of both methods are selected by cross validation. Note that the principle component analysis is just used for decorrelation instead of dimensionality reduction. Meanwhile, the results of Linear kernel SVDD is shown as a reference. A good performance of Linear-SVDD means the space distribution of reconstructed features is indeed a hypersphere, according to the principle of support vector data description.

It can be found that (1) the proposed hyper-spheroidization performs better than principle component analysis and normalization both on RBF-SVDD and SOM; (2) The performances of RBF-SVDD and SOM on original data are similar to the results after principle component analysis; (3) The performances of Linear-SVDD, RBF-SVDD and SOM are similar after hyper-spheroidization. These results are consistent with the conclusion that it is easier to describe a hypersphere than to describe a hyperellipsoid.

From the space point of view, the decorrelation is essentially a procedure of rotation. Suppose the spatial distribution of original features is a hyperellipsoid, then the principal axes of such hyperellipse will be parallel to coordinate axes after decorrelation. As mentioned in Section 4.1.1, the principle of the decorrelation method in this study is essentially the same as the principle of principal component analysis without consideration of dimensionality reduction. That is, a hyperellipsoid whose principal axes parallel to coordinate axes is obtained after principal component analysis. However, the spatial

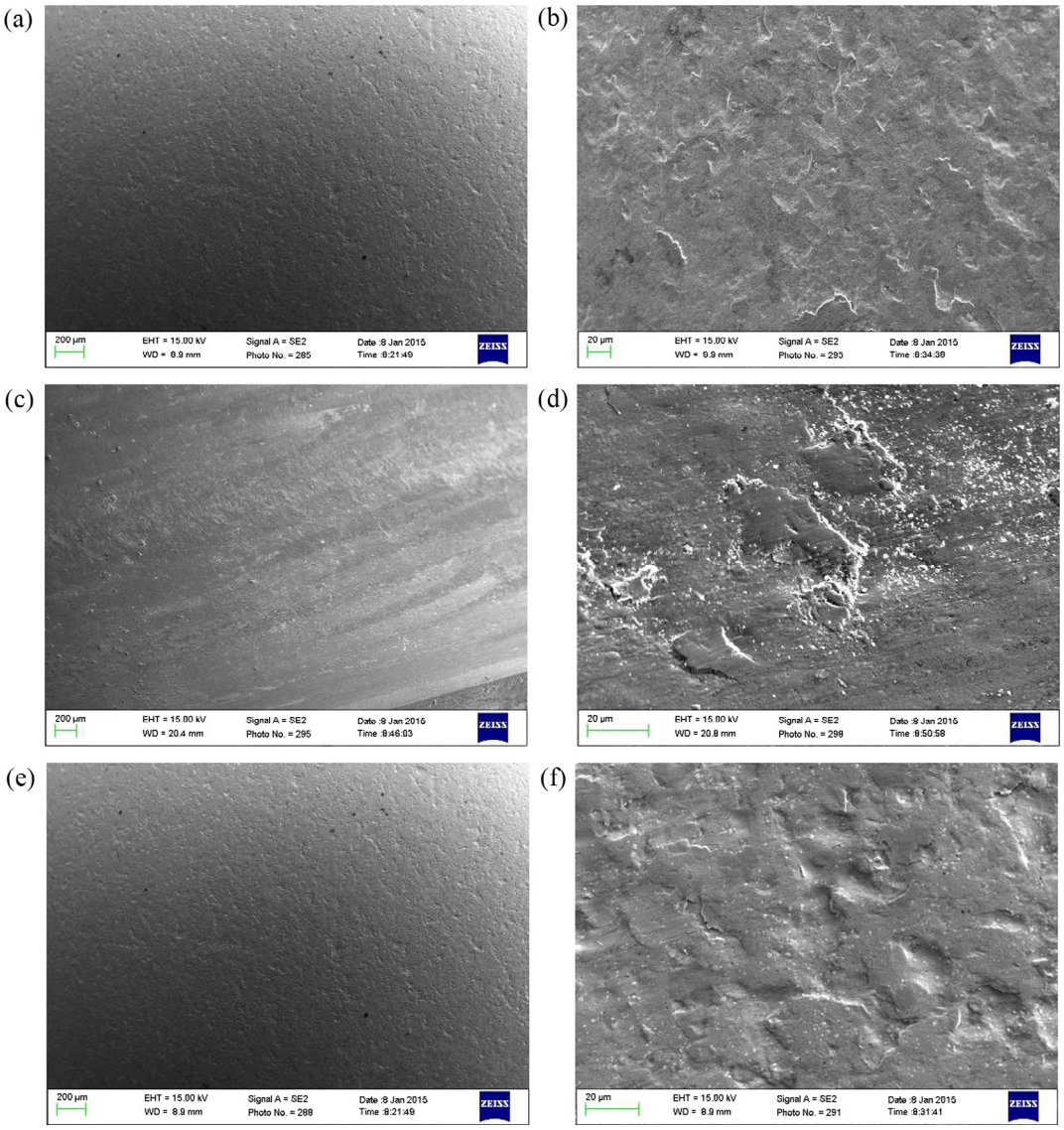


Fig. 15. Morphology of tested bearing: (a) outer ring; (b) magnification of outer ring; (c) inner ring; (d) magnification of inner ring; (e) rolling element; (f) magnification of rolling element.

distribution of rotated feature vectors is still a hyperellipsoid, and this is why the performances of RBF-SVDD and SOM on original data are similar to the results after principal component analysis. Generally, the PCA method performs well [17,18] when used for dimensionality reduction. It is because the noise in the data may be removed by dimensionality reduction while the classification will be difficult if the feature dimension is very high. Hence, only eliminating the correlation of features contributes little to data description.

The normalization is essentially a procedure of "stretching or compression". It stretches or compresses the hyperellipsoid from the directions of coordinate axes. If the distribution is a hyperellipsoid whose principal axes parallel to coordinate axes, then the distribution will tend to be a hypersphere after normalization. Otherwise, the distribution is still a hyperellipsoid after normalization. This is why hyper-spheroidization performs better than normalization. On the other hand, the normalized spatial distribution is sometimes more spherical than the original spatial distribution because the transformed features are of approximate scale. And this is why using normalization as a preprocessing method can get a better performance sometimes.

After hyper-spheroidization, even the performance of Linear-SVDD is competitive compare to the performances of RBF-SVDD and SOM. It implies that a simple data description method is accurate enough for classification in this situation. So the hyper-spherical distance discrimination (HDD) method combined with the hyper-spheroidization and distance

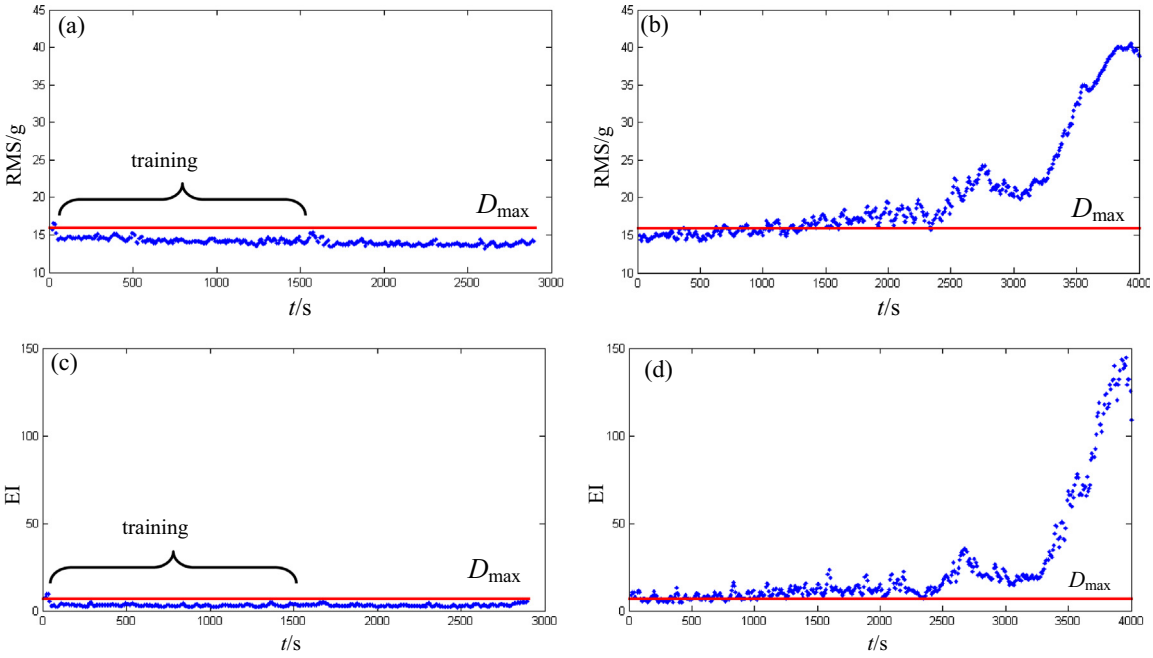


Fig. 16. The results of experiment 2: (a) RMS during Step 1; (b) RMS during Step 2; (c) EI during Step 1; (d) EI during Step 2.

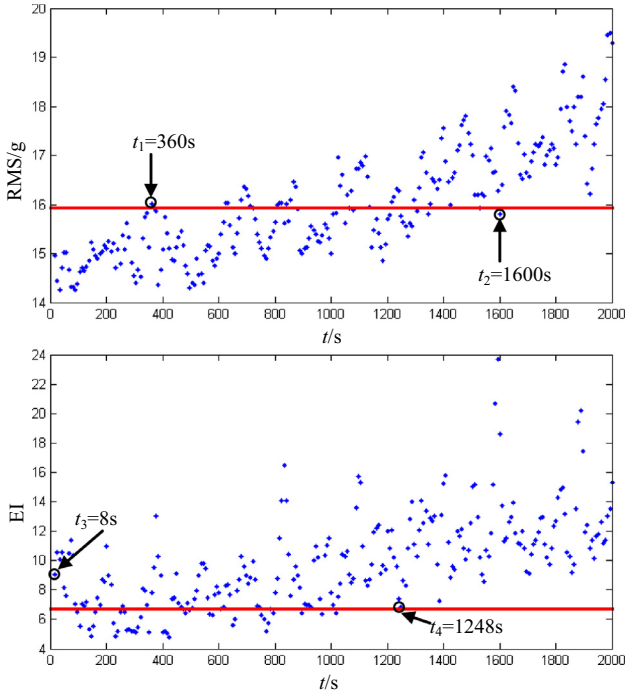


Fig. 17. The magnification figures from t = 0 s to t = 2000 s of Step 3.

discrimination analysis is proposed. By comparison, HDD is not only performs well in recognition rate, but also with the advantages of low computational complexity and no need to tuning parameters. Hence the HDD method is able to meet the requirement of aero-engine rolling bearing on-line monitoring.

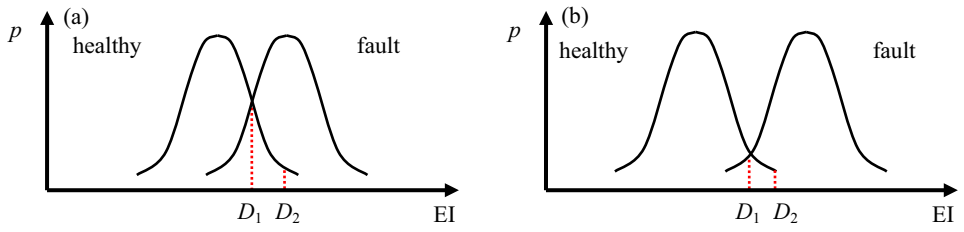


Fig. 18. Probability distributions of EI: (a) worse one; (b) better one.

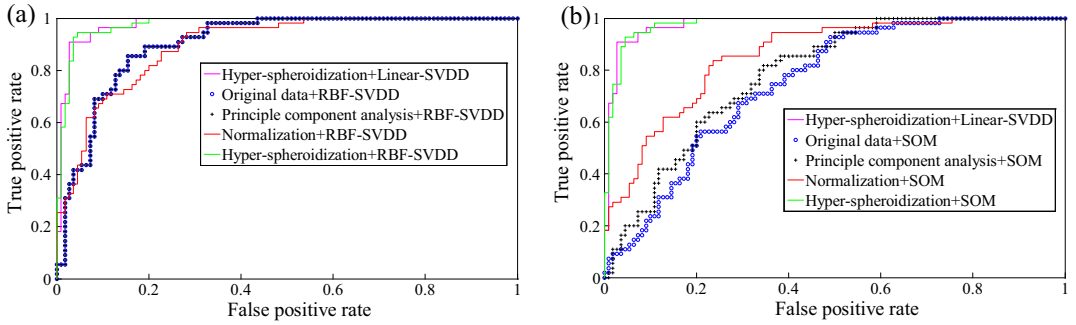


Fig. 19. ROC curves of different feature preprocessing methods: (a) on RBF-SVDD; (b) on SOM.

7. Conclusions

A novel method called HDD is proposed for aero-engine rolling bearing fault detection and on-line monitoring in this paper. HDD converts a complex boundary description problem to a 1-D threshold determination problem. Only calculating the 2-norm of reconstructed feature can it realize fault detection of rolling bearing. HDD is not only superior to the SVDD and SOM in recognition rate, but also with the advantages of low computational complexity and no need to tuning parameters. Therefore, the HDD method is able to meet the requirements of aero-engine rolling bearing fault detection. Especially, it solves the contradiction of limited computing resources in engineering and the complexity of the model. When reconstructed features vectors are used as input vectors, the recognition rate of SVDD and SOM are both improved significantly. It means that the classifier performance is able to be improved by pre-processing of hyper-spheroidization. In addition, the run-to-failure test of rolling bearing shows that the evaluating indicator obtained by proposed method is able to reflect the degradation tendency of rolling bearing, and it is also more sensitive to initial fault than RMS. Therefore, the evaluating indicator obtained by proposed method can adaptively use the most effective features for bearing degradation assessment without human intervention.

Acknowledgements

We would like to thank Mr. Zhixiang Zhao, Dr. Ke Song from Beijing Aeronautical Technology Research Center for helping in the experiments. Besides, we would like to thank the anonymous referees for valuable criticisms and useful suggestions that helped us to improve the quality of our present and future work. This study was supported by the Open Fund of the National Natural Science Foundation of China (Nos. 51675263).

References

[1] J. Lee, F. Wu, W. Zhao, et al, Prognostics and health management design for rotary machinery systems—Reviews, methodology and applications, Mech. Syst. Signal Process. 42 (1) (2014) 314–334.
[2] P.D. Mcfadden, J.D. Smith, Model for the vibration produced by a single point defect in a rolling element bearing, J. Sound Vib. 96 (1) (1984) 69–82.
[3] R.B. Randall, J. Antoni, S. Chobsaard, The relationship between spectral correlation and envelope analysis in the diagnostics of bearing faults and other cyclostationary machine signals, Mech. Syst. Signal Process. 15 (5) (2001) 945–962.
[4] R.B. Randall, J. Antoni, Rolling element bearing diagnostics—a tutorial, Mech. Syst. Sig. Process. 25 (2) (2011) 485–520.
[5] J. Yu, Bearing performance degradation assessment using locality preserving projections and Gaussian mixture models, Mech. Syst. Signal Process. 25 (7) (2011) 2573–2588.
[6] Tax, M.J. David, R.P.W. Duin, Support vector domain description, Pattern Recogn. Lett. 20 (11–13) (1999) 1191–1199.
[7] Y. Pan, J. Chen, L. Guo, Robust bearing performance degradation assessment method based on improved wavelet packet–support vector data description, Mech. Syst. Signal Process. 23 (3) (2009) 669–681.
[8] T. Benkedjouh, K. Medjaher, N. Zerhouni, et al., Fault prognostic of bearings by using support vector data description, in: Prognostics and Health Management (PHM), IEEE Conference, Colorado, 18–21 June, 2012.

[9] Z. Shen, Z. He, X. Chen, et al, A monotonic degradation assessment index of rolling bearings using fuzzy support vector data description and running time, Sensors 12 (8) (2012) 10109–10135.
[10] H. Qiu, J. Lee, J. Lin, et al, Robust performance degradation assessment methods for enhanced rolling element bearing prognostics, Adv. Eng. Inform. 17 (3) (2003) 127–140.
[11] R. Huang, L. Xi, X. Li, et al, Residual life predictions for ball bearings based on self-organizing map and back propagation neural network methods, Mech. Syst. Signal Process. 21 (1) (2007) 193–207.
[12] C. Lu, H. Yuan, Y. Tang, Bearing performance degradation assessment and prediction based on EMD and PCA-SOM, J. Vibroengin. 16 (3) (2014).
[13] G.F. Wang, Y.B. Li, Z.G. Luo, Fault classification of rolling bearing based on reconstructed phase space and Gaussian mixture model, J. Sound Vib. 323 (3–5) (2009) 1077–1089.
[14] H. Qiu, N. Eklund, H. Luoa, et al., Fusion of vibration and on-line oil debris sensors for aircraft engine bearing prognosis, in: 51st AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference 18th AIAA/ASME/AHS Adaptive Structures Conference 12th, AIAA, Florida, 12–15 April, 2010.
[15] H. Luo, H. Qiu, G. Ghanime, et al, Synthesized synchronous sampling technique for differential bearing damage detection, J. Eng. Gas Turbines Power 132 (7) (2010) 072501.
[16] G. Chen, T. Hao, H. Wang, et al, Sensitivity analysis and experimental research on ball bearing early fault diagnosis based on testing signal from casing, J. Dyn. Syst. Meas. Contr. 136 (6) (2014).
[17] A. Malhi, R.X. Gao, PCA-based feature selection scheme for machine defect classification, IEEE Trans. Instrum. Meas. 53 (6) (2005) 1517–1525.
[18] K.L. Tsui, N. Chen, Q. Zhou, et al, Prognostics and health management: a review on data driven approaches, Math. Probl. Eng. 1–17 (2015).
[19] R.O. Duda, P.E. Hart, D.G. Stork, Pattern Classification, John Wiley and Sons, 2012.
[20] C.C. Chang, C.J. Lin, LIBSVM: A library for support vector machines, Acm Trans. Intell. Syst. Technol. 2 (3) (2011) 389–396.
[21] J. Vesanto, J. Himberg, E. Alhoniemi, et al., Self-organizing map in Matlab: The SOM Toolbox, in: Proceedings of the Matlab DSP Conference, 1999.
[22] Tax, M.J. David, The Data Description Toolbox for Matlab, June, 2015, http://prlab.tudelft.nl/david-tax/dd_tools.html.