齿轮点蚀的多通道数据融合识别方法

熊 炘^{1,2}, 杨世锡¹, 甘春标¹, 叶红仙³

(1. 浙江大学液压传动及控制国家重点实验室 杭州,310027)

(2. 上海大学机电工程与自动化学院 上海,200072) (3. 杭州电子科技大学机械工程学院 杭州,310018)

摘要 针对齿轮箱振动信号中混杂其他零部件振动频率的问题,提出一种基于小波包分解独立分量分析(wavelet package independent component analysis,简称 WPICA)和多维经验模式分解(multivariate empirical mode decomposition,简称 MEMD)的齿轮箱齿面点蚀故障信号的多通道数据融合识别方法。首先,利用一种窄带独立分量分析 (sub-band decomposition independent component analysis,简称 SDICA)方法—WPICA,从水泵机组多通道信号中 提取齿轮箱振源,确定齿轮箱振动包含的特征频率成分;其次,借助 MEMD 分解多通道机组振动信号,将所获得的 多维固有模式函数(intrinsic mode function,简称 IMF)进行矩阵互信息运算,完成多通道数据的融合;最后,通过定义 IMF 故障敏感因子,确定故障敏感 IMF 的阶数并获得了齿轮点蚀故障的特征频率。数据分析结果证明了本研究方法的有效性。

关键词 齿轮箱;齿轮点蚀;小波包分解独立分量分析;多维经验模式分解中图分类号 TN911.7;TH165.3

引言

齿轮箱是机械设备中连接和传递动力的通用部件,其结构复杂,工作环境恶劣,极易产生损伤而发生故障。研究其振动机理并开展振动信号的特征分析,对于故障诊断工作意义重大。

Wang 等[1]利用小波变换成功提取了直升机齿轮箱的冲击振动特征。刘婷婷等^[2]将一种基于快速不动点迭代的独立分量分析 (independent component analysis, 简称 ICA)方法应用于旋转机械故障特征提取。毕锦烟等^[3]利用半监督学习的模糊核聚类方法将齿轮轻微点蚀样本与正常运行样本分离开。雷亚国等^[4]在国内外研究的基础上,对行星齿轮箱故障诊断的特点与难点进行了综述,指出了研究中存在的问题,并讨论了解决问题的途径。以上工作是将齿轮箱从整个机械设备中剥离开单独进行研究,但齿轮箱作为传动机构,通常与机械系统的动力输入部件及负载部件同时运转,其振动响应常常混杂了机组其他零部件的振动成分,因而上述研究具有一定的局限性。要准确判别齿轮箱运转是否正

常,必须从多路传感观测中提取包含齿轮箱相关振 动的频率成分。近年来,ICA 和经验模式分解(empirical mode decomposition, 简称 EMD) 信号分析 方法[5-6]的出现、发展及其在多维数据分析等领域的 应用[5-7],为上述问题的解决开辟了新途径。ICA 和 EMD被广泛应用于旋转机械振动信号的特征提 取[2,8-10]。ICA 要求源信号必须满足统计独立性假 设[5],因此对于相互之间不独立的机械振源只能进 行近似估计。Zhang 等[11] 提出了窄带分解 ICA (sub-band decomposition ICA, 简称 SDICA)方法。 该方法在源信号间存在频率交叠的情况下,不要求 源信号间严格独立,放宽了对源信号的假设条件,适 用于机械振源的分离。EMD 作为一种自适应的非 线性、非平稳数据分析方法,特别适用于提取复杂机 械振动信号的频率成分。以往的研究[10] 通常只利 用单一通道数据进行分析,没有借助多通道数据综 合分析机械系统的振动状态,容易造成特征信息的 丢失。笔者根据 Rehman 等[12] 提出的 MEMD 方 法,在分解多通道机组振动数据的基础上,提出了一 种齿轮点蚀故障的多通道数据融合识别方法。

^{*} 国家自然科学基金资助项目(11072214);上海大学理工类创新基金资助项目(K. 10-0109-13-007) 收稿日期:2012-03-09;修回日期:2012-05-17

1 基于 SDICA 的齿轮箱振源识别 方法

由于机组其他部件的振动与齿轮箱振动存在耦合现象,因此计划利用盲信号分离(blind source separation, 简称 BSS)方法进行水泵机组振源识别,进而分离出相对纯净的齿轮箱振动信号。

1.1 SDICA 基本原理

标准 ICA 方法作为 BSS 的一种实现手段,根据独立性测度函数获得源信号估计。

设 $\mathbf{s} = [s_1, s_2, \dots, s_p]^T$ 为相互独立的源信号; $\mathbf{x} = [x_1, x_2, \dots, x_q]^T$ 为观测信号。 \mathbf{x} 中的元素由 \mathbf{s} 中的元素线性组合而成,用矩阵形式可表示为

$$x = Hs + b \tag{1}$$

其中:H 为一列满秩矩阵;b 为加性测量噪声。

ICA 从已知的观测 x 出发,确定线性变换 $W = [w_1, w_1, \cdots, w_q] \cdot H^{-1}$,使得变换后的信号尽快能独立,即

$$\mathbf{v} = \mathbf{W}\mathbf{x} = \mathbf{W}\mathbf{H}\mathbf{s} = \mathbf{O}\mathbf{s} \tag{2}$$

在水泵机组的 BSS 应用中,因为机组各部件 (泵和齿轮箱)的耦合作用,其振动存在频率交叠。如果将不同部件的激励响应视为源信号,它们将不满足独立性条件;因此,用标准 ICA 方法无法准确分离出振源信号。考虑 SDICA 方法,将机组观测信号表示为其子带分量的组合,在不考虑噪声且观测信号个数和源信号个数相等时,式(1)可改写为

$$\begin{bmatrix} \mathbf{x}^1 \\ \mathbf{x}^2 \\ \vdots \\ \mathbf{x}^t \end{bmatrix} = \begin{bmatrix} \mathbf{H} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{H} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{H} \end{bmatrix} \begin{bmatrix} \mathbf{s}^1 \\ \mathbf{s}^2 \\ \vdots \\ \mathbf{s}^t \end{bmatrix}$$
(3)

其中: $\mathbf{x}^i = [x_1^i + x_2^i + \cdots + x_p^i]^T$ 和 $\mathbf{s}^i = [s_1^i + s_2^i + \cdots + s_p^i]^T$ 分别为观测信号和源信号的 i 个子带,i = 1, $2, \dots, I$ 。

机组待分离源信号 s 在通带内是相关的,但存在若干个相互独立的子带。假设在第 i 个子带上是独立的,则可在该子带内用标准的 ICA 方法估计分离矩阵 W。

$$\mathbf{y}^{i} = \mathbf{W}\mathbf{x}^{i} = \mathbf{W}\mathbf{H}\mathbf{s}^{i} \tag{4}$$

Kopriva 等^[13]在 SDICA 方法的基础上,提出了基于小波包分解的相关源分离(WPICA)方法。该方法首先将观测信号用小波包分解成子带信号,假定共轭滤波器 h(n)满足 $\sum h(n-2a)(n-2b)=$

 δ_{ab} , 令 $g(a) = (-1)_a h_{-a+1}$,则可定义递归函数 $\{\mu_a\}$ $(a=0,1,\cdots)$

$$\begin{cases} \mu_{2n}(t) = \sum h_a \mu_n (2t - a) \\ \mu_{2n+1}(t) = \sum g_a \mu_n (2t - a) \end{cases}$$
 (5)

其中: $\mu_0(t)$ 定义为尺度函数 $\Phi(t)$; $\mu_1(t)$ 为小波基函数 $\varphi(t)$; $\mu_n(t)$ (n=2b 或 n=2b+1,b=0,1,…)为正交尺度函数 $\Phi(t)$ 的小波包。

经 n 层小波包分解后,观测信号和相应的源信号可分别表示为

$$\mathbf{x} = \sum_{i=1}^{2^n} \mathbf{x}^i \tag{6}$$

$$\mathbf{s} = \sum_{i=1}^{2^n} \mathbf{s}^i \tag{7}$$

观测信号的子带分量可表示为

$$\sum_{i=1}^{2^n} x^i = \sum_{i=1}^{2^n} s^i \tag{8}$$

接下来计算各个子带内信号的互信息,选择互信息较小的几个子带信号重构成新的观测信号。文献[13]证明了子带观测信号x的互信息最小,意味着子带源信号的互信息最小,即统计独立。利用Cardoso^[14]提出的互信息近似计算方法,选择归一化互信息小于阈值的J个子带重构观测信号 x_{rec} ,则

$$\mathbf{x}_{rec} = \sum_{i=1}^{J} \mathbf{x}^{i} = \mathbf{H} \sum_{i=1}^{J} \mathbf{s}^{i} = \mathbf{H} \mathbf{s}_{rec}$$
 (9)

其中: s_{rec} 为重构的独立源信号。

最后,用 FastICA 算法[15] 从重构的机组观测信号 x_{rec} 中估计分离矩阵 W,用分离矩阵乘以原始观测信号 x 就可实现机组待分离的相关源信号分离。

1.2 齿轮箱振源识别

借助来源于荷兰 Lemmer 的 Buma 泵站的水泵 机组振动数据^[16]进行机组的振源识别。机组由电 机、齿轮箱和泵组成,齿轮箱由电机驱动,带动潜水 泵。机组示意图及传感器安装位置如图 1 所示。其

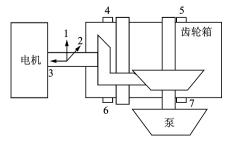


图 1 设备示意图及加速度计安装位置

Fig. 1 Device schematic and accelerometer location

中一台正常工作,另一台机组的齿轮箱发生齿面点蚀故障。齿轮箱输入轴的齿轮齿数为 13,转速为 997 r/min(216 Hz)。通过安装在机组上的 1~7 通道单向加速度计拾取振动数据,信号采样频率为 3.2 kHz。

对整台机组中的泵而言,其工作介质为液态,且受齿轮箱输出轴的激励,因此其结构振动被激发出来^[8]。由此判断机组存在两个振源:泵的结构振动和齿轮啮合相关振动。这两个信号源间存在交叠的频率成分,并在7个通道传感观测处分别产生激励响应。对故障机组的7个通道数据进行傅里叶分析,得到如图2所示的信号幅值谱。图2中:第1,2和4通道在50~150 Hz范围内存在一定能量的25 Hz振动频率成分;在所有7个通道中,都存在较强的216 Hz和286 Hz成分。因此估计系统有两个振源。

利用 WPICA 分离出两个机组振源的时域波形 及幅值谱如图 3 所示。可见,齿轮箱振源(图 3(d))能量主要集中在 216 Hz 和 286 Hz 两个频率成分上。泵的结构振动(图 3(b))能量主要集中在 0~150 Hz 频率范围内,成分以 25 Hz 及其高次谐波为主,其中以 126 Hz 的频率成分能量最大。图 3(b)中仍存在能量相对较弱的 216 Hz 和 286 Hz 频率成

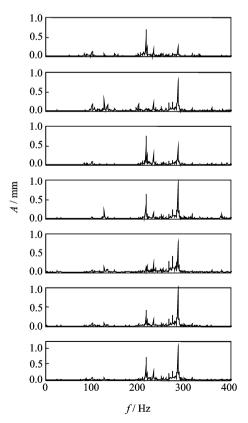


图 2 1~7 通道的齿轮点蚀信号幅值谱

Fig. 2 Amplitude spectrum gear pitting signals in channel $1 \sim 7$

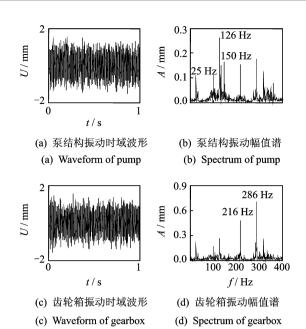


图 3 WPICA 所分离机组振源

Fig. 3 Separation of vibration sources based on WPICA

分。图 3 的分离结果表明,286 Hz 的频率成分和216 Hz 的输入轴齿轮啮合频率成分被共同分离到齿轮箱振源,25 Hz 及其倍频成分属于泵的结构振动源。利用 WPICA 识别出了机组中存在的两个振源,确定了故障特征频率的来源(泵或齿轮箱)。

2 MEMD 在齿轮箱故障特征提取中 的应用

EMD 作为一种自适应数据分析方法(详细算法 参见文献[6]),能将非线性、非平稳信号分解为一组 近似平稳的表征信号所含时间尺度的时间序列——固有 模式 函数 (intrinsic mode function,简称 IMF),进而描述信号的频率构成。设有一维信号 x(t)经 EMD 分解后可表示为

$$x(t) = \sum_{k=1}^{K} c_k(t) + r_K(t)$$
 (10)

其中: $c_k(t)$ 为第 k 阶 IMF; $r_K(t)$ 为信号趋势项。

然而,面对工程实际中诸多复杂的、具有不确定性的非线性动力学问题,以及一些多通道信号处理问题^[12],原始的一维 EMD 算法遇到了阻碍。这主要是因为一维 EMD 算法在处理多维数据时存在尺度排列的不确定性问题^[9,12],即一维算法在分别对多维数据中的单一维进行分解时,所得 IMF 的数量不尽相同,给后续的数据融合带来了困难。因此有必要将原有 EMD 算法从一维拓展到多维,方便多维数据的聚类融合。

2.1 MEMD 基本原理

Rehman 等^[12]提出了 MEMD 方法,该方法首 先将多维信号投影至多个方向向量,然后分别在各 方向上求取信号的投影包络线,最后通过计算包络 线均值的方式定义多维信号的均值。算法的关键问 题是利用准蒙特卡罗低偏差序列(Halton 序列和 Hammersley 序列),生成均匀分布于(n-1)球面的 点集(方向向量集)。因篇幅所限,准蒙特卡罗低偏 差序列生成均匀分布方向向量的算法参见文献 [15]。

定义 q 维观测信号 $\mathbf{x} = [x_1, x_2, \cdots, x_q]^T$ 和其方向向量集 \mathbf{D} 中的第 l 个方向向量 $\mathbf{D}^l = [d_1^l, d_2^l, \cdots, d_q^l]$,则 MEMD 的基本步骤为:

- 1) 确定方向向量集 **D**。
- 2)在 \mathbf{D}^l 上计算 \mathbf{x} 的第 l 个投影 $p^l(t)$, l=1,2, ..., L, 其中 L 为方向向量的个数并确定 $p^l(t)$ 最大值所在的时刻 t_m^l 。
 - 3) 在最大值点 $[t_m^l, \mathbf{x}_m^l]$ 上插值获得多维包络线

 $E^{l}(t)$.

4) 计算包络均值

$$M(t) = \frac{1}{L} \sum_{l=1}^{L} E^{l}(t)$$
 (11)

5) 计算剩余量 R(t) = V(t) - M(t),如果 R(t)满足多维 IMF 的迭代终止条件 [15],则定义 R(t)为 IMF,并对 V(t) - R(t)重复步骤 $2 \sim 5$,直至分离出下一阶 IMF。如果 R(t)不满足 IMF 迭代终止条件,则对其重复执行步骤 $2 \sim 5$,直至满足终止条件。

2.2 齿轮箱故障特征频率的提取

根据 WPICA 振源分析,确定了机组中泵结构振动(25 Hz 及其倍频成分)和齿轮箱啮合振动(216 Hz和 286 Hz)两个振源。接下来需要确定齿轮箱啮合振动中,哪个频率成分为点蚀故障特征频率。因此,将正常条件下通道 1~7 的数据和点蚀条件下通道 1~7 的数据不改变次序的组成 14 路振动数据,利用 MEMD 对其进行分解,每通道分别得到13 个 IMF。图 4(a)为 14 路振动观测中,正常工况下

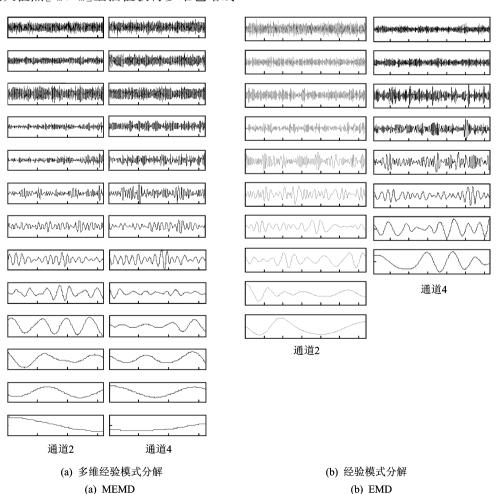


图 4 机组正常运行数据的 IMF 尺度排列对比

Fig. 4 Normal operation data comparison based on different IMF time scales

通道 2 和 4 的 IMF 时域波形。由图可知,MEMD 不存在一维 EMD 的尺度排列问题(图 4(a)~(b)),不同传感观测分解所得的 IMF 数量同为 13。与此形成对比的是图 4(b)中的 EMD 分解结果。第 2 通道共分解出 10 阶 IMF,而第 4 通道只分解出 8 阶 IMF,IMF 阶数互不相同,出现了尺度排列问题(图 4 只进行 IMF 尺度排列对比,不关心信号能量大小,因此不标注信号幅值)。例如,图 4(b)中通道 2 的第 8 阶 IMF 和通道 4 的第 7 阶 IMF 时间尺度近似,但所处的 IMF 阶数却不尽相同。在图 4(a)中,相近的时间尺度都被分解到了第 9 阶 IMF 中。

MEMD 分解后,为了从多个 IMF 中提取对故 障敏感的 IMF,进而确定其包含的频率成分,首先 将每个通道的对应阶 IMF 组成 13 个包含 7 个 IMF 列向量的矩阵,然后通过计算互信息对 IMF 进行聚 类,提取故障敏感频率成分。引入第 k 阶 IMF 的故 障敏感因子 λ_k ,这里定义正常信号为 x_{nor}^r ,其中r= $1,2,\cdots,q$ (本例中 q=7)表示信号来源的通道。 x'_{nor} 的第 k 阶 IMF 为 c'_{nork} 。齿轮点蚀信号为 x'_{gh} ,它的 第 k 阶 IMF 为 c^r_{gpk}。将正常信号和齿轮点蚀信号组 成观测信号矩阵 $\mathbf{x} = [\mathbf{x}_{nor}, \mathbf{x}_{gp}],$ 其中: $\mathbf{x}_{nor} = [\mathbf{x}_{nor}^1, \mathbf{x}_{nor}]$ \dots, x_{nor}^q]; $\mathbf{x}_{\text{gp}} = [x_{\text{gp}}^1, \dots, x_{\text{gp}}^q]$ 。对于 \mathbf{x}_{nor} ,将分解所 得每个通道的对应阶 IMF 组成 IMF 列向量的矩阵 $\mathbf{c}_{\text{nork}} = [c_{\text{nork}}^1, \cdots, c_{\text{nork}}^q]$ 。对于 \mathbf{x}_{gp} ,将分解所得每个通 道的对应阶 IMF 组成 IMF 列向量的矩阵 c_{gpk} = $[c_{\text{gpk}}^1, \cdots, c_{\text{gpk}}^q]$ 。分别计算 x_{nor} 和 $c_{\text{nork}}, x_{\text{gp}}$ 和 $c_{\text{gpk}}, c_{\text{nork}}$ 和 c_{gpk} 之间的互信息[14] u_k , v_k 和 w_k , 从而计算出敏 感因子

$$\lambda_k = \frac{u_k + v_k}{2} - w_k \tag{12}$$

 λ_k 的值越大,表明第 k 阶 IMF 的频率成分对齿轮箱故障越敏感。

选取占据信号总能量 98%的 2~7 阶 IMF 的故障敏感因子如表 1 所示。由表可知,2 阶 IMF 的故障敏感因子最大,表示该 IMF 中包含的频率成分对故障最为敏感,是故障特征频率。图5即为7个通

表 1 IMF 故障敏感因子 λ_k

Tab. 1 Fault sensitive factor of IMF

IMF #	故障敏感因子 λε
2	0.899
3	0.780
4	-1.221
5	-3.406
6	-5.612
7	-6.005

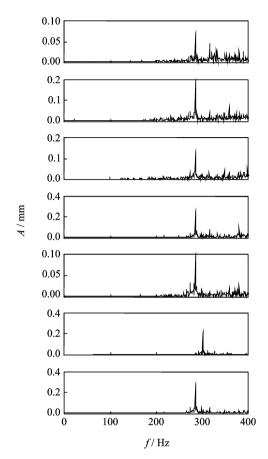


图 5 7 个通道对应 2 阶 IMF 的幅值谱,主要频率为 286 Hz

Fig. 5 Fault sensitive IMF # 2 each of the seven channel, with the principle frequency being 286 Hz

道对应 2 阶 IMF 的幅值谱。图中,2 阶 IMF 主要以 286 Hz 频率成分为主。由于已确定 286 Hz 频率成分和 216 Hz 的输入轴齿轮啮合频率成分同属于齿轮箱振源,在已知齿轮箱发生点蚀故障的前提条件下,286 Hz 成分可确定为齿轮点蚀故障的特征 频率。

3 结束语

笔者结合一种相关源窄带分解方法 WPICA 和MEMD方法,对水泵机组齿轮箱点蚀故障进行了信号特征提取和识别。WPICA 放宽了原始 ICA 算法对源信号统计独立的假设条件,在源信号间存在频率交叠的情况下,通过信号的小波包分解,分离出其中的独立成分,进而使用传统 ICA 算法实现源信号的分离。借助 WPICA 方法,首先将泵的结构振动和齿轮箱中齿轮的啮合振动两个振源从 7 通道传感观测中分离出来。利用 MEMD 的尺度排列优势分

解多维观测矩阵。分别计算分离所得各阶 IMF 矩阵的故障敏感因子,获得对故障敏感的 IMF 阶数。 多通道数据融合后的结果表明,2 阶 IMF 对齿轮点蚀故障最为敏感,所包含的 286 Hz 频率成分幅值变化最大,是齿轮箱故障的特征频率。

值得注意的是,WPICA 方法在分离观测信号时,假定其是由源信号线性混叠而成的。在实际情况下,齿轮箱内部存在固液耦合及冲击振动等非线性因素,用线性混叠模型很难准确地进行系统辨识。因此,对基于相关源的卷积混叠盲分离方法的研究具有重要意义。

参 考 文 献

- [1] Wang W J, McFadden P D. Application of wavelets to gearbox vibration signals for fault detection [J]. Journal of Sound and Vibration, 1996, 192(5): 927-939.
- [2] 刘婷婷,任兴民.独立分量分析在机械振动信号分离中的应用[J].振动、测试与诊断,2009,29(1):36-41. Liu Tingting, Ren Xingmin. Application of independent component analysis to vibration signal separation of rotational machine[J]. Journal of Vibration, Measurement and Diagnosis, 2009, 29(1):36-41. (in Chinese)
- [3] 毕锦烟,李巍华. 基于半监督模糊核聚类的齿轮箱离群检测方法[J]. 机械工程学报,2009,45(10): 48-52. Bi Jinyan, Li Weihua. Semi-supervised kernel-based fuzzy clustering for gear outlier detection[J]. Journal of Mechanical Engineering, 2009, 45(10): 48-52. (in Chinese)
- [4] 雷亚国,何正嘉,林京,等. 行星齿轮箱故障诊断技术的研究进展[J]. 机械工程学报,2011,47(19): 59-67.
 Lei Yaguo, He Zhengjia, Lin Jing, et al. Research ad-

vances of fault diagnosis technique for planetary gear-boxes[J]. Journal of Mechanical Engineering, 2011, 47(19): 59-67. (in Chinese)

- [5] Comon P. Independent component analysis, a new concept[J]. Signal Process, 1994, 36(3): 287-314.
- [6] Huang N E, Shen Z, Long S R, et al. The empirical mode decomposition and the Hilbert spectrum for non-linear and non-stationary time series analysis[J]. Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, 1998, 454(1971): 903-995.
- [7] Lin S L, Tung P C, Huang N E. Data analysis using a combination of independent component analysis and

- empirical mode decomposition[J]. Physical Review E, 2009, 79(6):1-6.
- [8] Ypma A, Leshem A, Duin P W R. Blind separation of rotating machine sources: bilinear forms and convolutive mixtures [J]. Neurocomputing, 2002, 49(1-4): 349-368.
- [9] Zhao X, Patel T H, Zuo M J. Multivariate EMD and full spectrum based condition monitoring for rotating machinery[J]. Mechanical Systems and Signal Processing, 2012, 27: 712-728.
- [10] 曹冲锋,杨世锡,杨将新.基于白噪声统计特性的振动模式提取方法[J].机械工程学报,2010,46(3):65-70.
 - Cao Chongfeng, Yang Shixi, Yang Jiangxin. Vibration mode extraction method based on the characteristics of white noise[J]. Journal of Mechanical Engineering, 2010, 46(3): 65-70. (in Chinese)
- [11] Zhang K, Chan L W. An adaptive method for subband decomposition ICA[J]. Neural Computation, 2006, 18(1): 191-223.
- [12] Rehman N, Mandic D P. Multivariate empirical mode decomposition [J]. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science, 2010, 466;1291-1302.
- [13] Kopriva I, Seršic D. Wavelet packets approach to blind separation of statistically dependent sources[J]. Neurocomputing, 2008, 71(7-9): 1642-1655.
- [14] Cardoso J F. Dependence, correlation and Gaussianity in independent component analysis[J]. Journal of Machine Learning Research, 2003, 4:1177-1203.
- [15] Hyvarinen A. Fast and robust fixed-point algorithms for independent component analysis[J]. IEEE Transactions on Neural Networks, 1999, 10(3): 626-634.
- [16] Ypma A. Vibration measurement dataset [EB/OL]. [1999-12-08]. http://www.ph. th. tudelft. nl/~ypma/mechanical.html.



第一作者简介:熊炘,男,1983 年 11 月 生,博士、讲师。主要研究方向为机电装 备故障诊断、机械动力学和智能检测技 术。曾发表《A new procedure for extracting fault feature of multi-frequency signal from rotating machinery》(《Mechanical Systems and Signal Processing》 2012 年第 32 卷)等论文。

E-mail: xxiong@shu. edu. cn