

基于 EEMD 和球 SVM 的高速列车转向架故障估计

翟冰, 金炜东, 秦娜

(西南交通大学 电气工程学院, 成都 610031)

摘要: 基于监测数据评估高速列车空气弹簧、抗蛇行减振器和横向减振器关键部件的运行状态, 针对故障状态下车体构架横向加速度的非平稳信号, 提出 IMF 能量矩与改进的多类超球支持向量机相结合对车体运行状态估计, 改进的超球支持向量机对球外与球内重叠区域的样本用不同的决策, 具有更好的分类效果; 实验数据仿真分析表明, 在速度变化下列车故障识别率稳定在 87% 以上, 证明所用方法能够提取到故障状态下的典型特征, 改进的支持向量分类器并能很好的估计出高速列车的故障状态。

关键词: 转向架; 阈值消噪; EEMD; IMF 能量矩; 超球支持向量机

Running State Estimation of High-speed Train Based on EEMD and Hyper-sphere Support Vector Machines

Zhai Bing, Jin Weidong, Qin Na

(School of Electrical Engineering, Southwest Jiaotong University, Chengdu 610031, China)

Abstract: Based on monitoring data, the running state of key components about air springs, resist sinusoidal vibration absorber and transverse shock absorbers damper of high-speed train were estimated. Aiming at Lateral acceleration for high-speed train bogie in fault, and a typical non-stationary complex vibration signal, a running state estimation method of combining IMF energy moment and improved Hyper-sphere support vector machines methods was proposed. The improved Hyper-sphere support vector machines methods based on improved samples in different parts with different classification policy. Experimental results show that under different speeds, the high speed train recognition rate of damper fault are steady at 87% above. It proves that the improved method can be used to extract the characteristics of the fault condition, and can estimate running state of the high-speed train effectively.

Keywords: bogie; threshold de-noising; ensemble empirical mode decomposition (EEMD); IMF energy moment; hyper sphere support vector machines

0 引言

高速列车的运行安全有其特殊性, 随着服役时间的增加, 高速列车的服役性能也在发生着变化, 其中一些性能的恶化威胁着高速列车的安全运行。转向架是高速列车走行部中的关键部分, 其机械性能的好坏对整个列车系统运行的可靠性有重要的影响, 因此高速列车转向架关键部件的性能检测及故障诊断方法在实际应用中具有重要意义。

转向架故障机理复杂, 故障特征信号微弱, 其振动信号属于非线性的随机信号, 对其提取特征也需要采用非线性随机信号的分析方法^[1]。集合经验模式分解方法 (ensemble empirical mode decomposition, EEMD) 是一种改进经验模式分解 (empirical mode decomposition, EMD) 的信号处理方法, 具有较强的自适应性, 适用于非线性、非平稳信号的分析, 在很多领域得到了广泛应用^[2-5]。

本文提出了将 EEMD 与能量矩特征相结合的故障特征提取方法。该方法首先将采集到的列车构架的振动信号进行小波包阈值滤波^[6-7], 消除噪声干扰。然后对处理后的信号进行

EEMD 分解, 分解出的每一个固有模式函数 (Intrinsic Mode Function, IMF) 所包含的频率成分与采样频率和信号本身有关; 再选取前 6 个 IMF 分量进行能量矩分析, 求出包含主要故障信息的能量矩特征, 最后将处理后的特征通过多类超球支持向量机进行状态估计。具体的信号处理流程如图 1 所示。

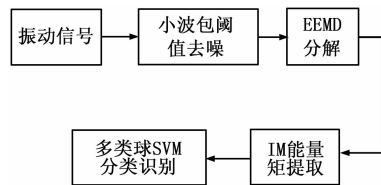


图 1 振动信号处理流程框图

IMF 分量上提取的能量矩特征是一个含有时间的能量分布, 因此 IMF 能量矩是一个信号局部的基于时间的能量统计值。考虑到此方法的实用时, 采集的信号有限, 而多类球 SVM 分类器在分类样本较少时的分类效果比较稳定^[8]。实验表明, 该方法的故障识别率较高。

1 高速列车走行部监测模型分析

在实际运行中, 轨道不平顺的激励是高速列车产生各种振动的主要根源^[9], 列车的垂向加速度随速度的增大而增大, 并通过构架作用于车体, 以较高频率激励出车体的弹性振动, 反过来车体又通过弹簧等反作用于构架, 影响列车的动态性能,

收稿日期: 2013-11-29; 修回日期: 2014-01-26。

基金项目: 国家自然科学基金重点项目 (61134002)。

作者简介: 翟冰 (1989-), 女, 四川南充人, 硕士研究生, 主要从事信号处理和智能信息处理方向的研究。

列车悬挂系统结构简图如图 2 所示。

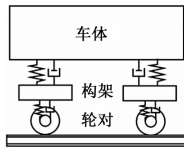


图 2 列车悬挂系统结构简图

空气弹簧、横向减振器和抗蛇行减振器都是列车系统的关键部件, 前两者分别属于为二系和一系悬挂。空气弹簧位于列车转向架中部, 主要功能就是减弱列车运行中发生的振动和抑制运行中发生的噪声, 提高乘坐的舒适度^[9], 但是空气弹簧无法得到左右方向的衰减力。横向减振器一般水平安装在车体和转向架之间, 它能减弱传递至车体的横向振动, 同时保证车体的横向平稳性指标满足乘坐舒适度标准要求, 维持良好的乘坐舒适性^[9]。抗蛇行减振器安置在车体和转向架之间沿纵向方向, 主要功能是抑制高速列车的蛇行运动^[9]。

选取高速列车转向架的正常状态和三种关键部件的典型故障模式(横向减振器故障、抗蛇行减振器故障与空气弹簧失气)进行对比试验。当发生故障时产生异常振动, 由实际可知, 车体振动时频率主要集中在低频段, 垂向振动的主频一般小于 4 Hz。为了提取列车在不同故障下的细微特征, EEMD 分解有很大的优势, 且提出用 IMF 能量矩算法, 训练分类器, 进行故障识别。

2 振动信号的特征提取方法

2.1 信号的 EEMD 分解

希尔伯特-黄变换(Hilbert-Huang transform, HHT)^[10]是近年来提出的一种非平稳信号处理方法, 该方法由 EMD 和 Hilbert 变换两部分组成。EMD 方法本质上是将信号进行平稳化处理, 其结果是把信号中不同尺度的波动或者趋势逐级分解开来, 从而得到一系列具有不同尺度具有真实物理意义的数据分量, 这些分量称为本征模函数(Intrinsic Mode Function, IMF), IMF 可以是调频或调幅信号。但 EMD 方法本身存在诸如模态混叠等缺陷。当高速列车发生故障时, 其振动信号往往会夹杂着较多冲击成分, 利用 EMD 处理故障信号时就更容易造成模态混叠。Huang 和 Wu 提出的 EEMD 有效地抑制了 EMD 分解中的模态混叠现象^[10]。EEMD 的本质是在 EMD 的基础上加了高斯白噪声辅助的分解方法, 利用白噪声频率均匀分布的统计特性, 消除原信号中的间歇现象, 从而有效地抑制模态混叠问题。EEMD 算法可以根据信号自身的特征自适应地分解个中非平稳信号, 并将其分解为中心频率由高到低变化的窄带分量。

EEMD 方法具体分解步骤^[4]为:

1) 在待分析的信号中加入随机高斯白噪声序列,

$$x_m(t) = x(t) + k \cdot n_m(t) \quad (1)$$

式中, k 为加入的白噪声的幅值系数。

2) 将加入白噪声的信号利用 EMD 分解为一组 IMF;

3) 每次加入不同的白噪声序列, 重复步骤 1)、2);

4) 计算分解后 IMF 的均值, 把分解得到的各个 IMF 的均值作为最终的结果。

$$\bar{c}_i = \sum_{m=1}^N c_{i,m} / N \quad (2)$$

式中, N 为 EMD 的集成次数; $c_{i,m}$ 为第 m 次 EMD 所产生的第 i 个 IMF。

如前所述, EEMD 利用白噪声的统计特性抑制模态混叠现象, 使得 EMD 方法成为真正的二进滤波器组, 是对 EMD 方法的重大改进。

2.2 IMF 能量矩特征提取

针对故障振动信号的非平稳特征, 当其发生故障时, 其振动信号在某些频段的能量必将发生变化。而经过 EEMD 分解的每个 IMF 分量包含了某些频段的信息, 为了更好地捕捉到隐藏在信号中的故障信息, 本文采用了一种 IMF 能量矩特征值的算法^[11]。它是传统 IMF 能量法的改进, 该方法用 IMF 能量矩代替 IMF 能量或是能量熵组成特征向量, 它采用了一种基于时间来计算 IMF 能量特征值的方法, 能有效、精准地提取出信号的本质特征, 其特征的提取步骤如下。

第 1 步: 对检测到的振动信号进行预处理: 小波包阈值消噪。

第 2 步: 对消噪后的信号进行 EEMD 分解, 得到若干个 IMF。

第 3 步: 从分解出的 IMF 中筛选几个, 按式 (3) 来计算 IMF 的能量矩, 离散信号, 则按式 (4) 来计算 IMF 的能量矩

$$E_i = \int |c_i(t)|^2 dt \quad (3)$$

$$E_i = \sum_{k=1}^n (k \cdot \Delta t) |c_i(k \cdot \Delta t)|^2 \quad (4)$$

式中, Δt 为采集周期; n 为总的采样点数; k 为采样点。

第 4 步: 构造特征向量并归一化为 T :

$$T = |E_1, E_2, E_3, \dots, E_n| / \sum_i E_i \quad (5)$$

从式 (3) 和式 (4) 中可以看出, IMF 能量矩 E_i 不仅考虑到了 IMF 能量的大小, 还考虑到了 IMF 能量随时间参数 t 的分布情况。因此, 相对于 IMF 能量而言, IMF 能量矩将可以更好地揭示能量分布的特征, 有利于故障特征的提取。

3 多类超球支持向量机

支持向量机(support vector machines, SVM)建立在统计学习理论的 VC 维理论(Vapnik-Chervonenkis dimension theory)和结构风险最小化原理(structural risk minimization)基础上, 根据有限的样本信息在模型的复杂性(即对特定训练样本的学习精度)和学习能力(即无错误地识别任意样本的能力)之间寻求最佳折中, 以获得最好的推广能力^[8]。设有一类模式样本 $Z = \{x_i | x_i \in R_n, i = 1, 2, \dots, m\}$, 其中: m 为样本数目; n 为空间维数。设法找一个最小球体包含所有样本点, 球体以 c 为中心, 以 R 为半径。为了使优化区域更加紧致, 采用核函数 Φ , 将样本点非线性映射到高维特征空间, 在高维特征空间内求解包含所有样本点的最小超球体。为了允许其中的一些数据点存在误差, 引入松弛变量 ξ , 同时将高维空间优化中的内积运算采用满足 Mercer 条件的核函数代替, 即寻找一个核函数 $K(x_i, x_j)$, 使得 $K(x_i, x_j) = \Phi(x_i) \Phi(x_j)$ 。常用的核函数包括: 多项式核函数, 径向基核函数, Sigmoid 核函数。根据结构风险最小化原则, 将问题优化为:

$$\min_{\tau}(R, \xi, c) = R^2 + \frac{1}{\tau m} \sum_{i=1}^m \xi_i \quad (6)$$

$$s. t. \|\Phi(x_i) - c\|^2 \geq R^2 + \xi_i$$

$$(\xi_i \geq 0; i = 1, \dots, m)$$

式中, R 为球半径, ξ_i 为松弛变量, c 为球心, v 为惩罚因子, 用来控制包围球半径与允许球外存在样本的个数之间折中。

使用 Lagrange 乘子法求解带约束的二次规划问题, 得到对偶 Lagrange 表达式, 求解 Lagrange 系数 α_i , 如果 $\alpha_i \neq 0$, 那么对应向量 x_i 为支持向量。根据 Karush—Kuhn—Tucker 最优优化条件, 使用最优理论, 原始问题 (6) 的对偶问题为:

$$\begin{aligned} \max \sum_{i=1}^m \alpha_i K(x_i \cdot x_i) - \sum_{i,j=1}^m \alpha_i \alpha_j K(x_i \cdot x_j) \quad (7) \\ \text{s.t.} \quad \sum_{i=1}^m \alpha_i = 1 \\ 0 \leq \alpha_i \leq \frac{1}{vm}; i = 1, 2, \dots, m \end{aligned}$$

对每一个类别都求解如上所述的二次规划问题, 这样就产生了 k 个最小超球体, 每个超球体代表了一类样本。对超球体的确定起关键作用的超球面上的点, 称之为支持向量, 根据上述结构构造决策函数。对每一个样本点 x 计算它到超球体中心 c_n 的距离的平方即:

$$R(x, c_n) = K(x, x) - 2K(x, c_n) + K(c_n, c_n) \quad (8)$$

$(n = 1, 2, \dots, k)$

并与第 n 类最小超球体半径的平方 R_n 比较, 即:

$$p(x, c_n) = \frac{R_n}{R(x, c_n)} (n = 1, 2, \dots, k) \quad (9)$$

根据式 (10) 来判断样本点 x 的位置:

$$G = \min(p(x, c_1), p(x, c_2), \dots, p(x, c_k)) (n = 1, 2, \dots, k) \quad (10)$$

当 $G > 1$ 时, 即样本点 x 在所有 k 超球体重叠区域, 则比较 $p(x, c_1), p(x, c_2), \dots, p(x, c_k)$ 的大小, 该值最大的球就是样本点 x 所属类别。

当 $G \leq 1$ 时, 即样本点 x 在有的超球体之外或是在所有 k 个球体之外, 这种情况, 需要考虑到各类样本分布疏密程度不同, 则各类超球体大小不同, 本文采用引力法^[12]确定点 x 所属的类别。比较各个超球体对点 x 的引力大小, 其计算公式为:

$$F = q \frac{R_n}{p + R_n} \quad (11)$$

式中, q 为引力系数, 对于不同的故障特征样本, 引力系数不同, 应根据具体的特征样本优化得到。点 x 属于各个超球体中引力最大的球代表的类别。

4 实验数据处理与分析

为验证该方法的有效性, 在动车构架一架一位位置由传感器收集的数据, 采样频率为 243 Hz。数据采集系统已经做了初步的 130 Hz 前置滤波。机车运行工况主要涉及横向减振器全拆、抗蛇行减振器全拆、空气弹簧失气和原车方案 (无故障状态)。每种工况下运行速度按照 80 km/h、120 km/h、140 km/h、160 km/h、200 km/h。每种速度下运行 3.6 min 并记录传感器数据, 采样频率为 243 Hz。

4.1 小波包阈值降噪

列车振动时很多是低频振动, 所以选择小波包阈值滤波器先将 15 Hz 以上的信号滤除掉, 这样可以在频域上剔除高频高峰干扰信号, 有利于特征提取。图 3 为 7 s 的 4 种工况预处理后的频域图。由图知, 动车前架空簧失气和横向减振器全拆故障幅值较大, 两种故障在 1 Hz 左右有敏感的振动频率, 动车原车和空簧无气故障的幅值较小, 动车原车在低频处有最大

幅值。

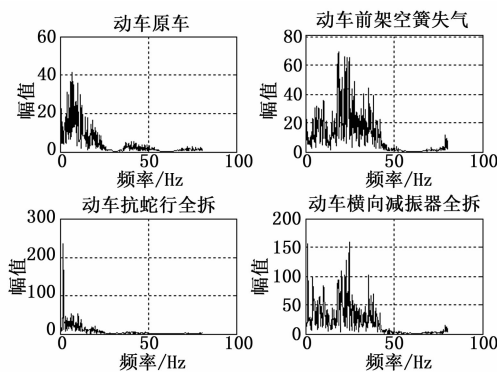


图 3 200 km/h 车体构架一架一位 4 种工况消噪后的频域信号

4.2 信号的 EEMD 分解

由于转向架关键部件故障而引起的减振失效会激起系统各部件不同频段的固有振动, 且 EEMD 方法具有类似正交二进滤波器组的性质, 所以 EEMD 会根据信号本身的特点自适应地将不同频段的固有振动成分分解到不同的 IMF 中, 分解出的 IMF 个数则与信号本身特点有关。高速列横向减振器全拆故障振动信号经过 EEMD 分解共得到 10 个 IMF。

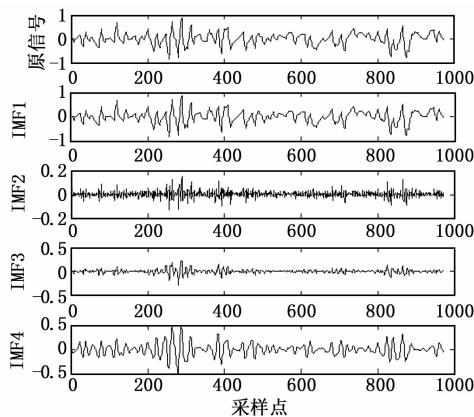


图 4 200 km/h 横向减振器全拆故障振动信号的 EEMD 分解结果的前 4 个 IMF

图 4 为 EEMD 分解结果的前 4 个本征模函数, 其中 IMF1 为 EEMD 方法中引入的高频噪声成分, IMF2、IMF3、IMF4 等为不同频段范围的响应振动, 频率范围从高到低排列。选取 IMF2、IMF3、IMF4、IMF5、IMF6 本征模函数来作为下一步特征提取的对象。

4.3 能量矩特征提取

对前 6 个 IMF 提取 IMF 能量矩, 从图 5 中可以看出, 4 种工况下 IMF2、IMF3、IMF4、IMF5、IMF6 中能量矩较大的都集中在第 3 IMF、4 IMF、5 IMF、6 IMF 上, 说明在这 4 个本征模函数基于时间轴的积分而得到的能量最高。实验发现同一种工况下能量矩的柱状图具有较为相似的形状, 不同工况下差异较为明显, 但是横向减振器全拆和空簧失气这两种工况能量矩值比较接近, 不容易区分。

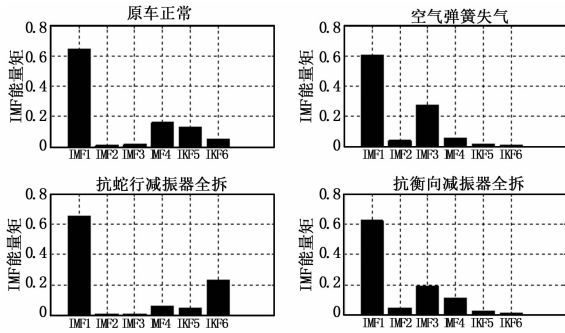


图 5 200 km/h 不同故障工况下的提取的 IMF 能量矩特征

4.4 分类器及分类结果

试验样本为 280 个, 即横向减振器全拆、抗蛇行减振器全拆、空簧失气和原车正常的试验样本各 70 个, 在训练和测试分类器时采用 4 折交叉验证。首先, 对每个试验样本进行 EEMD 分解, 由于 IMF1 含有添加的高频噪声, 所以选取 IMF2、IMF3、IMF4、IMF5、IMF6 频带提取能量矩, 这样每个样本被分别加工为 5 维的特征向量, 利用多类超球支持向量机分类算法构建四分类器, 对横向减振器全拆、抗蛇行减振器全拆、空气弹簧失气和原车正常状态的四分类识别。

本研究核函数采用径向基核函数, 200 km/h 的速度下对横向减振器全拆、抗蛇行减振器全拆、空簧失气和原车正常状态四分类的总识别率与径向基核函数参数的关系 σ (当 $v=0.1$ 时)^[13] 如图 6 所示。优化后核函数参数取 $\sigma=3.8$ 。

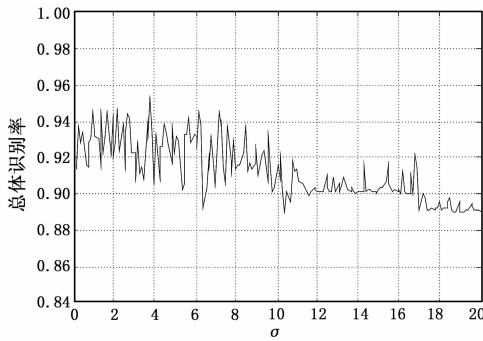


图 6 总体识别率随径向基核函数参数 σ 变化的趋势

都采用 60% 的样本做训练, 40% 的样本做测试, 采用 4 种模式识别方法, 第二种 IMF 能量矩多类球 SVM 采取参考文献 [13] 的决策。4 种故障的正确识别率结果如表 1 所示, 可以看出其中提取 IMF 能量矩特征采用改进的多类超球支持向量机试验结果最好, 随速度的提高, 其识别率比较稳定。

表 1 不同速度下 3 种不同模式识别方法的正确识别率

| | | | | | |
|---------------------|--------|--------|--------|--------|--------|
| 1. IMF 能量矩改进多类球 SVM | 87.23% | 87.78% | 89.80% | 92.45% | 96.23% |
| 2. IMF 能量矩多类球 SVM | 83.75% | 83.30% | 84.45% | 84.11% | 95.27% |
| 3. IMF 能量矩多类 SVM | 83.03% | 84.6% | 83.23% | 75.0% | 90.2% |
| 4. 小波能量矩多类 SVM 分类 | 69.3% | 75.11% | 82.23% | 81.54% | 78.56% |

5 结束语

提出了一种应用 EEMD 的 IMF 能量矩特征和改进的多类超球支持向量机进行故障诊断的方法, 该方法通过对 4 种工况的识别可以预测故障, 在较少的试验样本情况下基于多类超球支持向量机对横向减振器全拆、抗蛇行减振器全拆、空簧失气和原车正常状态振动信号在低速中速的总识别率都稳定在 87% 以上, 优于其他模式识别方法。4 种工况的振动时间序列表现出很强的非平稳性, 从试验可以看出, 基于 EEMD 提取特征向量的方法适合于处理非平稳信号, 有利于机器识别; 在多类超球支持向量机算法中, 对样本是否处于所有球体重叠区域, 采取两种决策, 降低误判率, 由于各类训练模型的参数是独立的, 因此其二次优化过程比利用二分类支持向量机进行多分类简单且计算步骤少。

由于本研究目前所采集到各工况的样本有限, 采用的 EEMD 的时间复杂度大, 还不能有效检测该方法自动识别和预测系统的可靠性。出现虚报和漏报的情况主要是因为横向减振器全拆和空簧失气工况的 IMF 能量矩特征相近, 造成支持向量机识别有误, 需进一步加强对振动信号特征提取方法及其优化的研究, 构造更合适的支持向量机模型, 并采集更多的测试样本来验证该系统的可靠性。

参考文献:

- [1] 张宝珍. 国外综合诊断、预测与健康技术的发展及应用 [J]. 计算机测量与控制, 2008, 16 (5): 591-594.
- [2] 陈 略, 唐歌实, 訾艳阳, 等. 自适应 EEMD 方法在心电信号处理中的应用 [J]. 数据采集与处理, 2011, 26 (3): 361-366.
- [3] 彭 畅, 柏 林, 谢小亮. 基于 EEMD、度量因子和快速峭度图的滚动轴承故障诊断方法波 [J]. 振动与击, 2012, 31 (20): 144-146.
- [4] 郝慧艳, 孙运强, 李晓峰, 等. 基于 EEMD 和 Choi-Willians 分布的侵入加速度信号时频分析 [J]. 中北大学学报 (自然科学版), 2012, 33 (5): 547-567.
- [5] Wu Z H, Huang N E. Ensemble empirical mode decomposition: a noise-assisted data analysis method [J]. Advances in Adaptive Data Analysis, 2009, 1 (1): 1-41.
- [6] 黄 建, 王 新. 小波阈值去噪的改进方法 [J]. 理论与研究测试技术卷, 2007, (9): 4-7.
- [7] 吴 伟, 蔡培升. 基于 MATLAB 的小波去噪仿真 [J]. 信息与电子工程, 2008, 6 (3): 220-229.
- [8] 刘 路, 王太勇, 蒋永翔, 等. 基于超球面支持向量机的刀具磨损状态识别 [J]. 农业机械学报, 2011, 42 (1): 218-222.
- [9] 任尊松, 张曙光. 车辆动力学基础 [M]. 北京: 中国铁道出版社, 2009.
- [10] 李海涛, 王成国, 许跃生, 等. 基于 EEMD 的轨道车辆系统垂向动力学的时频分析 [J]. 中国铁道科学, 2007, 28 (5): 24-30.
- [11] 秦太龙, 杨 勇, 程 珩, 等. 基于 IMF 能量矩和神经网络的轴承故障诊断 [J]. 振动、测试与诊断, 2008, 28 (3): 229-232.
- [12] 袁胜发, 褚福磊. 基于引力球结构支持向量机多类算法的涡轮泵故障诊断 [J]. 宇航学报, 2006, 27 (4): 635-639.
- [13] 吴 石, 刘献礼, 王艳鑫. 基于连续小波和多类球支持向量机的颤振预报 [J]. 振动、测试与诊断, 2012, 32 (1): 46-50.