

# Safety region estimation and state identification of rolling bearing based on statistical feature extraction

Yuan Zhang<sup>a,b</sup>, Yong Qin<sup>b,\*</sup>, Zongyi Xing<sup>c</sup>, Limin Jia<sup>b</sup> and Xiaoqing Cheng<sup>b</sup>

<sup>a</sup>*School of Traffic and Transportation, Beijing Jiaotong University, Beijing, China*

<sup>b</sup>*State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China*

<sup>c</sup>*Department of Automation, Nanjing University of Science and Technology, Nanjing, Jiangsu, China*

Received 16 October 2012

Revised 3 January 2013

Accepted 14 March 2013

**Abstract.** The idea of safety region was introduced into the rolling bearing condition monitoring. The safety region estimation and the state identification of the rolling bearing operational were performed by the comprehensive utilization of Empirical Mode Decomposition (EMD), Principal Component Analysis (PCA), and the Least Square Support Vector Machine (LSSVM). The collected vibration data was segmented according to a certain time interval, and then the Intrinsic Mode Functions (IMFs) of each piece of the data were obtained by EMD. The control limits of two statistical variables extracted by PCA were presented as state characteristics. The safety region estimation for the rolling bearing operational status was performed by two-class LSSVM. The states of normal bearing, ball fault, inner race fault, and outer race fault were identified by the multi-class LSSVM. The results show that the estimation accuracy for both the safety region and the states identification reached 95%, and that the validity of the proposed method was verified.

**Keywords:** Rolling bearings, condition monitoring, safety region, empirical mode decomposition, principal component analysis, least square support vector machine

## 1. Introduction

Rolling bearings are widely used in some industries such as railway vehicle, automobile, construction machinery and so on. Nevertheless its failure rate is high. According to the statistics, only 10%~20% of the rolling bearings can achieve their design life [1]. Accurate and effective condition monitoring and identification of the rolling bearings are very important for safety, work efficiency and operating cost.

Feature extraction and state identification are key issues while dealing with rolling bearing condition monitoring problems. For feature extraction, Empirical Mode Decomposition (EMD) is a relatively new signal processing method [2]. It is a proper technique for non-stationary and non-linear signal processing such as mechanical vibration signal because of its self-adaptive and high signal-to-noise ratio [3]. And, the intrinsic mode functions (IMFs) which are obtained by EMD can be used to extract fault feature information. Currently, calculation of the energy moment [1], energy entropy [4], Renyi entropy [5], and Shannon entropy [6] of IMFs and calculation of singular

---

\*Corresponding author: Yong Qin, State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China. Tel.: +86 10 51683973; Fax: +86 10 51684081; E-mail: zhangyuan111@gmail.com.

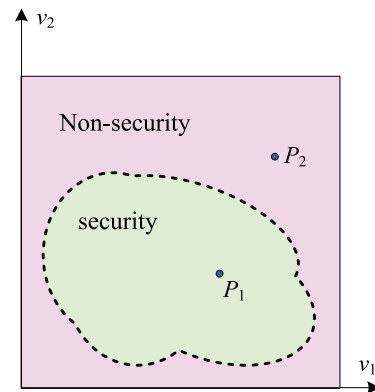


Fig. 1. Schematic diagram of safety region.

values of IMFs matrix have been used in fault feature extraction research. However the statistical features of the vibration signal, which usually contain rich operational status information, have seldom been considered in current researches. They can be used to detect the change of the object's operational status. Principal Component Analysis (PCA) is a typical multivariate statistical analysis and model recognition method. The PCA parameters have few limitations and its calculation process is simple [7]. PCA based statistical performance monitoring has many applications in quality control, process monitoring, and fault diagnosis [8–11]. In condition monitoring, after the best comprehensive simplified current status sample data based on PCA, statistical variable values, and their control limits in different subspaces are computed. For the state identification, Least Square Support Vector Machine (LSSVM) is an intelligent classification method which is an improvement of the general Support Vector Machine (SVM). LSSVM possess not only the abilities of good learning from small sample, good generalization, and avoiding trapping into local minima effectively like general SVM [12] but also a fast calculating speed with no need to specify approximate accuracy [13]. Owing to the above strengths, LSSVM is widely used in the pattern recognition and predictive control field [14–16].

Learning from the idea of region division, the method based on statistical features extraction was proposed for the safety region estimation and the operational state identification for the rolling bearings condition monitoring problem. Firstly, the state characteristic values were extracted by EMD and PCA. IMF matrixes are obtained by processing the vibration signal with EMD. The rolling bearings' state feature parameters, Hotelling's T-squared ( $T^2$ ) and Squared Prediction Error (SPE), are calculated by PCA. Then, the boundary of safety region was demarcated by two-class LSSVM, and the identification of the normal state and fault state for the rolling bearings was performed. Meanwhile, the identification of the normal state, the ball fault, the inner race fault, and the outer race fault of the rolling bearings are also performed by the multi-class LSSVM.

## 2. Safety region theory

The theory of safety region analysis and estimation was first put forward by Wu [17]. This theory was applied to safety evaluation for power systems, and then spread to network control, highway traffic, e-government affairs, etc. During the past few years, the authors' research team brought the concept of safety region into the security status monitoring and was introduced to evaluate the railway system and its key equipment using [18]. A safety assessment method was proposed for vehicle equipments based on safety region estimation.

In the research field of condition monitoring and state identification for rolling bearings, safety region is a quantitative model which fully describes an area of normal (no fault) operation of rolling bearings. Visually, a region is selected in the space determined by state characteristics, and its boundary is estimated. When the point determined by state characteristic values is in the aforementioned region, the rolling bearing operational status is normal (no fault). Otherwise the rolling bearing operational status is abnormal (fault). The schematic diagram of safety region is shown in Fig. 1 when the number of state characteristics is 2 (i.e., the state feature space is two-dimensional). In

Fig. 1,  $v_1$  and  $v_2$  are the two state characteristics,  $P_1$  and  $P_2$  are two points which represent normal and fault state of rolling bearing, respectively. The main work of safety region estimation is to get its boundary (i.e., a classification decision-making function which can distinguish between normal and fault states).

### 3. Method based on the statistical features extraction

The method based on the statistical features extraction is mainly concerned with three technologies: the EMD for signal processing, the PCA for statistical features extraction, and the LSSVM for state identification. In this section, the three technologies were described in detail, and the execution procedures of the method were given.

#### 3.1. The basic principles of EMD

EMD, proposed by Huang [19], is a method used to analyze time-frequency of signals. The IMF components and the residual component, extracted from original signal according to its local characteristic time scale, are capable of highlighting the local feature and reflecting the signal in slow variation, respectively. Hence, the characteristic information of the original signal can be accurately expressed by analyzing IMF and residual components [20].

Each IMF must meet the following two conditions: the difference between the number of zero points and the number of extreme points is less or equal to one; at any point, the average of the above envelope determined by local maximum and the below envelope determined by local minimum must be zero (i.e., the IMF has local symmetry about time axis). The sifting processes to get the components of IMF according to EMD are as follows [21]:

- 1) Sign original signal as  $x(t)$ , find out all its local extreme points, connect all the local maximum points and local minimum points respectively, then draw the above and below envelope of  $x(t)$ .
- 2) Sign  $m_1$  as the sequence which is composed of the local average of the above and below envelopes. Then

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

- 3) Determine whether  $h_1(t)$  meets the two conditions mentioned above. If not, make it the original signal and repeat step 1) and 2). That is

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

After repeating  $k$  times,

$$h_k(t) = h_{k-1}(t) - m_k \quad (3)$$

If  $h_k(t)$  meets the two conditions of IMF,  $h_k(t)$  is expressed as

$$c_1(t) = h_k(t) \quad (4)$$

Then the first IMF  $c_1(t)$  is obtained. In practice, Standard Deviation (SD) of two adjacent  $h(t)$  is usually used as a criterion to stop the iteration.  $SD$  is defined as

$$SD = \sum_{t=0}^T \frac{|h_{k-1}(t) - h_k(t)|^2}{h_k^2(t)} \quad (5)$$

where  $T$  is signal length.

- 4) Separate the IMF from the original signal, and get the original signal  $r_1(t)$

$$r_1(t) = x(t) - c_1(t) \quad (6)$$

5) Repeat Eqs (1) and (4),

$$\begin{cases} r_2(t) = r_1(t) - c_2(t) \\ r_3(t) = r_2(t) - c_3(t) \\ \vdots \\ r_n(t) = r_{n-1}(t) - c_n(t) \end{cases} \quad (7)$$

When IMF  $c_n(t)$  is less than the threshold value or  $r_n(t)$  become a monotonic function, stop the decomposition. The latter criterion is used in this research.

6) Substitute Eq. (6) into Eq. (7),

$$x(t) = \sum_{i=1}^n c_i(t) + r_n(t) \quad (8)$$

where  $r_n(t)$  is the residual component of the decomposition, representing the mean trend of the signal.

According to the above sifting processes, the original signal  $x(t)$  can be decomposed as a linear sum of  $n$  stationary IMFs  $c_i(t)$ ,  $i = 1, 2 \dots n$  and a residual component  $r_n(t)$ . The frequency components of IMFs are in a decreasing order. In other words,  $c_1(t)$  has the highest frequency while  $c_n(t)$  has the lowest. It shows that the IMF components are at different frequency bands and it's good for signal feature extraction.

### 3.2. Statistical features extraction based on PCA

PCA is a well-developed method for process monitoring and fault diagnosis. It partitions the multi-variable sample space into a low dimensional feature subspace composed of principal components and a residual subspace by multi-projector method. This partitioning is based on historical state data or process data. Two statistics which reflect the subspace changes are constructed. The observation vectors are projected into the two subspaces respectively, and the statistical indicators are calculated for state and process monitoring [22]. The feature subspace statistics reflect the main state variation, among which  $T^2$  is a commonly used. The residual subspace statistics reflect the interference and noise in state changes, and the SPE and Hawkins statistics are commonly used [23].

Based on the IMF components of vibration signals of normal and fault rolling bearings, two statistics,  $T^2$  and  $SPE$ , and their control limits are calculated. Then, the values of control limits, collected as rolling bearing's state characteristics, are used for safety region estimation and state identification. The basic PCA algorithm and the algorithm to calculate  $T^2$ ,  $SPE$  and their control limits are as follow [23,24].

- 1) Build multivariate statistical model based on data set  $\mathbf{Y}_{a \times b}$  ( $a$  is sample number and  $b$  is variable number). Standardize  $\mathbf{y} = [y_1, y_2, \dots, y_b]^T \in \mathbb{R}^b$  ( $\mathbb{R}$  represents the real number field) which is the vector of each moment in  $\mathbf{Y}_{a \times b}$  given by follow formula.

$$\bar{\mathbf{y}} = \mathbf{D}_\sigma^{-1}[\mathbf{y} - \mathbf{M}(\mathbf{y})] \quad (9)$$

where  $\mathbf{M}(\mathbf{y}) = [y_{m1}, y_{m2}, \dots, y_{mb}]^T$  is the mean vector corresponding to  $\mathbf{y}$ ,  $\mathbf{D}_\sigma = \text{diag}(s_{d1}, s_{d2}, \dots, s_{db})$  is variance matrix, and  $s_{dj} = \sqrt{E(y_j - y_{mj})^2}$  is the standard deviation for the  $j^{\text{th}}$  variable,  $j = 1, 2 \dots b$ .

Note data set after standardization as  $\bar{\mathbf{Y}}$ , then singular value decomposition is employed for correlation matrix  $\mathbf{R} = \frac{\bar{\mathbf{Y}}^T \bar{\mathbf{Y}}}{a-1}$  of  $\bar{\mathbf{Y}}$ .

$$\mathbf{R} = \mathbf{U} \mathbf{D}_\lambda \mathbf{U}^T \quad (10)$$

where  $\mathbf{U} = [u_1, u_2, \dots, u_b] \in \mathbb{R}^b$  is a unitary matrix and  $\mathbf{D}_\lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_b)$  is a diagonal matrix. The variances of  $\bar{\mathbf{Y}}$  in all directions for the new coordinate system meet  $\lambda_1 > \lambda_2 > \dots > \lambda_b$ . Call the subspace composed of the former  $d$  ( $d < b$ ) linearly independent vectors  $\mathbf{P} = [u_1, u_2, \dots, u_d]$  as principal components subspace  $\tilde{S}_P$ . The subspace composed of  $b-d$  vectors  $\mathbf{P}' = [u_{d+1}, u_{d+2}, \dots, u_b]$  is residual subspace  $\tilde{S}_E$ .  $d$  is usually calculated by the cumulative variance method. Then  $\bar{\mathbf{y}}$  can be decomposed as

$$\bar{\mathbf{y}} = \mathbf{y}_P + \mathbf{y}_E \quad (11)$$

where  $\mathbf{y}_P \in \tilde{S}_P$  and  $\mathbf{y}_E \in \tilde{S}_E$  are projections of  $\bar{\mathbf{y}}$  in  $\tilde{S}_P$  and  $\tilde{S}_E$  respectively.

- 2) Build statistic  $T^2$  and calculate its control limit in  $\tilde{S}_P$   
 $T^2$  is defined as

$$T^2 = \|\mathbf{D}_{\lambda_d}^{-0.5} \mathbf{t}\|^2 = \|\mathbf{D}_{\lambda_d}^{-0.5} \mathbf{P}^T \bar{\mathbf{y}}\|^2 = \bar{\mathbf{y}} \mathbf{P} \mathbf{D}_{\lambda_d}^{-1} \mathbf{P}^T \bar{\mathbf{y}}^T \quad (12)$$

where  $\mathbf{D}_{\lambda_d}$  is a matrix composed of the former  $d$  diagonal elements in  $\mathbf{D}_{\lambda}$  and  $\mathbf{t} = \mathbf{P}^T \bar{\mathbf{y}}$  is principal scoring vector. For vector  $\bar{\mathbf{y}}$  whose sample number is  $a$  and the number of principal components is  $b$ ,  $T^2$  follows  $F$  distribution with  $d$  and  $a-d$  degree of freedom. That is

$$\frac{a-d}{d} \cdot \frac{T^2}{a-1} \sim F(d, a-d) \quad (13)$$

where  $F(d, a-d)$  is the  $F$  distribution with  $d$  and  $a-d$  degree of freedom. Then control limit  $T_{CL}^2$  of  $T^2$  at  $\alpha$  confidence level is

$$T_{CL}^2 = \frac{d(a-1)}{a-d} \cdot F_{\alpha}(d, a-d) \quad (14)$$

where  $F_{\alpha}(d, a-d)$  is the  $100\alpha$  percentile in the  $F$  distribution, its value can be got from  $F$  distribution table.

- 3) Build statistic  $SPE$  and calculate its control limit in  $\tilde{S}_E$   
 $SPE$ , also known as the  $Q$  statistic, is defined as

$$SPE = \|\mathbf{y}_E\|^2 = (\bar{\mathbf{y}} - \bar{\mathbf{y}} \mathbf{P} \mathbf{P}^T)(\bar{\mathbf{y}} - \bar{\mathbf{y}} \mathbf{P} \mathbf{P}^T)^T = \bar{\mathbf{y}}(\mathbf{I} - \mathbf{P} \mathbf{P}^T) \bar{\mathbf{y}}^T \quad (15)$$

Accordingly  $SPE$  can be viewed as a measure of plant-model mismatch. The confidence limits for  $SPE$  are given by Jackson [25]. This test suggests the existence of an abnormal condition when  $SPE > SPE_{CL}$ , where  $SPE_{CL}$  is defined as follows:

$$SPE_{CL} = \theta_1 \left[ \frac{C_{\alpha} \sqrt{2\theta_2 h_0^2}}{\theta_1} + 1 + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} \right]^{\frac{1}{h_0}} \quad (16)$$

where,  $\theta_1 = \lambda_{d+1} + \lambda_{d+2} + \dots + \lambda_b$ ,  $\theta_2 = \lambda_{d+1}^2 + \lambda_{d+2}^2 + \dots + \lambda_b^2$ ,  $C_{\alpha}$  are the confidence limits for the  $(1-\alpha)$  percentile in a standard normal distribution, and  $h_0 = \frac{1-2\alpha\theta_3}{3\theta_2^2}$  (in which  $\theta_3 = \lambda_{d+1}^3 + \lambda_{d+2}^3 + \dots + \lambda_b^3$ ). These confidence limits are calculated based on the assumptions that the measurements are time independent and multivariate normally distributed.

In step 2) and step 3), the confidence level  $\alpha$  is set 0.95.

### 3.3. State identification based on LSSVM

The basic idea of SVM is to implement non-linear transformation by defining appropriate kernel function. The input space is transformed into a high dimensional space and then the optimal linear classification surface is calculated in the new space [26]. LSSVM is an expansion of SVM. By means of (1) using equality constraint instead of inequality constraint, and (2) making square-error and loss function as experiences loss of training set, the quadratic programming problem in SVM is transformed into a linear problem, which reduces the complexity of computation without losing accuracy [27].

For the given non-linear samples  $(g_l, o_l)$ ,  $l = 1, 2, \dots, N$ , where  $g_l$  is input data,  $o_l$  is output data, and  $N$  is the number of sample points. LSSVM can be described as the following optimization problem [28]:

$$\begin{cases} \min_{\omega, \eta} J(\omega, \varepsilon) = \frac{1}{2} \omega^T \omega + \frac{1}{2} \gamma \sum_{l=1}^N \varepsilon_l^2 \\ s.t. [\omega^T \varphi(g_l) + \eta] = 1 - \varepsilon_l; l = 1, 2, \dots, N \end{cases} \quad (17)$$

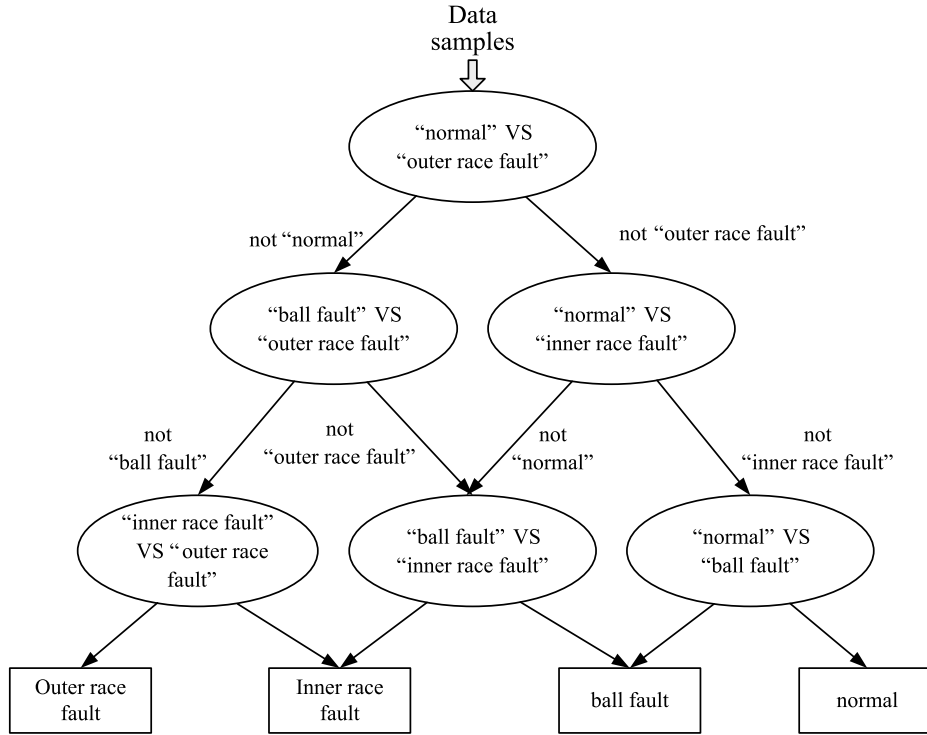


Fig. 2. DAGSVM multi-class method.

where  $J$  is objective function,  $\omega$  is weight vector,  $\eta$  is threshold,  $\varepsilon$  is slack variable, and  $\gamma$  is penalty coefficient. Sample  $g_l$  is projected into high-dimensional space by non-linear mapping  $\varphi(g_l)$ . Corresponding Lagrange function is

$$L(\omega, \eta, \varepsilon, \beta) = J(\omega, \varepsilon) - \sum_{l=1}^N \beta_l \{o_l [\omega^T \varphi(g_l) + \eta] + \varepsilon_l - 1\} \quad (18)$$

where  $\beta_l$  is Lagrange multiplier. Consider the optimal conditions

$$\left. \begin{aligned} \frac{\partial L}{\partial \omega} &= 0, \omega = \sum_{l=1}^N \beta_l \varphi(g_l) \\ \frac{\partial L}{\partial \eta} &= 0, \sum_{l=1}^N \beta_l = 0 \\ \frac{\partial L}{\partial \varepsilon_l} &= 0, \beta_l = \gamma \varepsilon_l \\ \frac{\partial L}{\partial \beta_l} &= 0, o_l [\omega^T \varphi(g_l) + \eta] + \varepsilon_l - 1 = 0 \end{aligned} \right\} l = 1, 2, \dots, N \quad (19)$$

Linear equations can be obtained as follows by eliminating  $\omega$  and  $\varepsilon$ .

$$\begin{bmatrix} 0 & \mathbf{1}^T \\ \mathbf{1} & \mathbf{K} + \gamma \mathbf{I} \end{bmatrix} \begin{bmatrix} \eta \\ \beta \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{O} \end{bmatrix} \quad (20)$$

where,  $\mathbf{1} = [1, 1, \dots, 1]_{1 \times N}^T$ ,  $\mathbf{I}$  is identity matrix,  $\mathbf{K}(g_l, g_q) = \mathbf{K}_{lq} = \varphi(g_l) \varphi(g_q)^T$  (in which  $l, q = 1, 2, \dots, N$ ),  $\beta = [\beta_1, \beta_2, \dots, \beta_N]^T$ , and  $\mathbf{O} = [o_1, o_2, \dots, o_N]^T$ . When  $h$  and  $\beta$  are available, then classification decision-making function is

$$f(g) = \text{sgn} \left[ \sum_{l=1}^N \beta_l K(g, g_l) + \eta \right] \quad (21)$$

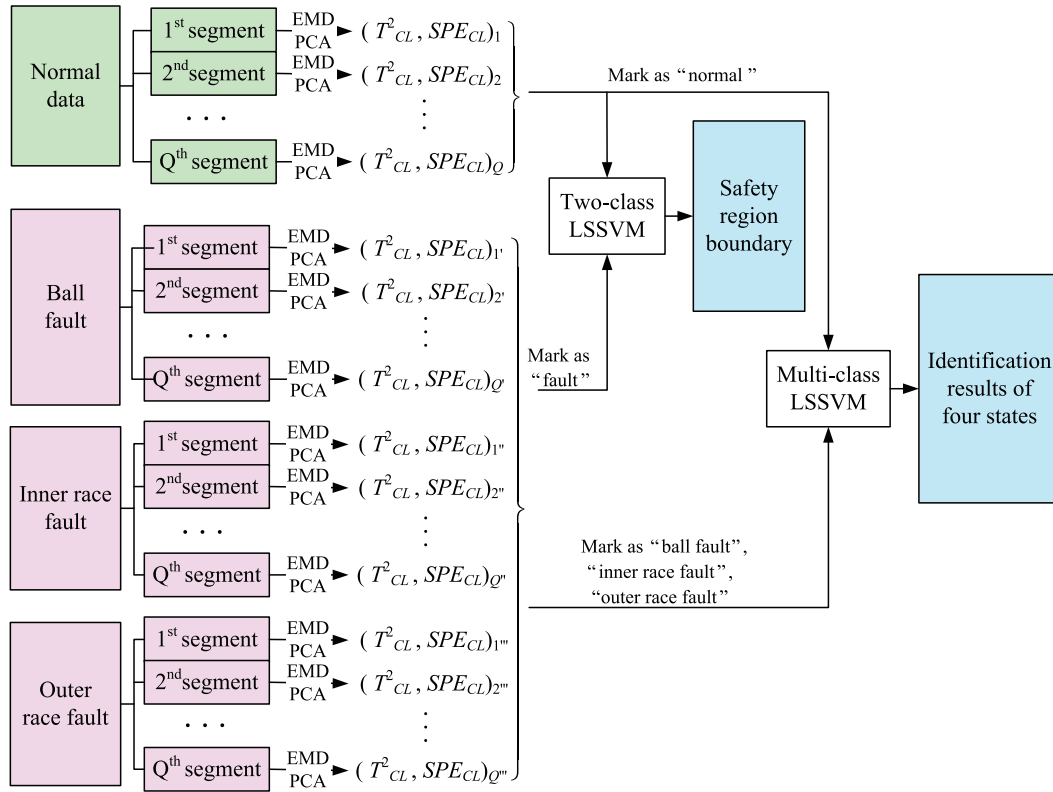


Fig. 3. The implementation process of safety region estimation and state identification.

Gauss kernel as shown in Eq. (22) behaves outstanding in many practical engineering problems, so it is selected as kernel function in this study.

$$K(g, g_l) = \exp\left(-\frac{\|g - g_l\|^2}{\rho^2}\right) \quad (22)$$

where  $\rho$  is the width of the radial basis function and  $\rho$  is 0.5 after tries.

For multi-classification problems, the multi-class SVM classifiers are constructed by combining multiple two-class SVM sub-classifiers. The common methods are one-against-one SVM, one-against-all SVM, and Directed Acyclic Graph SVM (DAGSVM). Among these methods, one-against-one SVM and DAGSVM are more suitable for practical application, but misclassifications and rejected classifications exist in one-against-one SVM [29]. So, DAGSVM is selected for the identification of rolling bearings under the conditions of normal, ball fault, inner race fault, and outer race fault.

For a classification problem containing  $M$  classes samples, DAGSVM sets up a hyper-plane between every two classes, i.e.  $M(M-1)/2$  two-class sub-classifiers are constructed. The sub-classifiers constitute a directed acyclic graph, including  $M(M-1)/2$  root nodes and  $M$  leaf nodes. In the graph, each root node represents a sub-classifier and connects with two leaf nodes of the next layer. When an unknown sample needs to be classified, the process starts from the root node (including two classes) and is continued by left or right node in the next layer until a bottom leaf is reached. The resulted leaf is the class of the unknown sample [30]. DAGSVM used for rolling bearing state identification is shown in Fig. 2.

### 3.4. Implementation procedures

The implementation procedures of safety region estimation for rolling bearings based on the proposed method are as follows.

Table 1  
Classification results for normal and fault states

	Case 1	Case 2
Testing <i>CR</i>	0.9979	0.9957
Testing <i>FK</i>	0.9943	0.9884

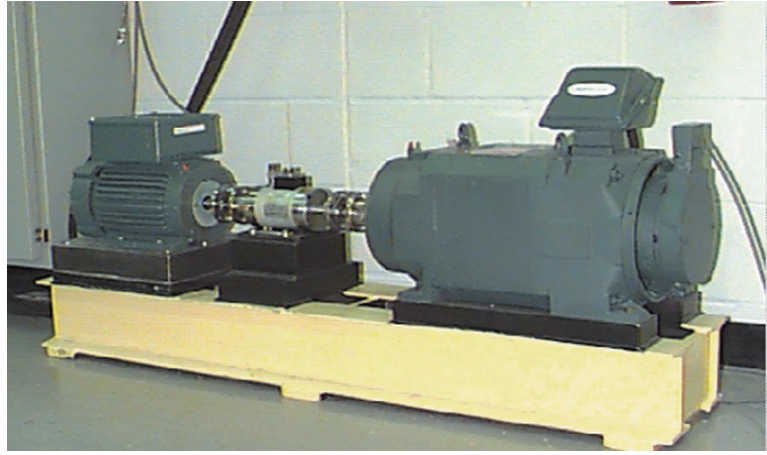


Fig. 4. The experimental apparatus of bearing fault diagnosis.

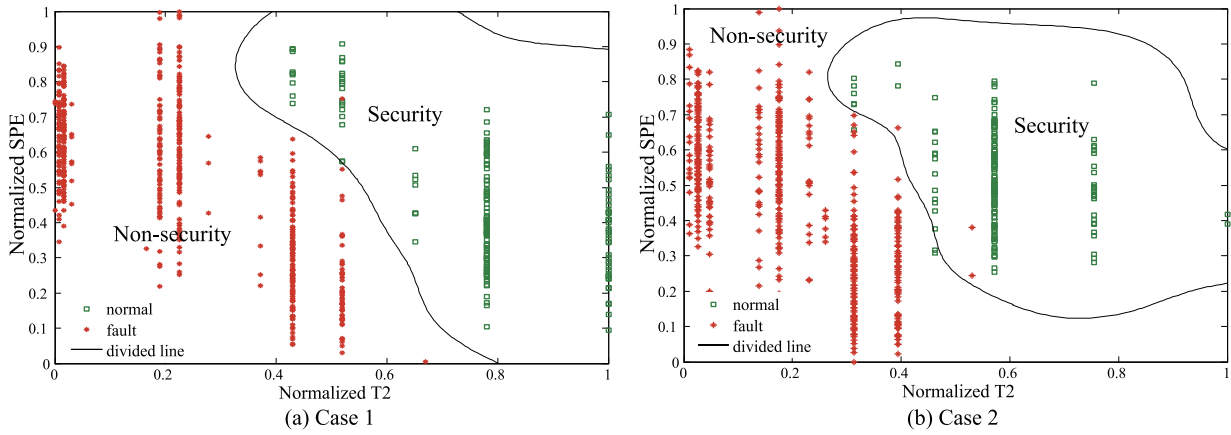


Fig. 5. Safety region estimation results for rolling bearing operational status.

- Step 1: Collect acceleration vibration signals when rolling bearing operation state is normal or fault (including ball fault, inner race fault and outer race fault), respectively.
- Step 2: Determine appropriate interval time and divide the collected data of each state into segments (for instance, the collected data is partitioned into  $T_S/t_I$  parts as its sampling time is  $T_S$  and interval time is  $t_I$ ). Each piece of data corresponds to a feature vector.
- Step 3: Process each piece of data by EMD and get their IMF matrixes.
- Step 4: Process each IMF matrix by PCA, calculate each IMF matrix's control limits of  $T^2$  and  $SPE$  statistics. And the two control limits form two-dimensional statistical feature vector of each data segment, each vector corresponds to a state point.
- Step 5: Scale the control limit data between zero and one, and then mark them as "normal" and "fault" respectively according to their operating status.
- Step 6: Classify the two-dimensional statistical feature vectors (state points) by LSSVM, and the best classified line which is the safety region boundary can be obtained.

When the rolling bearing states of normal, ball fault, inner race fault, and outer race fault are identified, acceleration vibration signals of the four states must be collected respectively. Then repeat Step 2~ Step 5, and train LSSVM sub-classifiers for every two states to obtain identified results by DAGSVM method.

The implementation process is shown in Fig. 3, where  $Q = T_S/t_I$  is the number of data segments in a specific state.



Table 2  
CR values of the sub-classifiers for every two classes of multi-class LSSVM

	CR	
	Case 1	Case 2
Normal vs outer race fault	1.0000	1.0000
Ball fault vs outer race fault	0.9871	0.9828
Inner race fault vs outer race fault	1.0000	1.0000
Normal vs ball fault	1.0000	0.9914
Normal vs inner race fault	1.0000	0.9914
Ball fault vs inner race fault	0.9485	0.9267

Table 3  
Classification results of normal, ball fault, inner race fault and out race fault

	Case 1				Case 2			
	Normal	Ball fault	Inner race fault	Outer race fault	Normal	Ball fault	Inner race fault	Outer race fault
DR	1.0000	0.9241	0.9626	0.9898	0.9966	0.9007	0.9517	0.9725
FAR	0.0023	0.0137	0.0252	0.0000	0.0034	0.0230	0.0298	0.0034
CR			0.9692				0.9553	
FK			0.9589				0.9404	

#### 4. Experiments and results

The rolling bearing experimental data was proposed by Loparo [31], and the photo of the experimental apparatus is shown in Fig. 4. The bearing type was 205-2RS JEM SKF, which was a kind of deep-groove ball bearings. The motor load was 3 horsepower and the speed was 1730 rpm. The fault diameters and the fault depths of the ball, the inner race, and the outer race were 0.1778 mm and 0.2794 mm, respectively. The number of each state of the testing bearing is one. It was a minor failure. The sampling time was 10 s. To prove the effectiveness of the proposed method comprehensively, experiments were carried out with the data of two different cases as mentioned below.

Case 1: The sampling frequency was 12 kHz with data from the drive end.

Case 2: The sampling frequency was 48 kHz with data from the fan end.

For Case 1 and Case 2, time interval  $t_I$  was determined by the rotation speed of the rolling bearing, i.e.  $t_I$  equals the time of per revolution of the rolling bearing. That is, a data segment contained the sample points collected during the time of per revolution. Then the data in Case 1 and Case 2 were both divided into 288 segments. There were 426 sample points of each segment in Case 1 and 1706 sample points of each segment in Case 2.

Then, the 288 data segments in each state (including the normal, the ball fault, the inner race fault, and the outer race fault) were decomposed by EMD to obtain the IMFs matrix of each segment. The  $T^2$  and  $SPE$  statistics control limits of each IMFs matrix were computed by PCA, and two-dimensional statistical features were extracted. Finally, LSSVM was used to achieve the safety region estimation and states identification based on the statistical features.

Additionally, input data are divided into two parts for training and testing with a proportion of 6:4 when training LSSVM.

##### 4.1. Assessment indicators

Three performance indicators, e.g., Detection Rate (DR), False Alarm Rate (FAR), and Classification Rate (CR) were used to measure the performance of proposed method [32].  $DR$  and  $FAR$  are indicators for a certain class of samples. The identification accuracy is considered better when  $DR$  is close to 1 and  $FAR$  is close to 0.  $CR$  is an indicator for all classes of samples. Classification performance is considered better when  $CR$  is close to 1.

$$DR = \frac{\text{the number of samples correctly classified for the given class}}{\text{the total number of samples in the given class}} \quad (23)$$

$$FAR = \frac{\text{the number of samples for all other classes classified as the given class}}{\text{the total number of samples in the given class}} \quad (24)$$

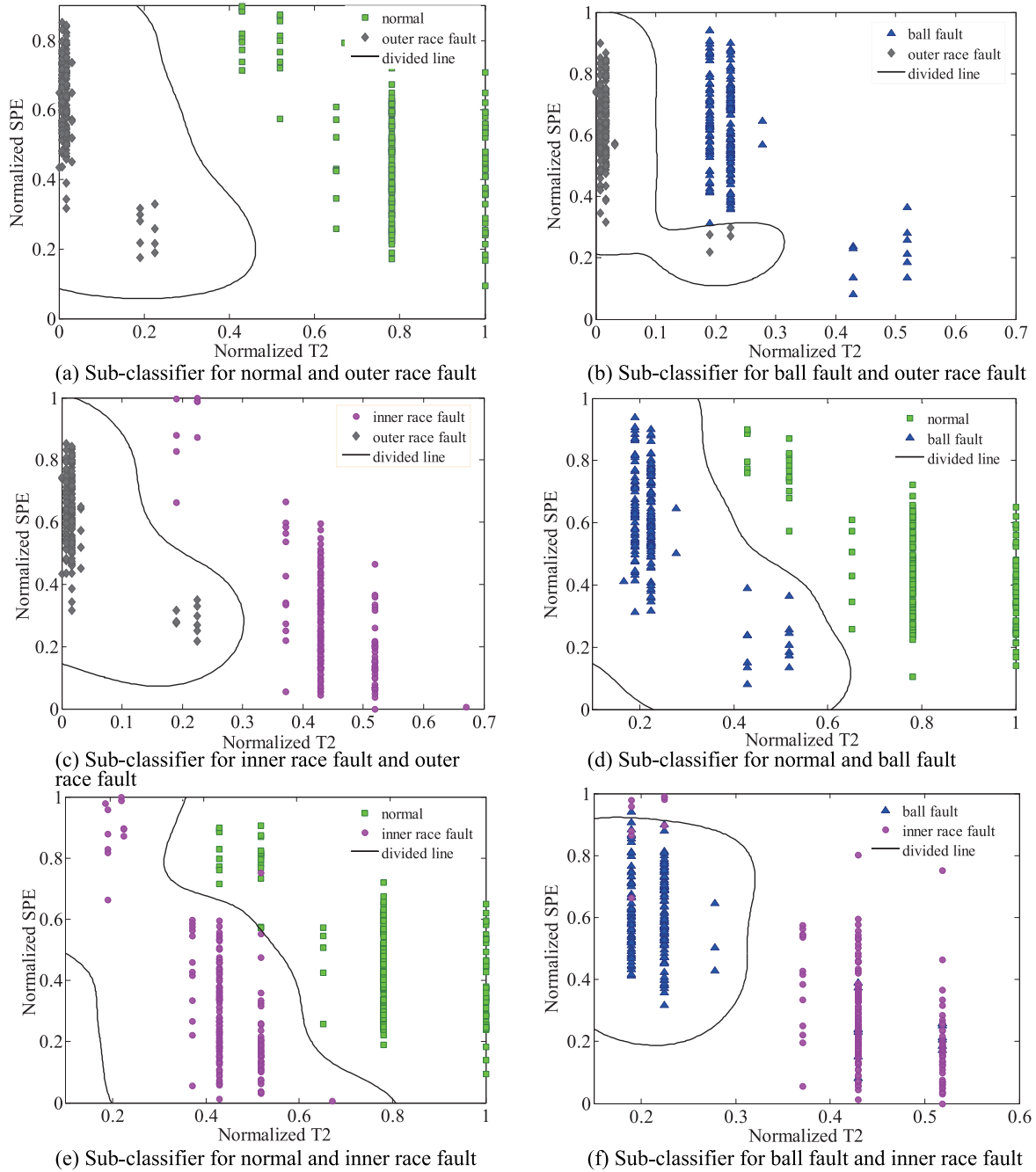


Fig. 6. Classification results of each sub-classifier in multi-class LSSVM in case 1.

$$CR = \frac{\text{the number of samples correctly classified for all classes}}{\text{the total number of samples}} \quad (25)$$

To measure the performance of state identification more comprehensively, Fleiss Kappa statistic ( $FK$ ,  $FK \in [-1, 1]$ ) is also used. It gives consistence quantitative assessment of the target of identification and the classifier outputs [33,34].  $FK$  statistic is a commonly used indicator in attribute measurement system for consistency analysis of classification data. The calculation method of  $FK$  can be found in [35]. A  $FK$  value between two groups of classification data is greater than 0.8 indicate that they belong to the same population.

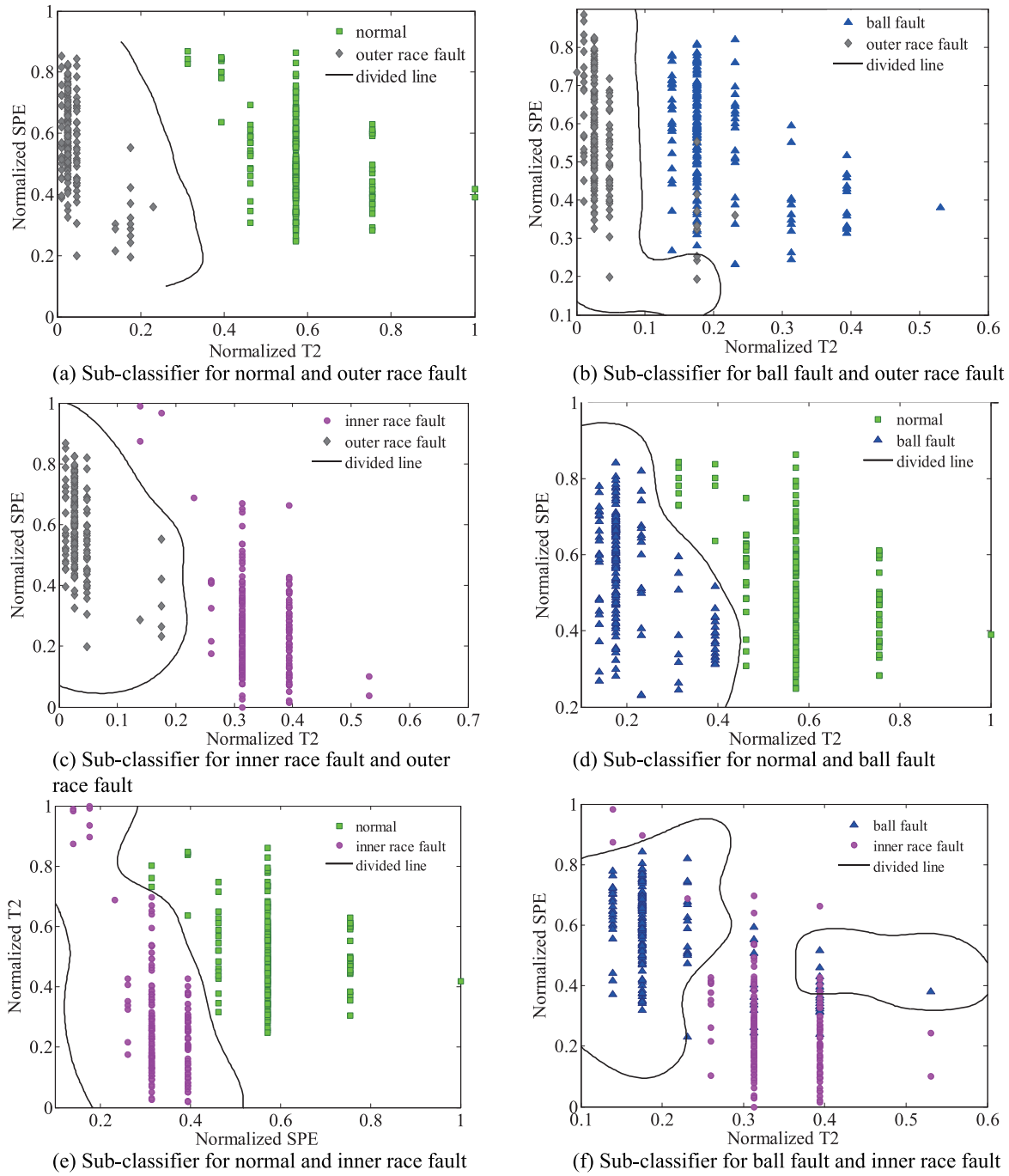


Fig. 7. Classification results of each sub-classifier in multi-class LSSVM in case 2.

#### 4.2. Experimental results

The classification results of testing data for normal and fault states are shown in Table 1. As seen in the Table 1,  $CR$  in Case 1 and Case 2 are 0.9979 and 0.9957, both of which are close to 1.  $FK$  is 0.9943 in Case 1 and 0.9884 in Case 2, both of which are greater than 0.8 and close to 1. The safety region estimation results of rolling bearing operation state are shown in Fig. 5. Figures 5(a) and (b) are the result for Case 1 and Case 2. One can see that

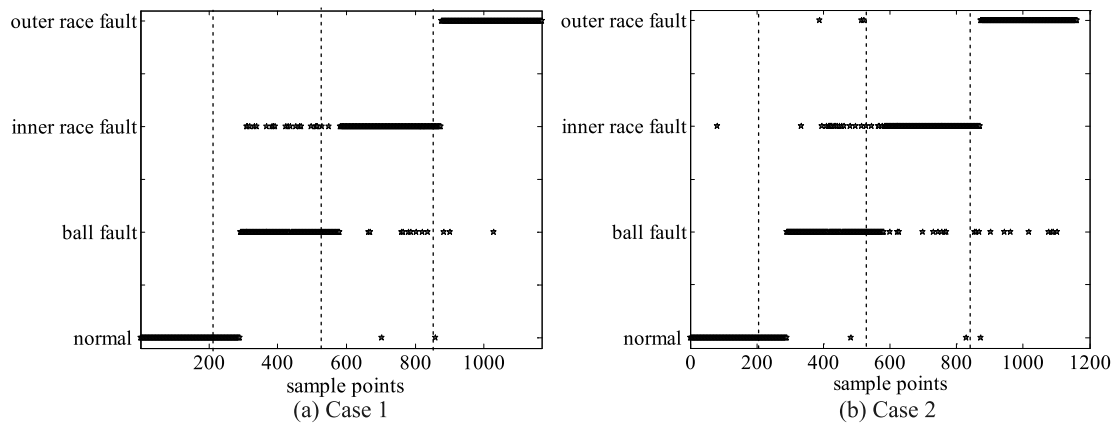


Fig. 8. Identification results of rolling bearings four states.

the dividing lines between security region and non-security region which represent normal state and fault state respectively is clear and the classification performance is good. The good accuracy of safety region classification and the high level consistency between outputs and targets prove the validity of the proposed method.

Table 2 shows the classification performance of the six sub-classifiers, which are obtained by multi-class LSSVM for the identification of every two of normal, ball fault, inner race fault, and outer race fault. It can be seen that  $CR$  values of the six sub-classifiers are higher than 0.9 for both cases. In Case 1, four  $CR$  values are equal to 1 while in Case 2 two of them reach 1 and another two are higher than 0.99. The classification results for Case 1 and Case 2 are shown in Figs 6 and 7, which show accurate classification of sub-classifiers. It indicates that the sub-classifiers have good classification performance of multi-class LSSVM.

Table 3 shows the identification results for normal, ball fault, inner race fault, and outer race fault of the rolling bearing. The  $DR$  values of the four states in both Case 1 and Case 2 are higher than 0.9. All  $FAR$  values are lower than 0.03. The  $CR$  values are 0.9692 and 0.9553 for all classes' samples in Case 1 and Case 2. The overall accuracy rate of identification exceeded 95%. The  $FK$  values in Case 1 and Case 2 are 0.9589 and 0.9404, being close to 1. The overall identification results agreed well with targets. Figure 8 shows the identification results of four states, with Fig. 8(a) for Case 1 and Fig. 8(b) for Case 2. In the figures, horizontal axis represents the number of sample points. The 1<sup>st</sup>~288<sup>th</sup> points are normal state samples, the 289<sup>th</sup>~576<sup>th</sup> points are ball fault samples, the 577<sup>th</sup>~864<sup>th</sup> points are inner race fault samples, and the 864<sup>th</sup>~1152<sup>th</sup> points are inner race fault samples. The whole area is divided into four subareas by dotted lines. Vertical axis from bottom to top represents the four states of bearing rolling. The figures shown the most samples of the four states can be correctly classified to the given class, misclassification samples are few. The results illustrated that using the statistical-features-extraction based method to identify the multiple states of rolling bearing was feasible and effective.

From Fig. 8(a) for Case 1, one can see that most misclassifications are ball fault samples which are misclassified into inner race fault. In Table 3, the ball fault has the lowest  $DR$  (0.9241) while inner race fault has the highest  $FAR$  (0.0252). In Table 2, the sub-classifier for ball fault and inner race fault has the lowest  $CR$  among the six sub-classifiers. The sub-classifier shown in Fig. 6(f) has relatively poor classification effect, which might caused by the relatively low accuracy of the sub-classifier for classifying ball fault and inner race fault. For Case 2 shown in Fig. 8(b), most misclassifications are ball fault samples misclassified into the other three classes, which resulted in the lowest  $DR$  (0.9007) in Table 3. As more samples are misclassified as ball fault and inner race fault, the ball fault and inner race fault have higher  $FAR$  than the other two states (0.0230 and 0.0298). This is because the sub-classifier for ball fault and inner race fault has the lowest  $DR$  (0.9267) and slightly lower classification efficiency, which can be found from the classification accuracy of the six sub-classifiers for Case 2 in Table 2 and the classification effect of the sub-classifiers in Fig. 7. Thus, the accuracy of ball fault and inner race fault identification need to be improved for both the two cases.

In summary, the results of safety region estimation for rolling bearing operational status and identification of normal bearing, ball fault, inner race fault and outer race fault show effectiveness and feasibility of the approach based on statistical features extraction.

## 5. Conclusions

The concept of safety region was introduced into the rolling bearing condition monitoring and the estimation methods integrated with EMD, PCA, and LSSVM were proposed based on statistical features extraction. IMF components decomposed by EMD were processed by PCA.  $T^2$  and SPE statistics control limits were proposed to represent state characteristics for rolling bearings. The safety region boundaries for the rolling bearings operational status were estimated by LSSVM based on extracted characteristics, and identification of normal and fault states was performed by the multi-class LSSVM. Experiment results show that the proposed method is effective and feasible for rolling bearing status monitoring.

The future researches will mainly concentrate on (1) how to enhance identification precision for ball fault and inner race fault. Besides, and (2) push forward the rolling bearing monitoring system based on the proposed method into practice.

## Acknowledgments

The authors would like to express their thanks to the editor and anonymous reviewers for their help in revising the manuscript. This research is sponsored by National High-tech R&D Program of China (863 Program, No. 2011AA110501) and the Specialized Research Fund for the Doctoral Program of Higher Education of China (Grant No. 20120009110035) and the State Key Laboratory of Rail Traffic Control and Safety (No. RCS2010ZZ002) of Beijing Jiaotong University. The supports are gratefully acknowledged.

## References

- [1] T.H. Qin, Y. Yang, H. Cheng and S. Xue, Rolling bearing fault diagnosis based on intrinsic mode function energy moment and bp neural network, *Journal of Vibration, Measurement & Diagnosis* **28** (2008), 229–232.
- [2] X.F. Du, Z.J. Li, F.R. Bi, J.H. Zhang, X. Wang and K. Shao, Vibration-based feature extraction of determining dynamic characteristic for engine block low vibration design, *Journal of Central South University of Technology* **19** (2012), 2238–2246.
- [3] V.K. Rai and A.R. Mohanty, Bearing fault diagnosis using FFT of intrinsic mode functions in hilbert-huang transform, *Mechanical Systems and Signal Processing* **21** (2007), 2607–2615.
- [4] Y. Yang, D.J. Yu and J.S. Cheng, A rolling fault diagnosis method based on EMD energy entropy and ANN, *Journal of Sound and Vibration* **294** (2006), 269–277.
- [5] T.Y. Wang, H.L. He, G.F. Wang, Y.G. Leng, Y.G. Xu and Q. Li, Rolling-bearings fault diagnosis based-on empirical mode decomposition and least square support vector machine, *Chinese Journal of Mechanical Engineering* **43** (2007), 88–92.
- [6] B. Li, P.L. Zhang, G.Q. Ren, D.S. Liu and S.S. Mi, Gear fault diagnosis using empirical mode decomposition genetic algorithm and support vector machine, *Journal of Vibration, Measurement and Diagnosis* **29** (2009), 88–92.
- [7] J.E. Jackson, A user's guide to principal components, Wiley, New York, 1991, pp. 30–35.
- [8] C.A. McNabb and S.J. Qin, Fault diagnosis in the feedback-invariant subspace of closed-loop systems, *Industrial and Engineering Chemistry Research* **44** (2005), 2359–2368.
- [9] L.H. Chiang, E.L. Russell and R.D. Braatz, *Fault Detection and Diagnosis in Industry System* Springer-Verlag, London, 2001, pp. 15–18.
- [10] R. Sharmin, S.L. Shah and U. Sundararaj, A PCA based fault detection scheme for an industrial high pressure polyethylene reactor, *Macromolecular Reaction Engineering* **2**(1) (2008), 12–30.
- [11] E.P.D. Moura, C.R. Souto and A.A. Silva, Evaluation of principal component analysis and neural network performance for bearing fault diagnosis from vibration signal processed by RS and DF analyses, *Mechanical Systems and Signal Processing* **25** (2011), 1765–1772.
- [12] K. Gopalakrishnan and S. Kim, Support vector machines approach to hma stiffness prediction, *Journal of Engineering Mechanics* **137**(2) (2011), 138–146.
- [13] M.T. Gencoglu and M. Uyar, Prediction of flashover voltage of insulators using least squares support vector machines, *Expert Systems with Applications* **36** (2009), 10789–10798.
- [14] L.M. Liu, A.N. Wang, M. Sha and F.Y. Zhao, Multi-class classification methods of cost-conscious ls-svm for fault diagnosis of blast furnace, *Journal of Iron and Steel Research, International* **18**(10) (2011), 17–23, 33.
- [15] S. Ari, K. Hembram and G. Saha, Detection of cardiac abnormality from PCG signal using LMS based least square SVM classifier, *Expert Systems with Applications* **37**(12) (2010), 8019–8026.
- [16] X. Li, G.Y. Cao and X.J. Zhu, Modeling and control of PEMFC based on least squares support vector machines, *Energy Conversion and Management* **47**(7–8) (2006), 1032–1050.
- [17] F. Wu and S. Kumagai, Steady-state security regions of power systems, *IEEE Transactions on Circuits and Systems* **29**(11) (1982), 703–711.

- [18] Y. Zhang, Y. Qin and L.M. Jia, Research on methodology of security region estimation of railway system operation safety assessment, *Proceedings of World Congress on Engineering and Technology* **6** (2011), 803–807.
- [19] N.E. Huang, Z. Shen and S.R. Long, The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis, *Proc R Soc Lond A* **454**(1971) (1998), 903–995.
- [20] Y.L. Xu and J. Chen, Characterizing nonstationary wind speed using empirical mode decomposition, *Journal of Structural Engineering* **130**(6) (2004), 912–920.
- [21] P.W. Tse, Y.H. Peng and R. Yam, Wavelet analysis and its envelope detection for rolling element bearing fault diagnosis their affectivities and flexibilities, *Journal of Vibration and Acoustic* **123** (2001), 303–310.
- [22] D.H. Zhou and Y.Y. Hu, Fault diagnosis techniques for dynamic systems, *ACTA Automatica Sinica* **35** (2009), 748–758.
- [23] S.J. Qin, Statistical process monitoring: Basics and beyond, *Journal of Chemometrics* **17**(8–9) (2003), 480–502.
- [24] A.J. Richard and W.W. Dean, Applied multivariate statistical analysis (sixth edition), *Prentice Hall* New Jersey (2007), 20–34.
- [25] M.A. Bin Shams, H.M. Budmann and T.A. Duever, Fault detection, identification and diagnosis using CUSUM based PCA, *Chemical Engineering Science* **66** (2011), 4488–4498.
- [26] V.N. Vapnik, The nature of statistical learning theory, Springer-Verlag, New York, 1999, pp. 58–70.
- [27] J.A.K. Suykens, Nonlinear modeling and support vector machines, *IEEE Instrumentation and Measurement Technology Conference Budapest* (2001), 287–294.
- [28] H.S. Tang, S.T. Xue, R. Chen and S. Tadanobu, Online weighted LS-SVM for hysteretic structural system identification, *Engineering Structures* **28**(12) (2006), 1728–1735.
- [29] W.H. Chih and J.L. Chih, A comparison of methods for multiclass support vector machines, *IEEE Transactions on Neural Networks* **13**(2) (2002), 415–425.
- [30] J.C. Platt, N. Cristianini and J. Shawe-Taylor, Large margin DAG's for multiclass classification, *Advances in Neural Information Processing Systems* **12** (2000), 547–553.
- [31] The case western reserve university bearing data center website, bearing data center seeded fault test data, <<http://csegroups.case.edu/bearingdatacenter/pages/download-data-file>>.
- [32] X.M. Tao, J. Xu, L.B. Yang and Y. Liu, Bearing fault diagnosis with a MSVM based on a garch model, *Journal of Vibration and Shock* **29**(5) (2010), 11–15.
- [33] M. Boumahdi, J.P. Dron, S. Rechak and O. Cousinard, On the extraction of rules in the identification of bearing defects in rotating machinery using decision tree, *Expert Systems with Applications* **37** (2010), 5887–5894.
- [34] P.K. Kankar, S.C. Sharma and S.P. Harsha, Rolling element bearing fault diagnosis using autocorrelation and continuous wavelet transform, *Journal of Vibration and Control* **17**(14) (2011), 2081–2094.
- [35] J. Sim and C.C. Wright, The kappa statistic in reliability studies: Use, interpretation, and sample size requirements, *Physical Therapy* **95**(5) (2005), 257–268.

