

# 基于集成隐马尔可夫模型的轴承故障诊断

蒋会明, 陈进, 董广明, 刘韬

(上海交通大学 机械系统与振动国家重点实验室, 上海 200240)

**摘要:** 对轴承振动信号进行时频分析获得全特征集; 运用距离补偿法提取轴承故障敏感特征获得敏感特征集。两种特征集在用于训练、测试轴承状态时不仅诊断率不同, 且误判样本亦不同。基于此, 提出基于集成隐马尔可夫模型的轴承故障诊断方法。采用两种特征集分别建立两独立隐马尔可夫模型; 运用平均法则、最大似然概率法集成隐马尔可夫模型分类效果; 对轴承信号进行故障诊断。实验结果表明, 与基于敏感特征集、全特征集的分类器相比, 该模型分类器在轴承故障诊断中识别精度更高。

**关键词:** 轴承故障诊断; 补偿距离评估技术; 隐马尔可夫模型

**中图分类号:** TH165.3

**文献标志码:** A

**DOI:** 10.13465/j.cnki.jvs.2014.10.018

## Integrated HMM-based bearing fault diagnosis

JIANG Hui-ming, CHEN Jin, DONG Guang-ming, LIU Tao

(State Key Laboratory of Mechanical System and Vibration, Shanghai Jiao Tong University, Shanghai 200240, China)

**Abstract:** Full features of a bearing vibration signal in time and frequency domain were extracted at first. A compensation method based on distance was used to choose features sensitive to bearing faults. Then full features and sensitive features vectors were built. The results using hidden markov model (HMM) based on those two features were different. Then the method of integrated HMM for bearing fault diagnosis was proposed. Based on independent HMM classifiers trained with those two different feature vectors, the average rule and the maximum likelihood probability method were used to integrate the two HMM classifiers. The experimental results showed that the proposed method has a higher recognition rate compared with the two independent classifiers based on different feature vectors.

**Key words:** bearing fault diagnosis; evaluation technique based on compensation distance; hidden markov model (HMM)

随科学技术进步及现代工业飞速发展, 机械设备逐步向大型、复杂、高速、高效及重载方向发展, 工作环境也更复杂、苛刻。若设备的关键部件发生故障, 很有可能发生连锁反应, 导致整台设备甚至整个生产过程被破坏, 造成重大经济损失, 甚至导致人员伤亡。因此, 作为旋转机械的轴承一直是设备状态监测与故障诊断领域热门研究对象<sup>[1]</sup>。

故障诊断关键为特征提取。对轴承故障, 通常由时域、频域、时频域提取多维特征, 以期更好表征轴承状态。由于各特征对故障模式表征能力不同, 通常用一定方法对敏感特征进行提取。周徐宁<sup>[2]</sup>用补偿距离评估技术提取敏感特征, 验证敏感特征在故障识别中的优越性。Rabinerd 等<sup>[3-4]</sup>建立的隐马尔可夫模型

(HMM) 为时间序列信号统计分析模型, 适用于随机过程时间序列统计建模, 尤其非平稳、重复再现性不佳的时间序列信号。并在语音识别、手写识别<sup>[5]</sup>、齿轮箱状态识别<sup>[6]</sup>、轴承故障诊断及性能退化评估<sup>[7-8]</sup>等领域的应用一直为研究热点。本文提出基于集成隐马尔可夫模型的轴承故障诊断方法。该方法提取轴承振动信号常用时、频域特征, 采用补偿距离技术提取敏感特征; 通过利用敏感特征集、全特征集训练 HMM 发现的两种模型分类效果不同不仅表现在识别率上, 且误判样本亦不同。尝试采用平均法则<sup>[5]</sup>、最大似然概率法决定样本所属故障类型; 通过对滚动轴承实验数据分析, 验证所提方法的有效性。

## 1 理论基础

### 1.1 特征提取

工程中用于监测轴承运行状态的时域指标有峰值、峰峰值、平均幅值、方根幅值、有效值、波形指标、脉

冲指标、峰值指标、裕度指标、歪度指标及峭度指标。轴承出现故障时,其信号内部特征及状态变化常体现在频率组成结构变化中,包括频率成分及其幅值变化。幅值谱熵度量信号频率分布均匀程度体现信号能量分布的频域复杂性。幅值谱熵计算式<sup>[9]</sup>为

$$\left. \begin{aligned} SE &= - \sum_{i=1}^N p_i \log(p_i) \\ p_i &= X(i) / \sum_{j=1}^N X(j) \end{aligned} \right\} \quad (1)$$

式中:  $\sum_{i=1}^N p_i = 1$ ;  $SE$  为谱熵;  $X(i)$  为序列  $\{x(i), i = 1, 2, \dots, N\}$  频率谱线幅值。

由于滚动轴承振动信号通常具有非平稳特性,为更准确提取信号中所含状态特征,采用时频分析及申波<sup>[10]</sup>小波能量谱熵概念,对时频信号能量分布进行定量描述。先对原信号进行连续小波变换,计算各尺度能量值  $E_i$ 。定义小波能谱熵为

$$\left. \begin{aligned} WE &= - \sum_{i=1}^N p_i \log(p_i) \\ p_i &= E_i / \sum_{j=1}^N E_j \end{aligned} \right\} \quad (2)$$

式中:  $\sum_{i=1}^N p_i = 1$ ;  $N$  为连续小波分解最大尺度;  $WE$  为小波能谱熵。

## 1.2 补偿距离评估技术

由于提取的特征对故障模式敏感度不同,不同特征在分类故障时所做贡献不同。因此采用有效特征选择方法 - 补偿距离评估技术对特征进行有效提取,选其中敏感性高的特征向量构成用于分类的敏感特征集<sup>[2]</sup>。设含  $C$  个模式类  $\omega_1, \omega_2, \dots, \omega_c$  的特征集为

$$\left\{ f_{c,m,k}, c = 1, 2, \dots, C; m = 1, 2, \dots, M_c; k = 1, 2, \dots, K \right\} \quad (3)$$

式中:  $f_{c,m,k}$  为在  $c$  状态下第  $m$  个样本第  $k$  个特征;  $M_c$  为  $c$  状态下样本总数;  $K$  为样本特征数。

基于补偿距离评估技术的特征选择步骤为:

(1) 计算  $\omega_c$  类中所有特征向量的平均距离:

$$d_{c,k} = \frac{1}{M_c(M_c - 1)} \sum_{i=1}^{M_c} \sum_{j=1}^{M_c} |f_{c,i,k} - f_{c,j,k}| \quad (4)$$

对  $d_{c,k} (c = 1, 2, \dots, C)$  求平均后得平均类内距离:

$$d_k^w = \frac{1}{C} \sum_{c=1}^C d_{c,k} \quad (5)$$

(2) 定义并计算  $d_k^w$  的方差因子:

$$v_k^w = \frac{\max(d_{c,k})}{\min(d_{c,k})} \quad (6)$$

(3) 计算  $C$  个模式类间距离:

$$d_k^b = \frac{1}{C(C-1)} \sum_{c=1}^C \sum_{e=1}^C |u_{e,k} - u_{c,k}| \quad (7)$$

式中:  $u_{c,k} = \frac{1}{M_c} \sum_{m=1}^{M_c} f_{c,m,k}$  为  $\omega_c$  类中所有第  $k$  个特征均值。

(4) 定义并计算  $d_k^b$  的方差因子:

$$v_k^b = \frac{\max(|u_{e,k} - u_{c,k}|)}{\min(|u_{e,k} - u_{c,k}|)} \quad (8)$$

$(c, e = 1, 2, \dots, C; c \neq e)$

(5) 定义并计算补偿因子:

$$\gamma_k = \frac{1}{\frac{v_k^w}{\max(v_k^w)} + \frac{v_k^b}{\max(v_k^b)}} \quad (9)$$

(6) 计算类间距  $d_k^b$  与类内距  $d_k^w$  的比值,得距离评估指标:

$$\alpha_k = \gamma_k \frac{d_k^b}{d_k^w} \quad (10)$$

对  $\alpha_k$  进行归一化处理,得补偿距离评估指标:

$$\bar{\alpha}_k = \frac{\alpha_k}{\max(\alpha_k)} \quad (11)$$

(7) 设定阈值  $\phi (\phi \in [0, 1])$ , 从特征集  $\{f_{c,m,k}\}$  中选距离评估指标  $\bar{\alpha}_k \geq \phi$  的特征作为敏感特征。

类间距离越大,类内距离越小,特征分类效果越明显,此为补偿距离评估技术本质。

## 1.3 隐马尔可夫模型

隐马尔可夫模型为双随机模型。不仅各状态间转换随机,且每个状态下所得观测值亦随机。其中,描述状态转移的马尔可夫过程一般无法直接观测到,而只能据观测值推测。故该模型称“隐”马尔可夫模型,其描述<sup>[3]</sup>为:

(1)  $N$  为模型状态数目。记模型  $N$  个状态分别表示  $S_1, S_2, \dots, S_N, t$  时刻模型所处状态记作  $q_t$ , 则  $q_t \in \{S_1, S_2, \dots, S_N\}$ 。

(2)  $M$  为每个状态下对应的观测值取值数。将  $M$  个观测值记为  $v_1, v_2, \dots, v_M, t$  时刻观测值为  $o_t$ , 则  $o_t \in \{v_1, v_2, \dots, v_M\}$ 。

(3)  $\pi$  为初始状态概率向量:

$$\pi = \{\pi_i\}, \pi_i = P(q_1 = S_i) \quad (12)$$

$(1 \leq i \leq N)$

(4)  $A$  为状态转移概率矩阵:

$$A = \{a_{i,j}\}, a_{i,j} = P(q_{t+1} = S_j | q_t = S_i) \quad (13)$$

$(1 \leq i, j \leq N)$

(5)  $B$  为观测值概率矩阵:

$$b_j(k) = P(o_t = v_k | q_t = S_j) \quad (14)$$

$(1 \leq j \leq N, 1 \leq k \leq M)$

通常将一个 HMM 简记为

$$\lambda = (\pi, A, B) \quad (15)$$

以上为观测值假设成离散值情况。工程中所遇绝大多数问题均为连续观测值。在连续 HMM 中,用高斯

混合模型模拟各状态下观测值概率密度函数为

$$b_j(o_i) = \sum_{m=1}^{M_j} w_{j,m} G(o_i, \mu_{j,m}, \delta_{j,m}) \quad (16)$$

式中:  $M_j$  为  $S_j$  状态下高斯分量数;  $w_{j,m}$  为  $S_j$  状态时第  $m$  个高斯分量权重;  $\mu_{j,m}, \delta_{j,m}$  分别为状态  $S_j$  时第  $m$  个高斯分量均值及协方差。

HMM 训练、诊断时常采用 Baum-Welch 算法完成模型参数估计, 前向-后卷算法可有效计算模型输出概率, Viterbi 算法用于寻找测试时间序列最优状态序列。

#### 1.4 集成隐马尔可夫轴承诊断模型

轴承振动信号特征可表征轴承故障状况, 从信号中提取时频域特征组成特征集  $\{X_1(n)\}$ 。采用补偿距离评估技术, 对特征集进行敏感特征提取获得  $\{X_2(n)\}$ 。据两种不同特征集训练模型可得两种类型分类器  $e_i (i=1, 2)$ 。有  $N$  种故障类型时, 每个  $e_i$  有  $N$  个 HMM 模型记作  $HMM_i^k$ ,  $k$  表示轴承故障类别,  $i$  表示分类器号。本文采用平均法则、最大似然概率法决定测试样本所属轴承故障类型。具体步骤如下:

(1) 对未知状态测试样本  $x$ , 获得  $\{X_1(n)\}, \{X_2(n)\}$  两个特征集。

(2) 将信号两特征集  $X_i$  依次代入两分类器  $e_i$  中, 获得两  $N$  维似然概率向量  $\{p_i^k\} (1 \leq k \leq N)$ 。

(3) 采用平均集成法则及最大似然概率法确定信号代表的故障类型<sup>[5]</sup>:

$$P_i^E = \sum_{k=1}^N P_i^k \quad (17)$$

$$\varphi = \arg \max_{1 \leq i \leq 2} P_i^E \quad (18)$$

$$\text{fault} = \arg \max_{1 \leq k \leq N} P_i^k \quad (19)$$

则测试轴承属于 fault 表征的故障类型。

## 2 轴承故障测试实验

为验证所提方法的有效性, 对故障轴承进行振动信号采集实验。实验装置见图 1。图中, 转子两端分别由支撑装置及试验轴承支撑。试验台自带液压定位与加紧装置, 用于固定测试轴承外圈。试验轴承选 GB203 单列深沟球轴承, 通过电火花在轴承表面加工单点点蚀, 获得外圈故障、滚动体点蚀、内圈故障。与正常轴承同进行 4 种状态模拟实验。轴承工作轴转速 720 r/min, 采样频率 25.6 kHz, 采样长度 8 s。传感器安装位置见图 2。

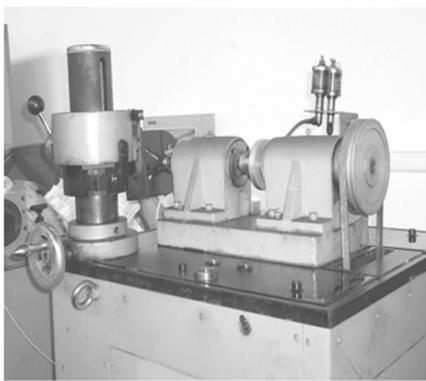


图1 滚动轴承试验台

Fig. 1 Rolling element bearing test rig

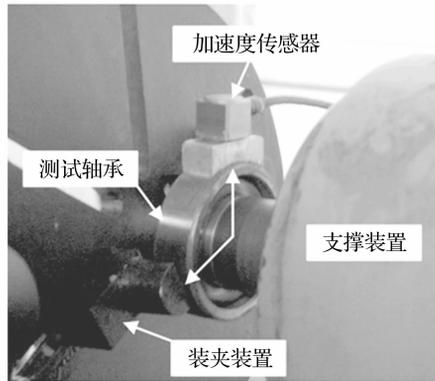


图2 传感器安装位置

Fig. 2 Location of accelerometer

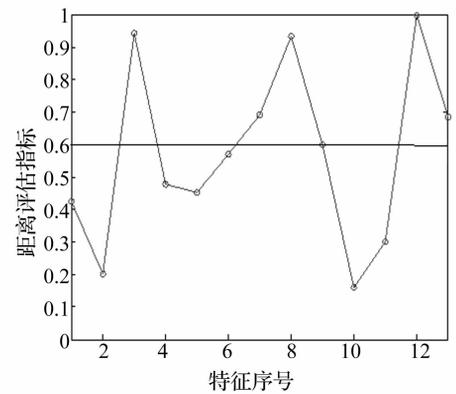


图3 各种特征距离评估指标

Fig. 3 Distance evaluation criteria of all features

对数据进行预处理, 1 024 点为一段, 共 200 段。在每段上提取时域特征、小波能量熵、幅值谱熵, 获得全特征集、敏感特征集。进行小波分析时, 选 db3 对信号进行 4 层小波包分解。对所有特征进行敏感特征选择。各特征距离评估指标见图 3。图 3 中横坐标序号代表各特征依次为峰值、峰峰值、平均幅值、方根幅值、有效值、波形指标、脉冲指标、峰值指标、裕度指标、歪度指标、峭度指标、幅值谱熵及小波能谱熵; 纵坐标为各特征补偿距离评估指标。综合考虑特征敏感程度及聚类效果, 设阈值为 0.6<sup>[2]</sup>, 提取敏感特征为平均幅值、脉冲指标、峰值指标、裕度指标、小波能谱熵及幅值谱熵。

每类轴承数据各有 200 组特征, 将其中 150 组训练 HMM, 余 50 组测试模型诊断率。用全特征集训练

HMM 后诊断结果见图 4。其中, 图 4(a) 为正常轴承振动信号在 4 种故障类型全特征集 HMM 中诊断结果, 17、18 号样本被误判为外圈故障; 图 4(b) 为轴承外圈故障振动信号在 4 种故障类型全特征集 HMM 中诊断结果, 32 号样本被误判为内圈故障; 图 4(c) 为轴承滚动体故障振动信号在 4 种故障类型全特征集 HMM 中诊断结果良好, 无误判样本; 图 4(d) 为轴承内圈故障振动信号在 4 种故障类型全特征集 HMM 中诊断结果, 17、22、33、44 号样本被误判为外圈故障; 正确分类综合诊断率为 96.5%。用敏感特征集训练 HMM 后所得诊断结果见图 5。其中图 5(a) 为正常轴承振动信号在 4 种故障类型敏感特征集 HMM 中诊断结果, 17、18 号样本被误判为外圈故障; 图 5(b) 为轴承外圈故障振动信

号在 4 种故障类型敏感特征集 HMM 中诊断结果,15、24、31、32 号样本被误判为内圈故障;图 5(c)为轴承滚动体故障振动信号在 4 种故障类型敏感特征集 HMM 中诊断结果,结果良好,无误判样本;图 5(d)为轴承内圈故障振动信号在 4 种故障类型敏感特征集 HMM 中诊断结果,2、3、4、26、49 号样本被误判为外圈故障;正确分类综合诊断率为 94.5%。采用平均集成法则、最大似然概率法集成 HMM 分类效果,测试结果见图 6。其中图 6(a)为正常轴承振动信号在 4 种故障类型集成 HMM 中诊断结果,17、18 号样本被误判为外圈故障;图 6(b)为轴承外圈故障振动信号在 4 种故障类型集成

HMM 中诊断结果,24 号样本被误判为内圈故障;图 6(c)为轴承滚动体故障振动信号在 4 种故障类型集成 HMM 中诊断结果,结果良好,无误判样本;图 6(d)为轴承内圈故障振动信号在 4 种故障类型集成 HMM 中诊断结果,33、44 号样本被误判为外圈故障;正确分类综合诊断率为 97.5%。

全部诊断结果见表 1。在诊断正常、滚动体故障时,三种分类器诊断率相同;对外、内圈诊断,集成 HMM 集合全特征 HMM 及敏感特征 HMM 性能,获得更好诊断效果。因此,集成 HMM 在诊断轴承故障时较全特征、敏感特征 HMM 更有效。

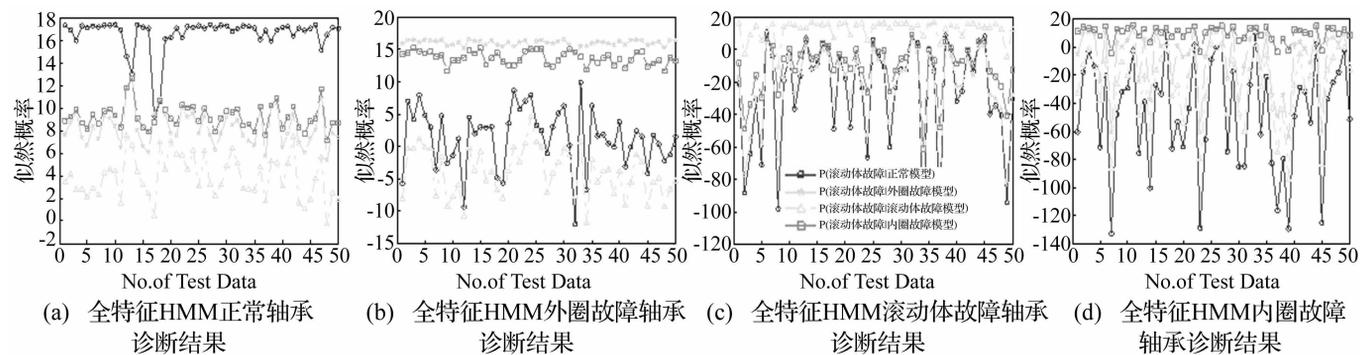


图 4 全特征 HMM 诊断结果

Fig. 4 Diagnosis results based on full features HMM

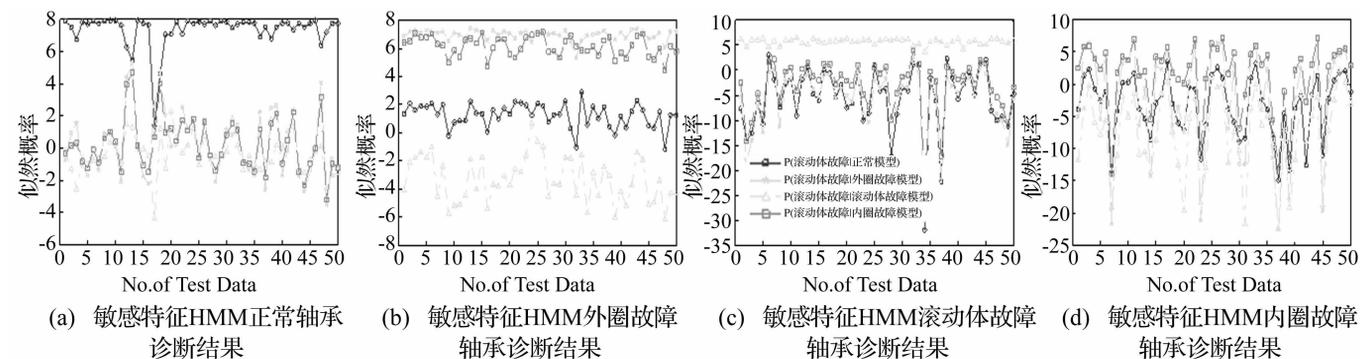


图 5 敏感特征 HMM 诊断结果

Fig. 5 Diagnosis results based on sensitive features HMM

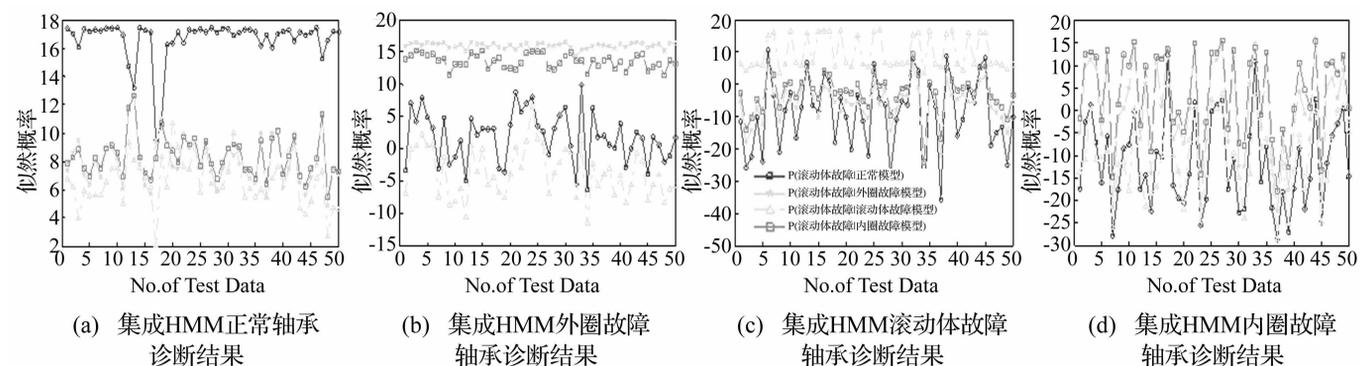


图 6 集成 HMM 诊断结果

Fig. 6 Diagnosis results based on integrated HMM

表1 全特征、敏感特征诊断结果

Tab.1 Diagnosis results of full and sensitive features HMM

故障类型	正常	外圈	滚动体	内圈	诊断率/%
全特征 HMM 误判样本号	17,18	32	无	17,22,33,44	96.5
敏感特征 HMM 误判样本号	17,18	15,24,31,32	无	2,3,4,26,49	94.5
集成 HMM 误判样本号	17,18	24	无	33,44	97.5

### 3 结论

(1) 本文鉴于用全特征、敏感特征表征轴承故障状态分类效果在诊断率与误判样本的不同,提出基于平均集成法则与最大似然概率法的集成 HMM 轴承故障诊断方法。

(2) 因全特征所含信息多,能全面表征轴承状态;敏感特征类间距离大,更有利于分类。但两特征集时间序列因所含特征维数不同,对轴承状态有各自特点。

(3) 集成 HMM 能更好表征轴承故障状态,并可集中全特征 HMM 及敏感特征 HMM 性能,降低误判个数,提高诊断性能。因此,本文所提方法可在轴承故障诊断中应用,亦为 HMM 在故障诊断应用提供新的研究思路。

#### 参考文献

[1] 陈进. 机械设备振动监测与故障诊断[M]. 上海:上海交通大学出版社,1999.

[2] 周徐宁. 基于特征加权连续隐马尔可夫模型的智能故障诊断方法研究[D]. 上海:上海交通大学,2012.

[3] Rabiner L R. A tutorial on Hidden Markov Models and selected applications in speech recognition[J]. Proceedings

of the IEEE,1989,77(2):257-286.

- [4] Rabiner L R, Juang B H. An introduction to hidden Markov models [J]. IEEE ASSP magazine, 1986, 3(1):4-16.
- [5] Liu Cai-feng, Tian Xue-dong, Yang Fang. Recognition of offline handwritten Chinese characters of amount in words based on integrated features and HMM [J]. Applied Mechanics and Materials, 2013, 263/266:2639-2642.
- [6] 张锐戈, 谭永红. 基于最优 Morlet 小波和隐马尔可夫模型的轴承故障诊断[J]. 振动与冲击, 2012, 31(12):5-8. ZHANG Rui-ge TAN Yong-hong. Fault diagnosis of rolling element bearings based on optimal morlet wavelet and hidden markov model [J]. Journal of Vibration and Shock, 2012, 31(12):5-8.
- [7] 滕红智, 赵建民, 贾希胜, 等. 基于 CHMM 的齿轮箱状态识别研究[J]. 振动与冲击, 2012, 31(5):92-96. TENG Hong-zhi, ZHAO Jian-min, JIA Xi-sheng, et al. Gearbox state recognition based on continuous hidden markov model [J]. Journal of Vibration and Shock, 2012, 31(5):92-96.
- [8] 肖文斌, 陈进, 周宇, 等. 小波包变换和隐马尔可夫模型在轴承性能退化评估中的应用[J]. 振动与冲击, 2011, 30(8):32-35. XIAO Wen-bin, CHEN Jin, ZHOU Yu, et al. Wavelet packet transform and hidden Markov model based bearing performance degradation assement [J]. Journal of Vibration and Shock, 2011, 30(8):32-35.
- [9] Qu L S, Li L M, Lee J. Enhanced diagnostic certainty using information entropy theory [J]. Advanced Engineering Informatics, 2003, 17(3/4):141-150.
- [10] 申弢, 黄树红, 韩守木, 等. 旋转机械振动信号的信息熵特征[J]. 机械工程学报, 2001, 37(6):94-98. SHEN Tao, HUANG Shu-hong, HAN Shu-mu, et al. Extracting information entropy features for rotating machinery vibration signals [J]. Chinese Journal of Mechanical Engineering, 2001, 37(6):94-98.