

# Health Status Identification of Connecting Rod Bearing Based on Support Vector Machine

Yongbin Liu<sup>1</sup>, Qingbo He<sup>1</sup>, Ping Zhang<sup>1</sup>, Zhongkui Zhu<sup>2</sup>, Fanrang Kong<sup>1</sup>

1. Department of Precision Machinery and Precision Instrumentation,

University of Science and Technology of China, Hefei, Anhui 230027, China

2. School of Urban Rail Transportation, Soochow University, Suzhou, Jiangsu 215021, China  
lyb@ustc.edu.cn

**Abstract:** Connecting rod bearing (CRB) is an important component which joins the reciprocating and rotating movements together in the internal combustion engine (ICE). It is very difficult to identify health status of CRB because of variable working process, complicated excitation and distribution sources, and lack of fault samples. Support vector machine (SVM), which has excellent capability in small data case, was introduced to identify the health status of CRB. In this paper, faults of the CRB were simulated in an ICE with the type of EQ6100. Vibration features were extracted from vibration signals acquired from the shell of ICE. And a SVM multi-classifier was designed to identify health status of CRB by using the radial basis kernel function. Experimental results indicated that the presented fault diagnosis method could effectively recognize different conditions of CRB.

**Keywords:** Support Vector Machine; Internal Combustion Engine; Connecting Rod Bearing; Fault Diagnosis

## 1. Introduction

Internal combustion engine (ICE), a complicated mechanical system, is composed of different parts, in which connecting rod bearing (CRB) is an important one joining the reciprocating and rotating movements together. It is difficult to identify the health status of CRB for several reasons, such as variable working process, complicated excitation and distribution sources, and lack of fault samples [1,2]. One crucial task in CRB fault diagnosis is the feature classification based on vibration signal. For this problem, methods based data learning have been investigated for fault diagnosis, such as neural network [3], Bayesian classification [4], etc. However, the vibration samples that represent the condition are difficult to collect in practice, which make the methods mentioned above are limited to being used. Therefore, it is critical to select an excellent classifier for reliable fault diagnosis.

Support vector machine (SVM), a machine learning method based on the statistical learning theory, was presented by Vapnik et al. [5-7]. The SVM implements the structural risk minimization principle which leads to the preferable generalization capability. Kernel function is used to map a sample input space into a high-dimensional feature space through some nonlinear mapping. With SVMs, nonlinear dependence between factors and objects can be described. Because of the excellent capability in small data case and in limiting over learning, SVM has been used in many applications, such as pattern recognition [8], regressive analysis [7], etc. Especially, it shows advantages in fault diagnosis as described by the referred studies [9-12]. In this paper, we address extraction vibration signal features of connecting rod bearing, and health status identification of connecting rod bearing using the SVM. The organization of this paper is arranged as follows. The theoretical background of the proposed method is introduced in Section 2. The feature extraction method and feature analysis are described in Section 3. The application example is presented in Section 4. Finally, conclusions are drawn in Section 5.

## 2. Support Vector Machine

SVM is a machine learning tool that is especially fit for classification in small-sample cases due to its good generalization capability. The basic idea of SVM for pattern recognition is to map data to a high-dimensional space by a nonlinear mapping, and find an optimum linear separating hyperplane with the maximal margin in this higher-dimensional space [6,7]. However, the calculation complexity does not increase almost in the high-dimensional space as a kernel function is used. “one-against-one” and “one-against-all” are the most commonly used methods of SVM classifier [8]. Multi-classification can be obtained by combining several binary classifiers. Considering the need of recognizing multiple statuses in fault diagnosis of ICE, one-against-all multi-class classifier is constructed in this paper and described in detail as follows [5].

Given a training set  $(x_i, y_i), i=1, 2, \dots, l$ , where  $l$  is the number of samples,  $x_i \in R^n$  is input data and  $y_i = \{-1, +1\}$  is output class data. Finding an optimum separating hyperplane is equivalent to solving the following quadratic programming problem:

$$\begin{aligned} \min_{w, b, \xi} \quad & \frac{1}{2} w^T w + C \sum_{i=1}^l \xi_i \\ \text{s.t.} \quad & \begin{cases} y_i (w^T \Phi(x_i) + b) \geq 1 - \xi_i, \\ \xi_i \geq 0. \end{cases} \end{aligned} \quad (1)$$

where  $w \in R^n$  and  $b \in R$  are the weighting factors,  $\xi_i$  is the slack variable, and  $C$  is the penalty parameter. Input vector  $\mathbf{x} = (x_1, x_2, \dots, x_l)$  in original space can be mapped to

high-dimensional feature space using the function  $\Phi(\cdot)$ . In feature space, optimum separating hyperplane can be found by maximizing the margin  $2//w//$ . The decision function can be finally written as:

$$\text{sgn} \left( \sum_{i=1}^l y_i \alpha_i \Phi(x_i)^T \Phi(x_j) + b \right) \quad (2)$$

$$S.t. \begin{cases} 0 \leq \alpha_i \leq C, & i = 1, 2, \dots, l; \\ \sum_{i=1}^l \alpha_i y_i = 0. \end{cases}$$

where the project function  $\Phi(x_i)$  could be conveniently replaced by a kernel function  $K(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j)$ . Then the decision function can be rewritten as:

$$\text{sgn} \left( \sum_{i=1}^l y_i \alpha_i K(x_i, x_j) + b \right) \quad (3)$$

There are several commonly used kernel functions: linear function, polynomial function, and radial basis function (RBF). The radial basis function was described as follows:

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2) \quad (4)$$

where  $\gamma > 0$ , and  $\gamma$  is the kernel parameter.

For classable  $k$ -class statuses of connecting rod bearing of internal combustion engine, a multi-class classifier can be constructed to identify health statuses of CRB using  $k$  classifiers, each of which is trained by binary-class data as above. To identify the health status of CRB, the feature vectors of a sample are firstly sent into the classifier 1. If the output of the decision function is 1, the sample will belong to class 1 and the process of classification is over. Otherwise, the feature vectors are secondly sent into the classifier 2. Then the following process does the same as above until the classification is finished. If the final output is -1, the sample will not be included in the  $k$  class statuses.

### 3. Signal Acquisition and Feature Analysis

#### 3.1. Signal Acquisition

Experiments were conducted on an internal combustion engine with the EQ6100 type. Faults were set to simulate abnormal statuses of connecting rod bearing in advance. Four wear conditions including normal wear, slight wear, medium wear and severe wear were set on the fifth cylinder of the ICE which has total six cylinders. As listed in Table 1, the clearances between bearing and axle neck corresponding to four statuses were set using

plug gauges except for normal condition. In the experiments, each time only one fault was set in order to make the simulated fault as real as possible and reduce disturbances as much as possible. Vibration signals were acquired by B&K4384 model accelerometer mounted on the shell of the fifth cylinder of the ICE. Then, the signals were amplified by B&K2635 model charge amplifier and saved in computer through the data acquisition card. The rotating speed of the engine was kept at 1000 rpm without loads, and the sampling frequency was set to be 10 kHz.

**Table 1.** Faults setting.

Status	Clearance(mm)
Normal Wear	0.04-0.098
Slight Wear	0.15
Medium Wear	0.30
Severe Wear	0.50

### 3.2 Feature Analysis of Vibration Signals

Due to the complex and stochastic nature of internal combustion engine, there are many excitation sources and complex interferences in its working process, such as transformation between rotating and reciprocating movements, ignitions and explosions, some unbalance impacts, differences and interferences in different working cycles, etc. The connecting rod bearing is a critical component which joins the rotating and reciprocating movements. The vibration signals of CRB are modulated and attenuated by multiple parts before being transferred to the sensor. Therefore, the measured vibration signals include lots of disturbance noise. To make the signal analysis results equitable, the measured signal firstly needs to be preprocessed to minimize the noise caused by the sensor itself and machine working environment. This can be done based upon the mean-variance standardization method as described below.

Given the measured vibration signal is one-dimensional data  $\{x_1, x_2, \dots, x_n\}$ , where  $n$  is the length of sample. The mean of the sample can be written as follows:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (5)$$

And the variance of the sample can be written as:

$$\sigma = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (6)$$

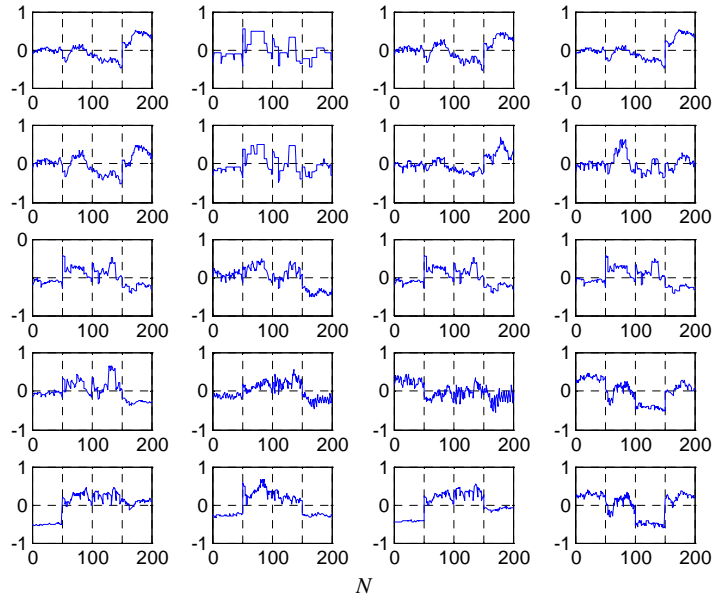
Then the sample is standardized as follows:

$$x'_i = \frac{x_i - \bar{x}}{\sqrt{\sigma}} \quad (7)$$

After mean-variance standardized, vibration signals will have zero mean and unit variance. In general, it is difficult to classify the vibration signals of CRB into different statuses using time-domain features only. The features in frequency-domain can also reflect different statuses of CRB to a certain extent. For example, the energy distribution in some typical frequency band could reflect CRB condition. Therefore, vibration features in both time-domain and frequency-domain were extracted to identify the health status of CRB comprehensively. These features include:

Time-Domain Features: absolute mean, maximum peak, the root mean square, the variance, peak-peak, the kurtosis, the skewness, the square root, the crest factor, the shape factor, the kurtosis factor, the impulse factor, and the tolerance factor;

Frequency-Domain Features: average frequency, stabilization factor of spectrum peak, as well as five spectral energy values on five equally divided frequency bands.



**Fig. 1.** Features of vibration signals of CRB.

The total 20 features described above, which calculated according to the reference [13], were extracted from vibration signals of internal combustion engine. Thus, the health status of connecting rod bearing could be diagnosed comprehensively by a combination of these features. Features as illustrated in Fig.1 were extracted from vibration signals of

connecting rod bearing with four conditions (normal wear, slight wear, medium wear and severe wear). As shown in Fig.1, the horizontal ordinate represents sample numbers, in which 50 samples for each type of signal are analyzed. It can be seen that these features are different from each other. The 20 features represent different physical signification respectively and the sensitivity and regularity of each feature is different. It is difficult to find which feature is more sensitive and stable intuitively than others in representing the status of CRB. In order to avoid the scale effect of the features and make all the features hold in a uniform scale, features were also preprocessed with the mean-variance standardization method.

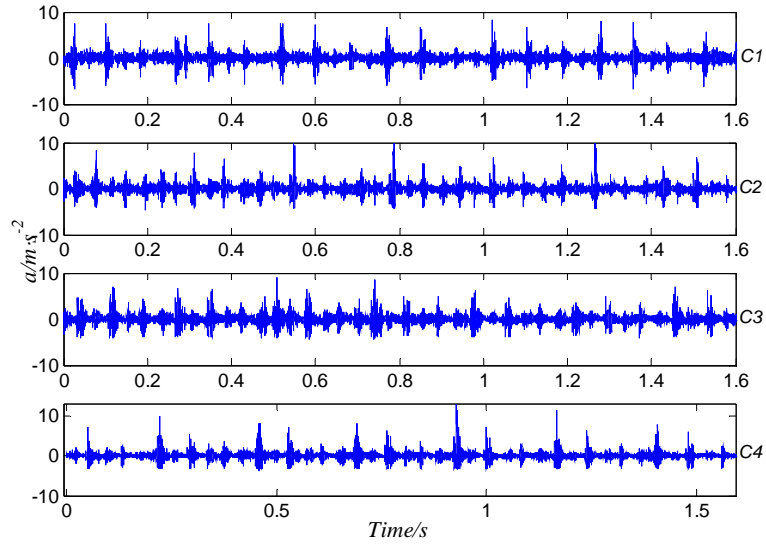
#### 4. SVM for Fault Diagnosis of CRB

After being preprocessed by the mean-variance standardization method, typical vibration signals of the CRB are shown in Fig. 2 with four wear conditions including normal wear (denoted as *C1*), slight wear (denoted as *C2*), medium wear (denoted as *C3*) and severe wear (denoted as *C4*). As seen from Fig. 2, it is difficult to find the difference among the four conditions. The corresponding frequency spectra are illustrated in Fig.3, from which we can see that the maximum spectral energy all locate around 1400Hz on four conditions. The energy is distributed widely on the spectrum except the severe wear condition, the energy of which is concentrated at 1400 Hz. From the frequency-domain, it is also difficult to see intuitively the difference among the four conditions.

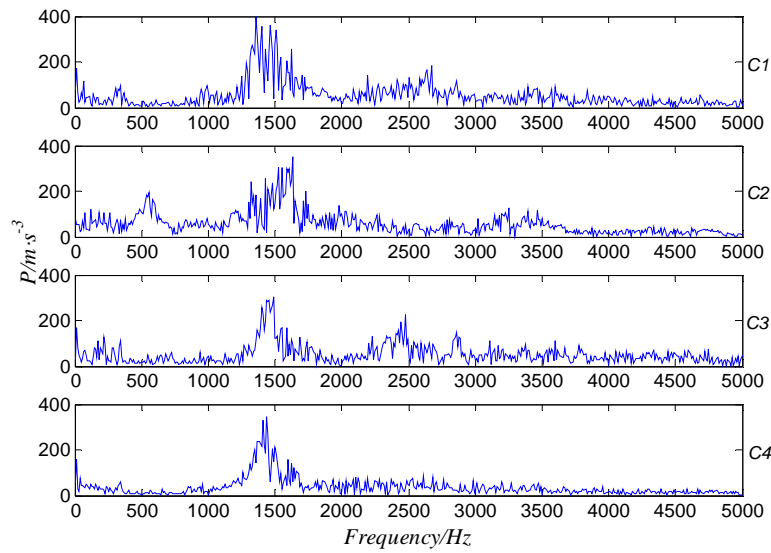
A set of data which is comprised of fifty vibration signals were acquired from each condition and twenty vibration features are extracted from each signal. Therefore, we can construct a  $50 \times 20$  vibration feature matrix for each condition. Fig. 1 illustrates the 20 vibration features on four conditions respectively. From Fig.1, it is difficult to find which feature is more sensitive and more stable than others in representing the status of connecting rod bearing. We selected the fore-half of the feature matrix of each condition, i.e., a  $25 \times 20$  sub-matrix, to train the SVM classifier and used the left half of the matrix for testing.

It is well known that the classification performance of the SVM depends on what kernel function is used to a certain degree. In the training, we compared linear kernel function, polynomial kernel function and radial basis kernel function. The training results using the three kernel functions are illustrated in Table 2, where the training time is denoted by TT, and the correct recognition rate is denoted by CRR. The parameters of the kernel functions are taken as: power index  $d=3$  for the polynomial kernel function and  $\gamma = 2.17$  for the radial basis function. It can be seen from Table 2 that recognition accuracy rates are almost the same when using all three kernel functions. But the training time of SVM classifier based on radial basis function is much less than that based on the other two functions. Meanwhile, the radial basis kernel function has a less parameter advantage over the polynomial one. The complexity of SVM would be affected directly

by the numbers of parameter. Therefore, in this paper, we constructed the SVM classifier using radial basis kernel function to identify the status of connecting rod bearing.



**Fig. 2.** Vibration signals of CRB on different conditions.



**Fig. 3.** Spectra of CRB on different conditions.

The essential problem of machine condition monitoring and fault diagnosis is pattern recognition [9]. The correct feature recognition is the basis of identifying the status of machine. The performance of the classifier will directly affect the fault diagnosis of machine. Therefore, the three-layer back propagation neural network (BPNN) and  $k$ -nearest neighbor classification (KNN) were used to compare with one-against-all SVM multi-classifier. In the training process of the BPNN, the transfer functions were chosen as log-sigmoid function and tan-sigmoid function respectively. Training times was set to 200, and network performance object was set to 0.001. Euclidean distance and one nearest neighbor were used in the KNN classification. When training the SVM classifier, the least square SVM algorithm was used. Radial basis function was selected as the kernel function with the kernel parameter  $\gamma$  being equal to 2.17. Table 3 shows the training time and correct recognition rate of the four statuses by using 3-layer BPNN, KNN and SVM classifier respectively. As seen from table 3, the correct recognition rates are almost the same on normal wear condition using BP neural network and SVM classifier. However, on other three conditions, the correct recognition rate of SVM classifier is higher than the rate of BPNN in evidence. Moreover, the training time of BPNN is far greater than SVM classifier used. Although the computer time of KNN is similar to SVM, its correct recognition rate is much lower than that of the other two methods. The results illustrates that the SVM is effective to identify the health status of connecting rod bearing of ICE.

**Table 2.** Diagnosis results by SVM with different kernel functions.

Status of CRB	Linear Function		Polynomial Function		Radial Basis Function	
	TT	CRR	TT	CRR	TT	CRR
<i>C1</i>	0.849s	100%	0.822s	99.4%	0.632s	99.3%
<i>C2</i>	0.799s	98.3%	0.843s	98.3%	0.527s	97.8%
<i>C3</i>	0.833s	99.5%	0.830s	99.1%	0.590s	98.5%
<i>C4</i>	1.486s	98.2%	0.849s	97.5%	0.591s	96.7%

**Table 3.** Comparison of BPNN, KNN and SVM results.

Status of CRB	KNN		BPNN		SVM	
	TT	CRR	TT	CRR	TT	CRR
<i>C1</i>	0.921s	71.1%	7.023s	98.5%	0.632s	99.3%
<i>C2</i>	0.824s	73.3%	7.316s	90.7%	0.527s	97.8%
<i>C3</i>	0.989s	77.3%	6.088s	90.8%	0.590s	98.5%
<i>C4</i>	0.709s	79.5%	8.487s	90.4%	0.591s	96.7%



## 5. Conclusions

This paper addresses on identification of the CRB health status of the ICE by using limited samples. By comparing, the SVM multi-classifier based on radial basis kernel function is chosen for identifying the health status of CRB. Furthermore, considering the training time and correct recognition rate in comparison with the BP neural networks, the SVM classifier based on statistical learning theory is especially fit for fault diagnosis in small-sample cases.

The results in this study validate the effectiveness of applying SVM multi-classifier to identifying the health status of connecting rod bearing. This illustrated that the SVM has great practical value and provide an efficient method for intelligent diagnosis of ICE.

## Acknowledgements

This work was partially supported by the National Natural Science Foundation of China (No.50905021) and by the Fundamental Research Funds for the Central Universities in China.

## References

1. Z.-Y. Wu and H.-Q. Yuan.: Research on Fault Diagnosis of Engine by Ant Colony Support Vector Machine. *Journal Of Vibration And Shock*, 28(3):83-86 (2009)
2. J. Jia, F. Kong, Y. Liu, et al.: Noise Diagnosis Research Based on Wavelet Packet and Fuzzy C-clusters about Connecting Rod Bearing Fault. *Transactions of the Chinese Society for Agricultural Machinery*, 36(6):87-91 (2005)
3. K. S. R. Rao and M. A. Yahya.: Neural Networks Applied for Fault Diagnosis of AC Motors. In: *International Symposium of Information Technology*, Vols 1-4, Proceedings: 2607-2612 (2008)
4. F. Sahin, M. C. Yavuz, et al. : Fault Diagnosis for Airplane Engines Using Bayesian Networks and Distributed Particle Swarm Optimization. *Parallel Computing*, Vol. 33( 2):124-143 ( 2007)
5. V. N. Vapnik.: *The Nature of Statistical Learning Theory*. Springer-Verlag, New York & Berlin (1995)
6. V. N. Vapnik.: *Statistical Learning Theory*, Wiley, New York (1998)
7. V. Vapnik, S. Golowich, et al.: Support Vector Method for Function Approximation Regression, Estimation, and Signal Processing. *Advances in Neural Information Processing Systems*, 9: 281-287 (1996)
8. Y. Liu, Z. You, et al.: A Novel And Quick SVM-Based Multi-Class Classifier. *Pattern Recognition*, 39: 2258-2264 ( 2006)
9. Q. Hu, Z. He, et al.: Fault Diagnosis of Rotating Machinery Based on Improved Wavelet Package Transform and SVMs Ensemble. *Mechanical Systems and Signal Processing*, 21: 688-

705 (2007)

10. S.-F. Yuan and F.-L. Chu.: Support Vector Machines and Its Applications in Machine Fault Diagnosis. *Journal Of Vibration And Shock*, 26:29-35 (2007)
11. S.-F. Yuan and F.-L. Chu. Support Vector Machines-Based Fault Diagnosis for Turbo-pump Rotor. *Mechanical Systems and Signal Processing*, 20: 939-952 (2006)
12. Y. Yang, D. Yu, et al.: A Fault Diagnosis Approach for Roller Bearing Based on IMF Envelope Spectrum And SVM. *Measurement*, 40: 943-950 (2007)
13. K. Chen and C. Li.: *Machine Condition Monitoring and Fault Diagnosis Technology*, Beijing Science and Technology Press, Beijing (1991)