

# 基于耦合隐马尔可夫的轴承故障诊断方法\*

夏裕彬<sup>1,2</sup>, 梁大开<sup>1</sup>, 郑国<sup>2</sup>, 王景霖<sup>2</sup>, 曾捷<sup>1</sup>

(1. 南京航空航天大学机械结构力学及控制国家重点实验室 南京, 210016)

(2. 上海航空测控技术研究所故障诊断与健康管理工作航空科技重点实验室 上海, 201601)

**摘要** 针对轴承故障信号比较微弱的特点,提出了一种基于耦合隐马尔可夫模型(coupled hidden Markov model, 简称 CHMM)的轴承故障诊断方法。首先,根据轴承传动结构特点,设计能够监测轴承振动状态的传感器网络;其次,通过非线性特征提取方法获取能够反映轴承健康状态的少数关键特征,利用传感信号的矢量量化代替提取其频域微弱特征的方法,可有效提高故障诊断效率和准确率;然后,在多通道信息融合中引入隐马尔可夫模型,采用左右型齐次隐马尔可夫链实现故障诊断;最后,通过对各种轴承故障状态构建其对应的耦合隐马尔可夫模型的方式,实现对轴承故障类型的辨识。试验结果表明,该方法能够有效地实现对轴承故障类型的识别。

**关键词** 非线性特征提取;矢量量化;耦合隐马尔可夫模型;故障诊断

**中图分类号** TH39;V216.3

## 引 言

现代工业及科学技术迅速发展造成了武器装备系统的结构随之越来越复杂,各系统之间的耦合程度也越来越深入,其使用环境也变的愈发恶劣。在这种状态下,偶然发生的一起故障可能引起系统之间的连锁反应,从而导致重大灾难性事故。因此,对武器装备运行的安全性和可靠性的监测要求也越来越高<sup>[1]</sup>。基于状态的维护因具有效率高、可操作性强等优点受到很高的关注,并运用于当今的故障诊断、预测与健康领域,极大地提高了武器装备的可监测性和使用安全性<sup>[2]</sup>。基于状态维护的关键技术是故障诊断和故障预测技术,其过程是根据监测到的武器装备的随机信号对其健康状况进行估计,具有一定的随机性。其既要描述装备不同健康状态之间变换的关系,还要根据测量过程来描述武器装备的健康状态。这样一种双重随机过程与隐马尔可夫模型(hidden Markov model, 简称 HMM)的处理过程相符。因此,隐马尔可夫被大量应用于故障诊断领域当中。

在 HMM 应用于故障诊断领域中,国内外学者们开展了一些研究工作。张西宁等<sup>[3]</sup>将 HMM 模型和主分量分析相结合,应用于轴承监测与故障诊断中。孙炎平等<sup>[4]</sup>提出利用经验模态分解和 HMM

方法在转盘故障诊断中,实现故障类型的识别与诊断精度的提高。王晓峰等<sup>[5]</sup>实现了利用隐马尔可夫的方法对汽轮机转子故障的诊断。Bunks 等<sup>[6]</sup>讨论了基于 HMM 的机械设备状态维护方法,并将其应用于 Westland 直升机齿轮箱数据的分类中。Ocak 等<sup>[7]</sup>将 HMM 应用于轴承的在线监测和故障诊断中。Bechhoefer 等<sup>[8]</sup>将 HSMM 用于轴的剩余寿命预测中,并用通用直升机的真实数据进行了验证。

为了确保航空装备安全可靠运行,目前多采用定期维修或维护的方式,由此带来了维护成本的大幅上升<sup>[9]</sup>。调查研究表明,故障诊断和预测技术是制约航空装备“视情维护”的主要因素<sup>[10]</sup>。笔者开展的非线性特征提取、故障诊断及故障预测技术,能快速准确地识别故障,并对故障进行有效预测,非常适合于故障样本数据难以获取、故障样本数据少的航空装备故障模式识别中,对提高航空装备的故障诊断能力具有一定的作用。

## 1 基本原理

### 1.1 时域同步平均

时域同步平均的实现过程如下:信号  $x(t)$  以采样频率  $f$  在起始时刻  $t$  开始采样,可以得到  $N$  段相同长度( $j$  个数据点)的数据段  $\{x_i(j)\}$ <sup>[11]</sup>

\* 国防基础科研计划资助项目(JCKY2016205A004);航空科学基金资助项目(20173333001)  
收稿日期:2018-01-08;修回日期:2018-06-30

$$x_i(j) = x\left(j \frac{1}{f_s} - t\right) \quad (1)$$

已知这些数据段为周期性的,即采样具有周期性,因此可认为这些数据段是同步的,对这些数据段进行平均可得到

$$y(j) = \frac{1}{N} \sum_{i=1}^N x_i(j) \quad (2)$$

其中: $y(j)$ 为时域同步平均信号。

## 1.2 小波包分解

经过时域同步平均处理信号之后,利用小波包对预处理之后的信号进行分解提取信号特征。小波包是由一系列小波函数  $\varphi^i(t)$  通过线性组合形成<sup>[12]</sup>,即

$$\varphi_{i,k}^j(t) = 2^{\frac{j}{2}} \varphi^i(2^j t - k) \quad (i=1,2,\dots) \quad (3)$$

其中: $i, j, k$  分别为频率因子、尺度因子和平移因子。

对于采集到的任何一个时域信号  $y(t)$  都可以进行如下分解

$$y(t) = \sum_{i=0}^{2^j-1} y_j^i(t) \quad (4)$$

其中: $y_j^i(t)$  为小波包第  $j$  层分解的第  $y$  个频带信号。

$y_{j+1}^i(t)$  具有如下关系<sup>[13]</sup>

$$y_j^i(t) = y_{j+1}^{2i-1}(t) + y_{j+1}^{2i}(t) \quad (5)$$

$$\begin{cases} y_{j+1}^{2i-1}(t) = H y_j^i(t) \\ y_{j+1}^{2i}(t) = G y_j^i(t) \end{cases} \quad (6)$$

其中: $\{h(k)\}, \{g(k)\}$  为尺度序列,为正交共轭低通和高通滤波器,即利用滤波的方式获取信号在不同频带范围类的分解信号。

通过对航空轴承振动信号的频谱分析,确定其特征信号所处的频段,由此可以确定小波包需进行分解的层数  $n$ , 共计  $2^n$  个。由于每一个分解信号与原始信号之间的关联程度不一样,因此利用核主元分析,以非线性处理方法对提取的分解信号进行约减获得主元特征向量。

## 1.3 隐马尔可夫

隐马尔可夫模型的一般形式为  $\lambda = (\mathbf{A}, \mathbf{B}, \boldsymbol{\pi})$ <sup>[14]</sup>。

1)  $\mathbf{A}$  为状态转换概率矩阵,矩阵中的每一个元素  $a_{ij}$  表示从状态  $i$  转换为状态  $j$  的概率

$$a_{(i,d),(j,d)} \stackrel{\text{def}}{=} P[S_{[t+1,t+d]} = j \mid S_{[t-d+1,t]} = i] \quad (7)$$

其中: $d$  和  $d'$  分别为状态转移的时间跨度; $S_{[\cdot,\cdot]}$  表示当前时刻的状态; $P[\cdot]$  表示由状态  $i$  转移到状态  $j$  的概率。

2)  $\mathbf{B}$  为观察概率矩阵,矩阵中每一个元素  $b$  表示在状态  $S_j$  下的特定观察值的概率

$$b_{j,d}(o_{t+1,t+d}) \stackrel{\text{def}}{=} P[o_{t+1,t+d} \mid S_{[t+1,t+d]} = j] \quad (8)$$

3)  $\boldsymbol{\pi}$  为初始时刻的状态向量的概率分布

$$\pi_{j,d} \stackrel{\text{def}}{=} P[S_{[t-d+1,t]} = j] \quad (9)$$

在 HMM 模型各主要参数确定之后即可完成其模型构建,同时将两条 HMM 链通过耦合算法耦合在一起组成包含两条马尔科夫链的耦合隐马尔可夫模型 CHMM。利用 K-means 聚类的优点,通过 Viterbi 算法和 K-means 聚类算法来对 CHMM 模型的构造参数进行初始化,再在模型训练中通过 Baum-Welch 算法对模型参数进行优化,以使最大似然概率达到最优。参数重估步骤<sup>[15]</sup>如下。

1) 根据实际应用情况确定状态数据  $N$  和观测值的个数  $M$ 。

2) 根据前向-后向算法计算前向变量  $\alpha_t(i)$  和后向变量  $\beta_{t+1}(i)$ ,进而计算  $\xi_t(i, j)$  和  $\gamma_t(i)$

$$\xi_t(i, j) = \frac{\alpha_t(i) a_{ij} b_j(o_{t+1}) \beta_{t+1}(j)}{\sum_{i=1}^N \sum_{j=1}^N \alpha_t(i) a_{i,j} b_j(o_{t+1}) \beta_{t+1}(j)} \quad (10)$$

$$\gamma_t(i) = \frac{\alpha_t(i) \beta_t(i)}{\sum_{j=1}^N \alpha_t(j) \beta_t(j)} = \sum_{j=1}^N \xi_t(i, j) \quad (11)$$

3) 参数重估为

$$\begin{cases} \bar{\pi}_i = \gamma_1(i) \\ \bar{a}_{ij} = \frac{\sum_{t=1}^{T-1} \xi_t(i, j)}{\sum_{t=1}^{T-1} \gamma_t(i, j)} \\ \bar{b}_j(k) = \frac{\sum_{t=1}^T \gamma_t(j) \delta(o_t, v_k)}{\sum_{t=1}^T \gamma_t(j)} \end{cases} \quad (12)$$

其中: $r_1(i)$  为初始状态的次数; $\xi$  为状态转移的次数; $r$  为该状态的次数; $\delta$  为某一状态下观测值为  $v$  的次数。

4) 将步骤 3 中的参数  $\bar{\lambda} = (\bar{\boldsymbol{\pi}}, \bar{\mathbf{A}}, \bar{\mathbf{B}})$  代入原模型中,重复上述过程,直至似然  $P(O|\bar{\lambda})$  收敛。

在数据融合过程中可能会出现个别通道数据缺失的状况,这就需要隐马尔可夫模型具有自适应推理能力,其过程如下:

1) 前向变量初始化,缺失通道  $X$  数据时

$$\alpha_1(i) = P(o_1^X, q_1 = S_i \mid \lambda) = \pi_i b_{i,X}^X(o_1^X) \quad (13)$$

2) 递推  $t$  时刻

$$\begin{aligned} \alpha_t(j) &= P(o_{1,t-1}, o_t^X, q_t = S_j \mid \lambda) = \\ &= \sum_i P(o_{1,t-1}, o_t^X, q_{t-1} = S_i, q_t = S_j \mid \lambda) = \\ &= \sum_i \alpha_{t-1}(i) a_{i,j} b_{j,X}^X(o_t^X) \end{aligned} \quad (14)$$

3) 终止

$$P(o \mid \lambda) = \sum_i P(o_{1:T}, q_t = S_i) = \sum_i \alpha_T(i) \quad (15)$$

## 2 诊断系统

隐马尔可夫模型能通过较少的样本数据获得较为可靠评估模型,但在实际应用过程中也存在着故障特征参数选择、初始模型参数初始化、算法下溢及多观测值序列训练等方面较为难处理的问题。

针对机械故障诊断存在的上述问题,开展相应的体系化研究,利用非线性特征提取方法建立特征集,构建包含驻留时间状态、初始模型参数选取以及适用多通道信息融合的耦合隐马尔可夫模型,如图 1 所示。

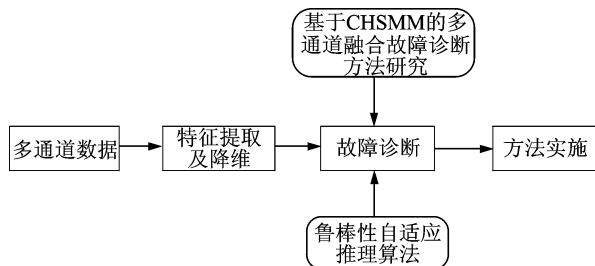


图 1 诊断方法示意图

Fig. 1 The diagram of diagnosis method

验证试验以动力传动系统故障诊断综合试验台为基础,试验平台由变速驱动电机、二级行星齿轮箱、二级齿轮传动齿轮箱、可编程磁力制动器、转速传感器及 IMC 数据采集器等构成,其外形结构如图 2 所示。整个试验台被固定在台架上,由变速驱动电机提供动力。驱动电机和磁力制动器由相应调节器来控制,用以变更不同转速及负载。动力传动系统故障诊断综合试验台采用模块化设计,可快速方便地更换齿轮和轴承部件,能模拟轴承故障(内圈故障、外圈故障及滚动体故障等),因此,该试验设备能满足本次试验目的的需求。

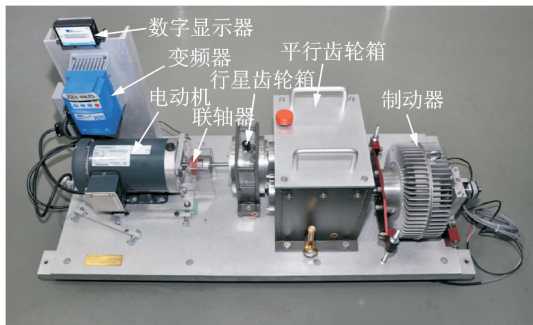


图 2 动力传动系统故障诊断综合试验台外形图

Fig. 2 The experiment system for power transmission fault diagnosis

## 3 诊断结果

获取试验数据后,将试验数据按故障分为 5 种

类别:正常状态、滚动体故障、外圈故障、内圈故障和联合故障(内圈与外圈均有故障)。将每一类数据进行预处理,根据小波包处理后获取其对应的主元特征向量,对其进行频谱分析。假定频谱分析得到一个包含 512 个元素的向量;以每 64 个元素为一帧,前向移动 16 个元素,则每个频谱信号可得到 29 帧向量;将其再做矢量量化,将每一帧信号转换成 24 个点的矢量特征(无量纲量)。其过程如图 3~图 5 所示:图 3 为普通状态下原始信号的振动幅度与点数之间的关系;图 4 为图 3 信号对应的频谱分析与点数之间的关系;图 5 为矢量量化后的离散数列图。

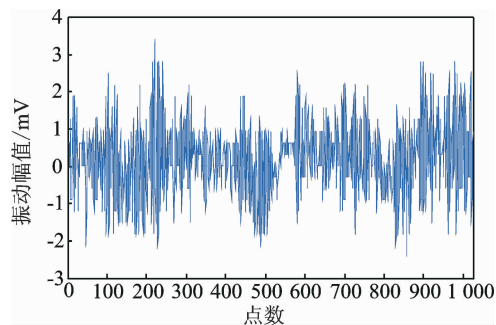


图 3 原始工况图

Fig. 3 Original signal map

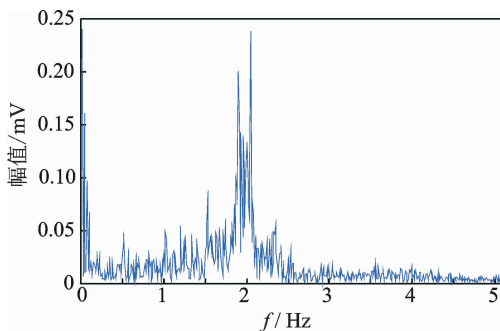


图 4 原始信号频谱分析

Fig. 4 Frequency spectrum of original signal

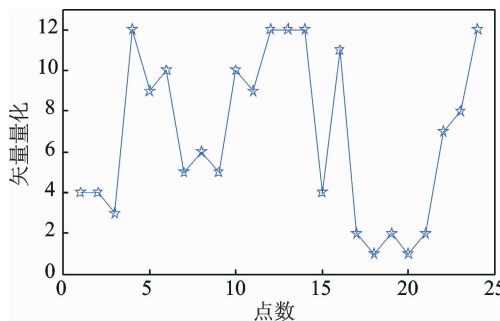


图 5 矢量量化后的离散数列

Fig. 5 Discrete series by vector-quantization

经过矢量量化后的特征被分为两类:a. 作为训练 CHMM 模型的训练数据,其数量占矢量量化特

征总数的 2/3;b. 用作测试模型准确度的测试数据, 为非训练数据。建立 5 个 CHMM 架构模型, 分别用 5 种对应轴承状态的测试数据的特征向量训练建立好的 5 个 CHMM 模型, 可得到分别对应 5 种状态的 CHMM 模型, 最后利用测试数据进行测试, 其识别结果(无量纲)如图 6~图 10 所示。

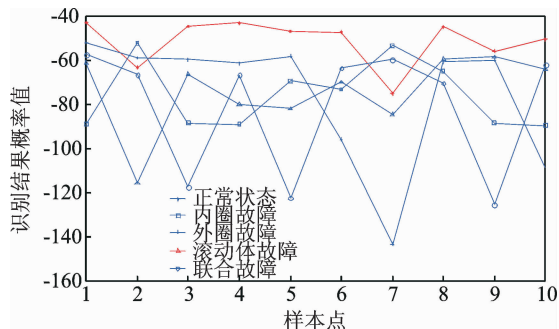


图 6 正常状态识别结果

Fig. 6 Identification result of normal model

图 6 为将 5 组测试数据输入进正常状态对应的 CHMM 模型, 其中红线代表的是正常状态的测试数据得到的输出结果。

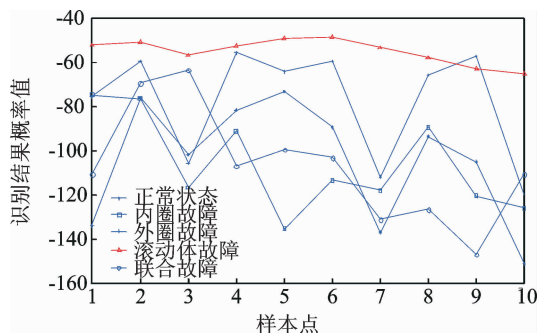


图 7 滚动体故障识别结果

Fig. 7 Identification result of ball-fault

图 7 为将 5 组测试数据输入进滚动体故障状态对应的 CHMM 模型, 其中红线代表的是滚动体故障的测试数据得到的输出结果。

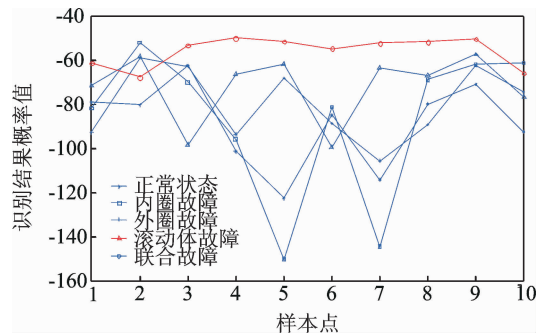


图 8 联合故障识别结果

Fig. 8 Identification result of union-fault

图 8 为将 5 组测试数据输入进联合故障状态对应的 CHMM 模型, 其中红线代表的是联合故障状态的测试数据得到的输出结果。

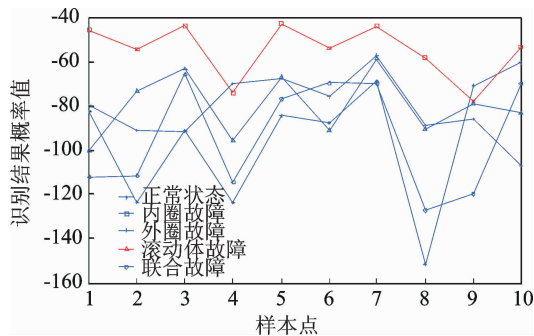


图 9 内圈故障识别结果

Fig. 9 Identification result of inner-fault

图 9 为将 5 组测试数据输入进内圈故障状态对应的 CHMM 模型, 其中红线代表的是内圈故障状态的测试数据得到的输出结果。

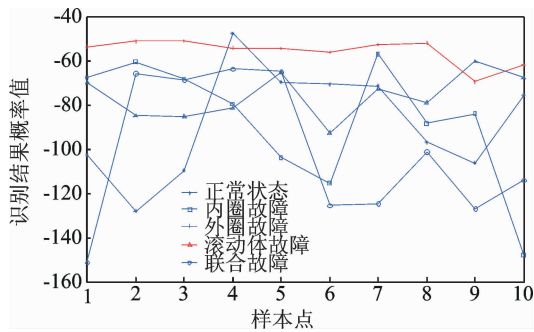


图 10 外圈故障识别结果

Fig. 10 Identification result of outer-fault

图 10 为将 5 组测试数据输入进外圈故障状态对应的 CHMM 模型, 其中红线代表的是外圈故障状态的测试数据得到的输出结果。

### 4 结束语

笔者提出的基于耦合隐马尔可夫模型的故障诊断技术, 利用非线性特征提取方法对轴承在不同健康状态下的振动信号进行特征提取, 获取与健康状态存在联系的表征特征参数, 进而采用多通道信息融合的耦合隐马尔可夫模型来构建表征参数与健康状态之间的映射关系, 实现以振动监测来标定轴承的健康状态以及故障类型的识别。

### 参 考 文 献

[1] Tian Zhigang, Wang W. Special Issue on machine fault diagnostics and prognostics[J]. Chinese Journal

- of Mechanical Engineering, 2017(30):1283-1284.
- [2] 莫固良,汪慧云,李兴旺,等. 飞机健康监测与预测系统的发展及展望[J]. 振动、测试与诊断, 2013, 33(6): 925-929.  
Mo Guliang, Wang Huiyun, Li Xingwang, et al. Health monitoring and prognostics system for aircraft [J]. Journal of Vibration, Measurement & Diagnosis, 2013, 33(6): 925-929. (in Chinese)
- [3] 张西宁,雷威,李兵. 主分量分析和隐马尔科夫模型结合的轴承监测诊断方法[J]. 西安交通大学学报, 2017, 51(6):1-7.  
Zhang Xining, Lei Wei, Li Bing. Bearing fault detection and diagnosis method based on principal component analysis and hidden Markov model[J]. Journal of Xi'an Jiaotong University, 2017, 51(6):1-7. (in Chinese)
- [4] 孙炎平,陈捷,洪荣晶,等. 基于 EMD-HMM 的转盘轴承故障诊断方法[J]. 轴承, 2017(1):41-45.  
Sun Yanping, Chen Jie, Hong Rongjing, et al. Application of empirical model decomposition and hidden Markov model in fault diagnosis of slewing bearing [J]. Bearing, 2017(1):41-45. (in Chinese)
- [5] 王晓峰,夏静,韩捷,等. 基于隐马尔可夫模型的汽轮机故障诊断方法研究[J]. 中国工程机械学报, 2016, 14(6):541-544.  
Wang Xiaofeng, Xia Jing, Han Jie, et al. Fault diagnosis on turbine generators via hidden markov model[J]. Chinese Journal of Construction Machinery, 2016, 14(6): 541-544. (in Chinese)
- [6] Bunks C, Mccarthy D, Al-Ani T. Condition-based maintenance of machines using hidden Markov models [J]. Mechanical Systems and Signal Processing, 2000, 14(4):597-612.
- [7] Ocak H, Loparo K A. HMM-based fault detection and diagnosis scheme for rolling element bearings [J]. Journal of Vibration and Acoustics, 2005, 127(4):299-306.
- [8] Bechhoefer E, Bernhard A, He D. Use of hidden semi-Markov models in the prognostics of shaft failure [C] // American Helicopter Society 62th Annual Forum. Phoenix, USA: [s. n. ], 2006.
- [9] 吴明辉,阮永贵,韩海舰. 机载设备视情维修及其决策建模分析[J]. 国外电子测量技术, 2015, 34(3):27-31.  
Wu Minghui, Ruan Yonggui, Han Haijian. Analysis of on condition maintenance and decision modeling for airborne equipment [J]. Foreign Electronic Measurement Technology, 2015, 34(3):27-31. (in Chinese)
- [10] 孙旭升,周刚,于洋,等. 机械设备故障预测与健康管理综述[J]. 兵工自动化, 2016, 35(1):30-33.  
Sun Xusheng, Zhou Gang, Yu Yang, et al. Overview of prognosis and health management of mechanical equipment [J]. Ordnance Industry Automation, 2016, 35(1):30-33. (in Chinese)
- [11] 沈勇,翟秀梅,李兴旺. 时域同步平均技术在直升机主减速器故障诊断中的应用分析[J]. 计算机测量与控制, 2014, 22(11):3473-3475.  
Shen Yong, Zhai Xiumei, Li Xingwang. Time synchronous averaging (TSA) technology applied in helicopter main reducer fault diagnosis analysis [J]. Computer Measurement & Control, 2014, 22(11): 3473-3475. (in Chinese)
- [12] 彭畅,柏林,刘小峰. 基于鲁棒性小波包峭度图的滚动轴承故障诊断[J]. 振动、测试与诊断, 2016, 36(1): 11-16.  
Peng Chang, Bo lin, Liu Xiaofeng. Rolling bearing fault diagnosis based on robust wavelet packet kurtosis diagram [J]. Journal of Vibration, Measurement & Diagnosis, 2016, 36(1):11-16. (in Chinese)
- [13] 张培林,李胜. 基于小波包变换和 GA-PLS 算法的故障特征选择方法[J]. 振动、测试与诊断, 2014, 34(2): 385-391.  
Zhang Peilin, Li Sheng. Fault feature selection method based on wavelet and GA-PLS algorithm [J]. Journal of Vibration, Measurement & Diagnosis, 2014, 34(2):385-391. (in Chinese)
- [14] 朱晓东,袁坤杰,王艳玲. 基于隐马尔可夫模型的石油钻井事故预警方法[J]. 郑州大学学报, 2018, 39(4): 51-57.  
Zhu Xiaodong, Yuan Kunjie, Wang Yanling. An accident prewarning method of oil drilling based on hidden Markov model [J]. Journal of Zhengzhou University, 2018, 39(4):51-57. (in Chinese)
- [15] 季云,王恒,朱龙彪,等. 基于 DPMM-CHMM 的机械设备性能退化评估研究[J]. 振动与冲击, 2017, 36(23): 170-174.  
Ji Yun, Wang Heng, Zhu Longbiao, et al. Performance degradation assessment for mechanical equipment based on DPMM- CHMM [J]. Journal of Vibration and Shock, 2017, 36(23):170-174. (in Chinese)



**第一作者简介:**夏裕彬,男,1963年5月生,研究员。上海市“领军人才”,享受国务院特殊津贴专家。国防科工委“511人才工程”高级经营管理人才,“国防科技工业有突出贡献的中青年专家”,故障诊断与健康管理技术航空科技重点实验室主任,《振动、测试与诊断》副主编。主要从事航空测试、诊断、预测与健康管理等技术研究,先后作为主要负责人主持了多项航空预先研究项目和装备研究工作。曾发表《航空故障诊断与健康管理技术》等论著。获得国防科技进步奖5项,部级科技进步奖10余项。

E-mail: xiayubin@vip. sina. com