

基于多传感器数据融合的轨道车辆 齿轮箱异常检测

刘玉梅¹, 乔宁国¹, 庄娇娇¹, 刘鹏程¹, 胡 婷¹, 陈立军²

(1. 吉林大学 交通学院, 长春 130022; 2. 吉林大学 大数据和网络管理中心, 长春 130022)

摘 要: 针对轨道车辆齿轮箱监测测点多、数据量大、数据融合程度低等特点, 提出一种基于相关函数融合算法与模糊C均值聚类结合的齿轮箱异常检测方法。相关函数融合算法用于将齿轮箱多个测点采集的振动信号融合为一个能全面反映齿轮箱运行状态的信号; 对融合信号进行聚合经验模态分解, 并计算奇异熵、能量熵; 采用模糊C均值聚类算法对特征集进行簇划分, 判断齿轮箱含有几类异常情况。通过实际线路运行数据的采集与分析, 验证了本文方法的有效性。

关键词: 铁路运输; 轨道车辆; 齿轮箱异常检测; 相关函数融合算法; 模糊C均值聚类

中图分类号: U260 **文献标志码:** A **文章编号:** 1671-5497(2019)05-1465-06

DOI: 10.13229/j.cnki.jdxbgxb20180559

Anomaly detection of rail vehicle gearbox based on multi-sensor data fusion

LIU Yu-mei¹, QIAO Ning-guo¹, ZHUANG Jiao-jiao¹, LIU Peng-cheng¹, HU Ting¹, CHEN Li-jun²

(1. College of Transportation, Jilin University, Changchun 130022, China; 2. Big Data and Network Management Center, Jilin University, Changchun 130022, China)

Abstract: Because the characteristics of many monitoring points, large data and low degree of data fusion of the gearbox in a rail vehicle, an anomaly detection method of the gearbox is proposed based on the correlation function fusion algorithm and fuzzy C means clustering. The correlation function fusion algorithm is employed to fuse the vibration signals collected from many monitoring points into a signal which can fully reflect the running state of the gearbox. The ensemble empirical mode decomposition (EEMD) method is applied to decompose the fused signal, and the energy entropy and singular entropy of the first n intrinsic mode functions (IMFs) are calculated. The fuzzy C means (FCM) clustering algorithm is employed to cluster the feature set to determine several abnormal situations in the gearbox. Through collection and analysis of the data during actual operation, the effectiveness of the proposed method is verified.

Key words: railway transportation; rail vehicle; anomaly detection of gearbox; correlation function fusion algorithm; fuzzy C means clustering

收稿日期: 2018-06-04.

基金项目: 国家自然科学基金项目(51575232); 吉林省科技厅重点科技攻关项目(20160204018GX).

作者简介: 刘玉梅(1966-), 女, 教授, 博士生导师. 研究方向: 轨道车辆工程, 车辆智能化检测及诊断.

E-mail: lymls@163.com

通信作者: 陈立军(1960-), 男, 副研究员. 研究方向: 计算机应用. E-mail: clj@jlu.edu.cn

0 引言

齿轮箱作为轨道车辆传动系的重要组成部分,在列车运行中除了受到齿轮啮合作用,还受轮对传递的轨道不平顺激励,工作环境极其复杂。我国在运的轨道车辆已发生过轴承高温报警、箱体裂纹等故障^[1],因此,齿轮箱的异常检测对于行车安全意义重大。

随着传感器技术的发展,轨道车辆齿轮箱监测呈现多测点、数据量大、数据融合程度低等特点。因此,将多传感器数据融合技术与齿轮箱状态监测结合,是提高故障诊断准确率的有效手段。多传感器数据融合分为三级,数据级、特征级和决策级^[2-4]。由于本文采集同一类型信号,为分析齿轮箱振动特性,采用相关函数融合算法。齿轮箱异常检测首先需要提取故障敏感特征。由于轨道车辆传动系统各零部件间振动信号相互叠加、耦合现象严重,齿轮箱振动信号呈现非线性、非平稳性特点^[5]。而EEMD能量熵、奇异熵能很好地反映部件的工作状况,因此选择EEMD熵作为敏感特征。聚类算法选择是异常检测的另一项重要内容。常用算法有K-means算法、K-Medoid算法和FCM算法等^[6,7]。由于引入了模糊理论,因此选用理论成熟的FCM算法。

综上所述,本文提出一种相关函数融合算法与FCM算法相结合的轨道车辆齿轮箱异常检测方法。首先,基于相关函数融合算法得到一个能综合反映齿轮箱运行状态的信号;其次,对融合信号进行EEMD分解,并计算IMF奇异熵、能量熵;最后,通过设定不同的聚类数对特征集进行簇划分,通过比较聚类结果判定齿轮箱有无异常及含有几类异常,实现齿轮箱的异常检测。

1 异常检测方法

1.1 相关函数融合算法

相关函数体现了信号间的相关关系,其定义如下:假设两个离散信号 $x(n)$ 和 $y(n)$,则 $x(n)$ 和 $y(n)$ 之间互相关函数为:

$$R_{xy}(m) = \frac{1}{N-m} \sum_{n=1}^{N-m} x(n)y(n+m) \quad (1)$$

$$m=0,1,2,\dots,k$$

式中: N 为数据个数。

$x(n)$ 和 $y(n)$ 之间相关系数为:

$$\rho_{xy} = \frac{R_{xy}}{\sqrt{\sum_{n=0}^{\infty} x^2(n) \sum_{n=0}^{\infty} y^2(n)}} = \frac{R_{xy}}{\sqrt{E_x E_y}} \quad (2)$$

式中: E_x 、 E_y 分别为 $x(n)$ 、 $y(n)$ 的能量。

根据许瓦兹不等式,当 $x(n)$ 和 $y(n)$ 完全相等时, $\rho_{xy}=1$;当 $x(n)$ 和 $y(n)$ 相关时, $0 < \rho_{xy} < 1$;当 $x(n)$ 和 $y(n)$ 完全无关时, $\rho_{xy}=0$ 。

假设采集的振动信号为 $x_1(n)$ 、 $x_2(n)$ 、 $x_3(n)$ 、 \dots 、 $x_m(n)$,互相关运算式为:

$$R_{ij} = \frac{1}{N-m} \sum_{i=1}^{N-m} x_i(n)x_j(n+m) \quad (3)$$

$$m=0,1,2,\dots,k$$

则振动信号的能量为:

$$E_{ij} = \sum_{i=1}^n [R_{ij}(i)]^2 \quad (4)$$

信号 $x_i(n)$ 与其他信号的总相关能量为:

$$E_i = \sum_{j=1, j \neq i}^n E_{ij} \quad (5)$$

根据权值与相关程度(即相关函数的能量)成正比原则有:

$$[p_1:p_2:\dots:p_n = E_1:E_2:\dots:E_n] \quad (6)$$

式中: $p_1+p_2+\dots+p_n=1$ 。

则各信号的权值为:

$$X = p_1x_1 + p_2x_2 + \dots + p_nx_n \quad (7)$$

式中: p_i 为第 i 个振动信号的权值。

1.2 EEMD熵特征提取

EEMD是由Wu等^[8]在EMD基础上提出的方法,即在EMD基础上加上高斯白噪声。具体步骤为:

Step1 在振动时序 $x(t)$ 上加入高斯白噪声 $z(t)$:

$$y(t) = x(t) + z(t) \quad (8)$$

Step2 将 $y(t)$ 分解为 J 个IMF分量 c_i 及余项 r 之和:

$$y(t) = \sum_{j=1}^J c_j + r \quad (9)$$

Step3 重复 M 次Step1、Step2;

Step4 取IMF均值作为最后的结果,即:

$$\begin{cases} \bar{c}_j = \frac{1}{M} \sum_{i=1}^M c_{i,j} \\ \bar{r}_j = \frac{1}{M} \sum_{i=1}^M r_i \end{cases} \quad (10)$$

$$i=1,2,\dots,M; j=1,2,\dots,J$$

式中: M 为聚合次数; $c_{i,j}$ 为经 i 次EMD获得的第 j 个IMF分量; r_i 为经 i 次EMD获得的余项; J 为IMF分量个数; \bar{c}_j 、 \bar{r}_j 分别为 M 次EMD获得的IMF及余项的均值。

EEMD能量熵:计算前 n 个有效IMF分量的能量:

$$E_i = \int |c_i(t)|^2 dt = \sum_{j=1}^J |x_{ij}^2| \quad (11)$$

式中: $i=1,2,\dots,n$; x_{ij} 为各离散点的幅值。

设 $p_i = E_i/E$, 且 $\sum_{i=1}^n p_i = 1$, $E = \sum_{i=1}^n E_i$, 则EEMD能量熵为 $W_E = -\sum_{i=1}^n p_i \lg p_i$ 。

EEMD奇异熵为:

$$W_S = -\sum_{i=1}^n (E_i \ln E_i) \quad (12)$$

式中: $E_i = \frac{\sigma_i}{E}$, $E = \sum_{i=1}^n \sigma_i$, σ_i 为前 n 个IMF奇异值分解得到的奇异值。

1.3 模糊 C 均值聚类

给定数据集 $X = \{x_1, x_2, \dots, x_N\}$, $x_i \in R^m$, 样本数为 N ; 特征维数为 m 。FCM算法通过求取目标函数的最小值将数据集 X 划分为 C 类^[9]:

$$F = \sum_{i=1}^N \sum_{j=1}^C u_{ij}^p \|x_i - c_j\|^2, p > 1 \quad (13)$$

式中: C 为聚类数; c_j 为聚类中心:

$$c_j = \frac{\sum_{i=1}^N u_{ij}^p x_i}{\sum_{i=1}^N u_{ij}^p}, 1 \leq j \leq C \quad (14)$$

式中: u_{ij}^p 为数据点 x_i 属于 j 类的隶属度:

$$u_{ij} = \left[\frac{\sum_{k=1}^C \left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{p-1}}}{\sum_{k=1}^C \left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{p-1}}} \right]^{-1} \quad (15)$$

$$1 \leq i \leq N, 1 \leq j \leq C$$

式中: p 用于控制数据点隶属度的模糊性。

使用FCM算法对数据集 X 聚类得到如下 $C \times N$ 隶属度矩阵 U :

$$U = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1N} \\ u_{21} & u_{22} & \cdots & u_{2N} \\ \vdots & \vdots & & \vdots \\ u_{C1} & u_{C2} & \cdots & u_{CN} \end{bmatrix} \quad (16)$$

2 轨道车辆齿轮箱早期异常检测方法

基于相关函数融合算法和FCM聚类轨道车辆齿轮箱异常检测流程如图1所示。

具体步骤如下:

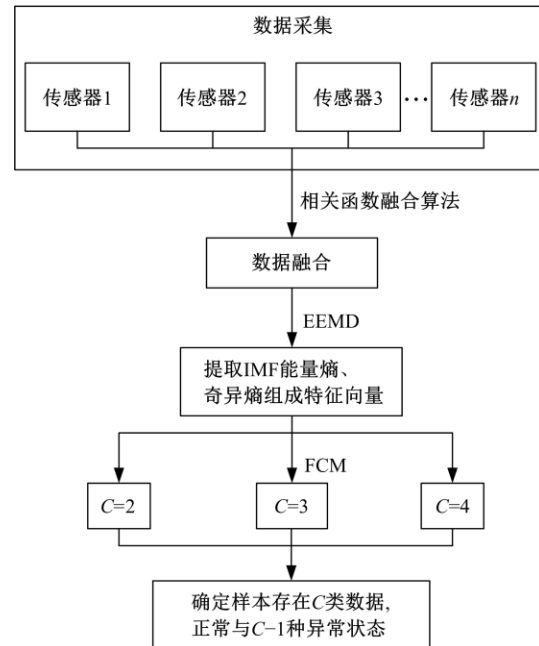


图1 轨道车辆齿轮箱异常检测流程图

Fig. 1 Abnormal detection flow chart of gearbox in rail vehicle

Step1 采集齿轮箱故障敏感位置的振动信号。

Step2 利用相关函数融合算法计算测点间信号的相关性,并计算相关函数能量,根据能量比推导出信号的加权系数,得到融合信号。

Step3 通过FFT变换,检验融合方法的有效性。

Step4 对融合信号进行EEMD分解,并计算前 n 个IMF奇异熵、能量熵。

Step5 $C=2\sim4$,采用FCM聚类算法对特征集进行簇划分,通过比较不同的聚类结果判定齿轮箱有无异常及含有几类异常。

3 实例验证

3.1 数据采集

本文数据来源于某型高速动车组实际线路运行采集的振动加速度数据。由于该型列车在运营里程达1 700 000 km时,经有关部门检测发现齿轮箱润滑油Fe含量超标,但对Fe含量超标齿轮箱远程温度数据进行查询,未发现温度异常情况也未出现故障报警。对Fe含量超标的齿轮箱进行分解检测,发现齿轮箱输出端电机侧轴承外滚道出现剥离情况,如图2所示。因此,对Fe含量超标齿轮箱监测的振动数据进行进一步分析研究,在无故障数据作为训练样本的情况下,提出了

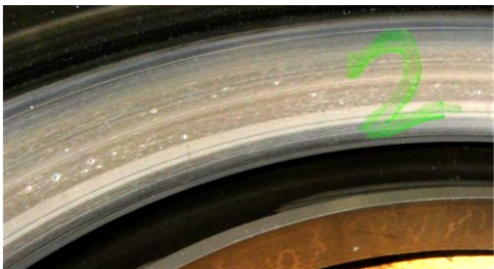
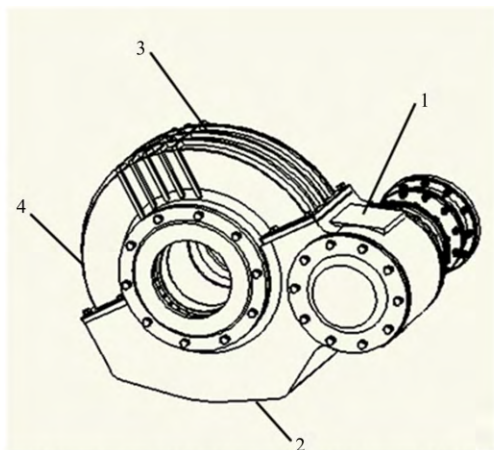


图 2 1轴输出端轴承外滚道剥离情况

Fig. 2 Exfoliation of outer ring of 1 axis output bearing
一种基于相关函数融合算法与FCM聚类相结合的齿轮箱异常检测方法。

齿轮箱上振动加速度传感器安装位置如图 3 所示。采样频率为 2048 Hz,对信号进行 FFT 变换,从图 4 中可见齿轮箱的振动频率为 343.9 Hz,另外,从大齿轮下方和大齿轮上方的频谱图中能明显看到齿轮箱轴承的故障特征频率(632.8 Hz)及其边频带。



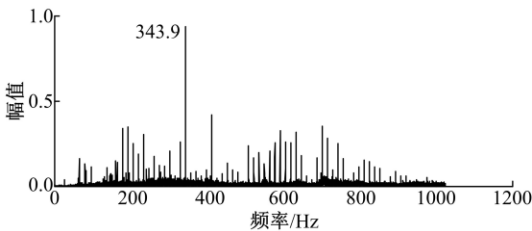
1-小齿轮上方; 2-大齿轮下方; 3-大齿轮上方; 4-齿轮箱端部

图 3 齿轮箱上传感器测点

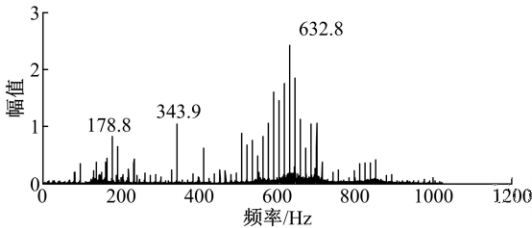
Fig. 3 Layout of sensors on gearbox

3.2 数据融合

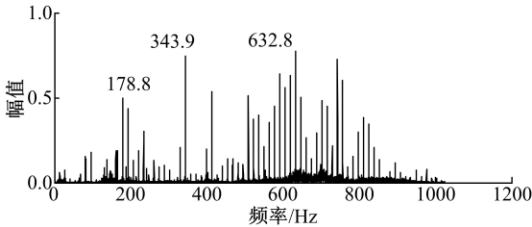
表 1 为根据式(1)(2)计算得到的信号间相关系数,可见大齿轮上方和大齿轮下方测点的信号具有强相关性,而其他测点间的相关性较弱。根据式(4)(5)得到各信号总相关能量分别为 1.1162×10^{12} 、 4.0348×10^{12} 、 3.3464×10^{12} 、 4.2899×10^{12} 。根据式(6)推导出各测点信号加权的权重分别为 0.0873、0.3155、0.2617、0.3355,得到的融合结果如图 5 所示。从图 6 中可见齿轮箱的振动频率为 343.9 Hz 及轴承故障的特征频率为 632.8 Hz。由此可以证明,相关函数融合算法得到的融合信号能较好地反映齿轮箱运行状况。



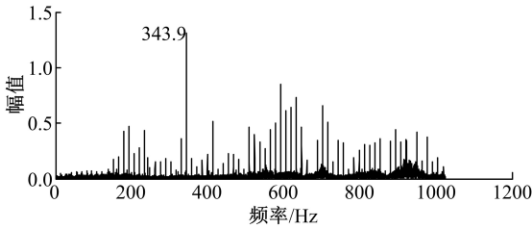
(a)小齿轮箱上方



(b)大齿轮下方



(c)大齿轮上方



(d)齿轮箱端部

图 4 频谱图

Fig. 4 Frequency-amplitude spectrums

表 1 信号间相关系数

Table 1 Correlation coefficients between vibration signals

测点	1	2	3	4
1	1	0.4910	0.4008	0.4184
2	0.4910	1	0.6951	0.4473
3	0.4008	0.6951	1	0.4679
4	0.4184	0.4473	0.4679	1

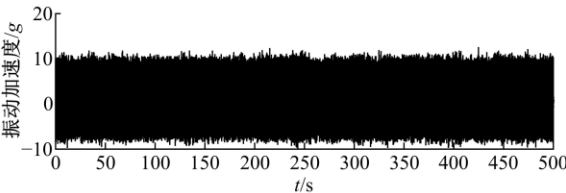


图 5 融合信号

Fig. 5 Fusion signal

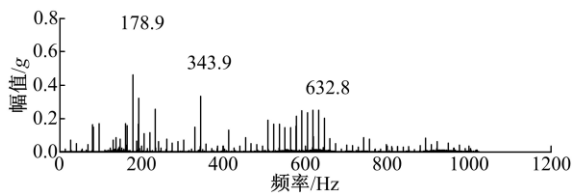


图6 融合信号频谱图

Fig. 6 Frequency-amplitude spectrum of fusion signal

3.3 EEMD熵特征提取

将监测数据分为300组,每组数据长度为20 480个点。文献[8]添加的白噪声幅值为待分解信号标准差的0.2倍,聚合次数进行100次,对融合信号进行分解。经计算,振动能量主要集中在前几个IMF分量,而最后几个分量意义不大,因此本文提取前8个IMF计算能量熵、奇异熵。从图7中可见,第211组数据出现故障,但在故障发生之后的一小段时间后熵值出现上升,这是因为在轴承外圈发生故障之后出现以特征频率为周期的振动,波形变得有规律。

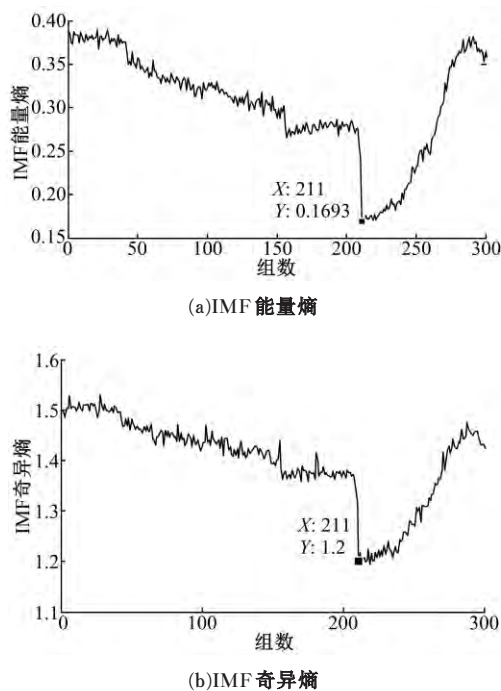


图7 EEMD熵特征

Fig. 7 EEMD entropy characteristics

3.4 模糊C均值聚类

将聚类数 C 设置为2~4,利用FCM聚类算法对特征集进行簇划分。图8为隶属度矩阵 U 和聚类结果,当 $C=2$ 时FCM聚类结果最好,两个簇被明显区分开,无模糊区域;当 $C=3$ 、 $C=4$ 时都存

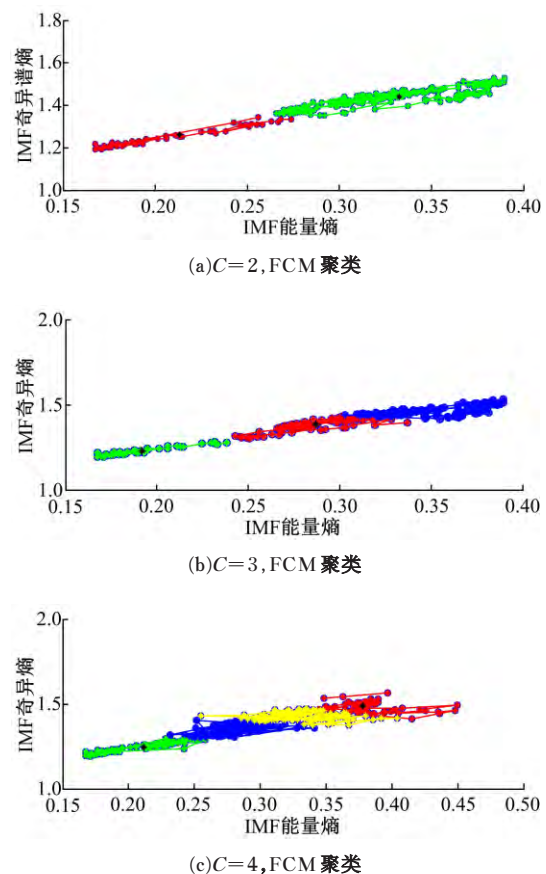
在模糊区域,说明数据中只含有两类数据,因此可判定齿轮箱除了正常状况外还出现了1类明显异常,同时说明恶化是一个渐变过程,具有模糊性:

$$U_2 = \begin{bmatrix} 0.933 & 0.946 & 0.955 & \cdots & 0.979 & 0.99 & 0.977 \\ 0.067 & 0.054 & 0.045 & \cdots & 0.021 & 0.01 & 0.023 \end{bmatrix} \quad (17)$$

$$U_3 =$$

$$\begin{bmatrix} 0.0166 & 0.0102 & 0.0066 & \cdots & 0.0173 & 0.0145 & 0.0189 \\ 0.0789 & 0.053 & 0.0367 & \cdots & 0.15 & 0.144 & 0.162 \\ 0.9045 & 0.9367 & 0.9567 & \cdots & 0.833 & 0.841 & 0.819 \end{bmatrix} \quad (18)$$

$$U_4 = \begin{bmatrix} 0.002 & 0 & 0 & \cdots & 0.032 & 0.019 & 0.022 \\ 0.97 & 0.99 & 0.99 & \cdots & 0.704 & 0.217 & 0.725 \\ 0.005 & 0 & 0 & \cdots & 0.074 & 0.086 & 0.06 \\ 0.022 & 0.01 & 0.01 & \cdots & 0.19 & 0.678 & 0.193 \end{bmatrix} \quad (19)$$

图8 $C=2\sim 4$,FCM的聚类结果Fig. 8 $C=2\sim 4$, FCM clustering results

4 结束语

针对轨道车辆齿轮箱监测测点多、数据量大、数据融合程度低,故障报警出现延迟的情况,提出

一种基于相关函数融合算法与模糊C均值聚类结合的齿轮箱异常检测方法。在数据融合过程中充分考虑了各测点数据间的相关性,并且通过FFT变换验证了得到的融合信号能很好地反映齿轮箱运行状态。在没有故障数据作为训练样本的情况下,选取不同C值进行聚类分析,可成功将数据划分为2簇,即正常和1类异常状态,与齿轮箱实际分解检测结果一致。

参考文献:

- [1] 袁雨青,李强,杨光,等. 高速列车齿轮箱线路测试与异常振动分析[J]. 铁道机车车辆, 2016, 36(1): 24-29.
Yuan Yu-qing, Li Qiang, Yang Guang, et al. Line test and abnormal vibration analysis of high speed train gearbox[J]. Railway Locomotive & Car, 2016, 36(1): 24-29.
- [2] 雷亚国,林京,何正嘉. 基于多传感器信息融合的行星齿轮箱故障诊断[C]//2010年全国振动工程及应用学术会议,沈阳, 2010: 216-218.
- [3] Liu Y, Fan X, Lyv C, et al. An innovative information fusion method with adaptive kalman filter for integrated INS/GPS navigation of autonomous vehicles[J]. Mechanical Systems & Signal Processing, 2017, 100: 605-616.
- [4] 徐旺敏,张宇驰,刘伍根,等. 基于相关函数的泄洪闸闸墩结构振动响应多测点融合研究[J]. 长江科学院院报, 2015, 32(11): 110-114.
Xu Wang-min, Zhang Yu-chi, Liu Wu-gen, et al. Research on multi point fusion of vibration response of floodgate pier structure based on correlation function[J]. Journal of Yangtze River Scientific Research Institute, 2015, 32(11): 110-114.
- [5] Liu Y M, Zhao C C, Xiong M Y, et al. Assessment of bearing performance degradation via extension and EEMD combined approach[J]. Journal of Central South University (English Edition), 2017, 24(5): 1155-1163.
- [6] 贾凡,严妍,张家琪. 基于K-means聚类特征消减的网络异常检测[J]. 清华大学学报:自然科学版, 2018, 58(2): 137-142.
Jia Fan, Yan Yan, Zhang Jia-qi. K-means based feature reduction for network anomaly detection[J]. Journal of Tsinghua University (Science and Technology), 2018, 58(2): 137-142.
- [7] 王军辉,贾嵘,谭泊. 基于EEMD和模糊C均值聚类的风电机组齿轮箱故障诊断[J]. 太阳能学报, 2015, 36(2): 319-324.
Wang Jun-hui, Jia Rong, Tan bo. Fault diagnosis of wind turbine gearbox based on EEMD and fuzzy C means clustering[J]. Acta Energaie Solaris Sinica, 2015, 36(2): 319-324.
- [8] Wu Z, Huang N E. Ensemble empirical mode decomposition: a noise-assisted data analysis method[J]. Advances in Adaptive Data Analysis, 2009, 1(1): 1-41.
- [9] Mok P Y, Huang H Q, Kwok Y L, et al. A robust adaptive clustering analysis method for automatic identification of clusters[J]. Pattern Recognition, 2012, 45(8): 3017-3033.