Multi-sensor information fusion method for vibration fault diagnosis of rolling bearing

Jing Jiao, Jianhai Yue and Di Pei
School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, Beijing, China
E-mail: 15121285@bjtu.edu.cn

Abstract. Bearing is a key element in high-speed electric multiple unit (EMU) and any defect of it can cause huge malfunctioning of EMU under high operation speed. This paper presents a new method for bearing fault diagnosis based on least square support vector machine (LS-SVM) in feature-level fusion and Dempster-Shafer (D-S) evidence theory in decision-level fusion which were used to solve the problems about low detection accuracy, difficulty in extracting sensitive characteristics and unstable diagnosis system of single-sensor in rolling bearing fault diagnosis. Wavelet de-noising technique was used for removing the signal noises. LS-SVM was used to make pattern recognition of the bearing vibration signal, and then fusion process was made according to the D-S evidence theory, so as to realize recognition of bearing fault. The results indicated that the data fusion method improved the performance of the intelligent approach in rolling bearing fault detection significantly. Moreover, the results showed that this method can efficiently improve the accuracy of fault diagnosis.

1. Introduction
As an important part of rotating machinery, rolling bearing plays an important role in the accuracy, performance, status and life of the equipment. With the wide application of the mechanical equipment, the fault caused by damage of mechanical equipment bearing occupy a large proportion, so the significance of rolling bearing fault diagnosis is obvious. Kun Yu et al. [1] employs a combined EEMD and SVD technique to extract useful fault features from the condition monitoring data of rolling element bearings. Ye Tian et al. [2] proposed an intelligent fault diagnosis method based on LMD, SVD and ELM. However, the rolling bearing vibration signals are often non-stationary and represent non-linear processes. Information fusion technology can make full use of the information resources of multi sensors to make the best decision. Yan Tingai et al. [3] proposed a fusion method based on n-DCPD for rolling bearing fault diagnosis with two types of signals. Xiaoyun Gong et al. [4] suggests a new feature frequency extraction approach for gear malfunctions diagnosis using the FCS and EEMD method. Chuan Li et al. [5] addresses the use of DRFF technique to improve fault diagnosis performance for gearboxes by using acoustic emission sensor and accelerometer. In this research, the vibration signal is collected from the high-speed EMU transmission test bench. And eight parameters in both time domain and frequency domain were used to obtain the feature sets. LS-SVM is employed in feature fusion. In order to increase the classification accuracy, data fusion method using D-S evidence theory is employed. The advantage of the proposed procedure is to detect fault in rolling bearings.
2. Multi-sensor Information Algorithm

2.1. Least Square Support Vector Machine (LS-SVM)
Proper kernel function and optimum kernel parameters are of importance in LS-SVM classifier. Given the training sample set \( \{(x_1, y_1), \ldots, (x_n, y_n)\} \), \( x_i \in \mathcal{R}^n \), \( y_i \in [-1,1] \). The regressive function is shown as
\[
y(X) = \alpha^T X + b
\]
where \( X = (x_1, x_2, \ldots, x_n) \) is a nonlinear transform function, which maps the input space to higher dimension space, \( \omega = (\omega_1, \omega_2, \ldots, \omega_n) \) is the weight coefficient, and \( b \) is threshold value.

Based on structural risk minimization (SRM), the regressive optimization problem is summarized to minimize the function:
\[
\min \frac{1}{2} \| \theta \|^2 + \frac{1}{2} \sum_{i=1}^{n} \xi_i
\]
which is subjected to the constraint:
\[
y_i = \langle \omega^T X \rangle + b + \xi_i
\]
where \( \alpha \) is penalty factor, \( \xi_i \) is the slack variable.

The optimization classification function can be obtained by introducing Lagrange function in:
\[
f(X) = \sum_{i=1}^{n} \alpha_i K(X, X_i) + b
\]
where \( \alpha_i \) is lagrange multiplier and \( K(X, X_i) \) is kernel function.

2.2. Dempster-Shafer (D-S) Evidence Theory
Evidence theory is the main means to deal with uncertain, incomplete and inaccurate information [6]. Assuming a finite nonempty set of mutually exclusive alternatives \( \Theta \) is the frame discernment of \( \theta \), if the function \( m: 2^\Theta \rightarrow [0,1] \) satisfies \( m(\emptyset) = 0, \sum_{A \in \Theta} m(A) = 1, \forall A \subseteq \Theta, m \) is called the basic probability assignment (BPA). \( Bel(A) \) is used to express the belief levels of hypothesis in D-S evidence theory, and the expression formula is as follows:
\[
\forall A \subseteq \Theta, Bel(A) = \sum_{B \subseteq A} m(B)
\]
However, it is not comprehensive of using the \( Bel(A) \) to describe the belief levels of hypothesis. Therefore, \( Pl(A) \), which represents the plausibility belief level, is introduced to complement the description and can be defined as follows:
\[
\forall A \subseteq \Theta, Pl(A) = 1 - Bel(\overline{A}) = \sum_{B \supset A} m(B)
\]
In order to facilitate the judgment and decision making, the algorithm of synthesizing and combining two or more different BPAs were proposed, called as Dempster’s combination rule [7]. Assuming that \( E_1 \) and \( E_2 \) belong to the same recognition framework \( \Theta \), and the corresponding basic probability assignment is \( m_1 \) and \( m_2 \). The combination rule can be described as follows:
\[
m_{\cap_2} (A) = \begin{cases} \sum_{A \cap_1 B = A} m_1(A) \cdot m_2(B) / (1 - K), & A \neq \emptyset \\ \sum_{A \cap_1 B = \emptyset} m_1(A) \cdot m_2(B), & A = \emptyset \end{cases}
\]
where \( A, B, A \in \Theta, K = \sum_{A \cap_1 B = \emptyset} m_1(A) \cdot m_2(B) \).

3. Fault Diagnosis Based on Information Fusion

3.1. Experimental Setup
The research of this paper is based on the high-speed EMU transmission test bench, as is shown in figure 1. It simulates the transmission mode of a certain type of EMU running gear, which is mainly composed of motor, gear box, driving wheel, track wheel, bearings, drive axle and generator. The operation box has the function of start/stop, emergency stop, speed control and real-time parameters.
(including given speed, actual frequency, output speed, actual pressure and actual current). The test rolling bearing is 351306 which are normal bearing (N), out-ring fault bearing (O) and inner-ring fault bearing (I).

3.2. Structure of Fusion Prediction Model

The idea of fault diagnosis based on LS-SVM and D-S evidence theory can be summarized as follows. Firstly, wavelet de-noising was applied in processing stage to remove the signal noise. Afterward, some statistical feature parameters were used for extracting the useful diagnostic information from the signals. LS-SVM classifiers were employed for fault detection and classification. Data fusion method performed by D-S theory of evidence shows the flowchart of the proposed procedure. The used techniques are briefly explained in the following. The flowchart of the signal processing is shown as figure 2.

3.3. Bearing Fault Diagnosis

There are three kinds of rolling bearings near the motor side of the test bench, which are normal, outer-ring fault and inner-ring fault. The other bearings are normal. The speed of test bench is 3000 rpm, which is close to the speed of 250 km/h in real EMU running. All signals were simultaneously sampled with a 10k Hz sampling frequency. And the length of each original signal captured was 40,960. The processed signals contain a large number of data points and could not be used as inputs of the classifier, because high-dimensional data increases computational complexity so, training of the classifier becomes very difficult [8]. Therefore, some statistical features should be applied to reduce the dimensionality of data.
The time domain parameters are peak index $K_{CF}$, impulsion index $K_{IF}$, tolerance index $K_{CLF}$ and kurtosis index $K_{u}$; the frequency domain parameters are mean frequency $u_{mf1}$, frequency center $ufc_1$, frequency root mean square (RMS) $u_{rmsf_1}$ and frequency standard deviation of frequency $u_{stdf_1}$. Because of the different definitions of these parameters, it is necessary to normalize these feature sets [9]. The results of the fault identification of the LS-SVM in time domain and frequency domain are tabulated in table 1.

**Table 1.** Classification results of LS-SVM in time domain and frequency domain

| Bearing condition | time domain | frequency domain |
|-------------------|-------------|------------------|
|                   | N           | I                | O    | Average |
|                   | 0.83        | 0.10             | 0.07 |         |
|                   | 0.09        | 0.71             | 0.20 | 0.75    |
|                   | 0.23        | 0.72             |      |         |
|                   | 0.89        | 0.03             | 0.08 |         |
|                   | 0.02        | 0.91             | 0.07 | 0.89    |
|                   | 0.05        | 0.08             | 0.87 |         |

It can be seen from table 1 that frequency domain characteristic parameters are better than time domain characteristic parameters in classification. Therefore, in order to obtain higher fault recognition accuracy, the fault recognition result obtained by frequency domain characteristic parameters is used as the characteristic input of D-S evidence theory. The accuracy of decision-making layer fusion are shown in table 2.

**Table 2.** Classification accuracy of decision-making layer fusion

| Bearing condition | N           | I          | O          | Average |
|-------------------|-------------|------------|------------|---------|
|                   | 0.98        | 0.02       | 0          |         |
|                   | 0.01        | 0.96       | 0.03       | 0.97    |
|                   | 0.01        | 0.01       | 0.98       |         |

4. Conclusion
In this paper, LS-SVM and D-S evidence theory was proposed to solve the problem of bearing fault diagnosis. On the basis of the high-speed EMU transmission test bench, the multi-source vibration signals of the bearing were analysed by information fusing method. The experimental results show that the LS-SVM fault identification based on the frequency domain feature parameters is more accurate. The decision-making layer fusion result shows that the accuracy of the fault diagnosis is improved greatly through the combination of the feature layer and the decision layer information fusion, which is of great significance for the bearing fault diagnosis.

Acknowledgments
The study is supported by Technology Division of China Railway (M15D00230).

References
[1] Yu K, Lin T Rn and Tan J W 2017 *Applied Acoustics* 121 33-45
[2] Tian Y, Ma J, Lu C and Wang Z 2015 *Mechanism and Machine Theory* 90 175-186
[3] Ai Y, Guan J, Fei C, Tian J and Zhang F 2017 *Mechanical Systems and Signal Processing* 88 123-136
[4] Gong X, Ding L, Du W and Wang H 2017 *Procedia Engineering* 174 918-926
[5] Li C, Sanchez R, Zurita G, Carrera D and Vásquez R E 2016 *Mechanical Systems and Signal Processing* 76–77 283-293
[6] Shafer G 1976 *A mathematical theory of evidence* Princeton University Press, Princeton 10-15
[7] Si L, Wang Z, Tan C and Liu X 2014 *Measurement* 54 140-151
[8] Zhang L, Xiong G, Liu H, Zou H and Guo W 2010 *Expert Systems with Applications* 37 6077-6085
[9] Jing J and Jianhai Y 2016 *1st International Workshop on Structural Health Monitoring for Railway System* 296-303