

# 基于 EMD 和选择性集成学习算法的磨机负荷参数软测量

汤健<sup>1</sup> 柴天佑<sup>2</sup> 丛秋梅<sup>3</sup> 苑明哲<sup>3</sup> 赵立杰<sup>4</sup> 刘卓<sup>2</sup> 余文<sup>5</sup>

**摘要** 针对磨机筒体振动和振声信号组成复杂难以解释、蕴含信息存在冗余性和互补性、与磨机负荷参数映射关系难以描述等问题,提出了基于经验模态分解(Empirical mode decomposition, EMD)技术和选择性集成学习算法分析筒体振动与振声信号组成,建立磨机负荷参数软测量模型的新方法.首先从机理上定性分析了筒体振动及振声信号组成的复杂性;然后采用 EMD 技术将原始信号自适应分解为具有不同时间尺度的系列组成成分,即本征模态函数(Intrinsic mode function, IMF);接着在频域内基于互信息(Mutual information, MI)方法分析并选择 IMF 频谱特征;最后采用基于核偏最小二乘(Kernel partial least square, KPLS)建模方法、分支定界优化算法的选择性集成学习方法建立磨机负荷参数软测量模型,实现了多源多尺度频谱特征的选择性信息融合.基于实验球磨机的实际运行数据仿真验证了该方法的有效性.

**关键词** 经验模态分解,选择性集成建模,磨机负荷参数,选择性信息融合,频谱特征

**引用格式** 汤健,柴天佑,丛秋梅,苑明哲,赵立杰,刘卓,余文.基于 EMD 和选择性集成学习算法的磨机负荷参数软测量.自动化学报,2014,40(9):1853-1866

**DOI** 10.3724/SP.J.1004.2014.01853

## Soft Sensor Approach for Modeling Mill Load Parameters Based on EMD and Selective Ensemble Learning Algorithm

TANG Jian<sup>1</sup> CHAI Tian-You<sup>2</sup> CONG Qiu-Mei<sup>3</sup> YUAN Ming-Zhe<sup>3</sup> ZHAO Li-Jie<sup>4</sup> LIU Zhuo<sup>2</sup> YU Wen<sup>5</sup>

**Abstract** The components of shell vibration and acoustical signals of ball mill are complexity and difficult to interpret. Moreover, the useful information contained in these signals is redundancy and complementary, and the mapping relationships between these signals and mill load parameters are difficult to describe. Aiming at these problems, a new soft sensor approach is proposed, which analyzes shell vibration and acoustical signals for modeling mill load parameters based on empirical mode decomposition (EMD) technology and selective ensemble learning algorithm. At first, the complexity of the shell vibration and acoustical signals are analyzed based on the production mechanism. Then, these signals are adaptive decomposed into a number of intrinsic mode functions (IMFs) with different time-scales using EMD technology, and the spectral features of IMFs are analyzed and selected based on the mutual information (MI) method. At last, the selective ensemble learning algorithm based on kernel partial least square modeling approach and the brand and bound optimal algorithm are used to construct soft sensor models of mill load parameters. Thus, the selective information fusion based on multi-source frequency spectrum features is realized. The simulation results based on operating data from the laboratory ball mill validate the proposed approach.

**Key words** Empirical mode decomposition (EMD), selective ensemble learning, mill load parameters, selective information fusion, frequency spectrum

**Citation** Tang Jian, Chai Tian-You, Cong Qiu-Mei, Yuan Ming-Zhe, Zhao Li-Jie, Liu Zhuo, Wen Yu. Soft sensor approach for modeling mill load parameters based on EMD and selective ensemble learning algorithm. *Acta Automatica Sinica*, 2014, 40(9): 1853-1866

收稿日期 2013-06-14 录用日期 2013-11-26  
Manuscript received June 14, 2013; accepted November 26, 2013

国家自然科学基金(61034008, 61004051, 61203102, 61020106003, 61134006), 111 计划(B08015), 国家支撑计划(2012-BAF19G00), 中国博士后科学基金(2013M532118, 2013M530953, 2013M541820)资助

Supported by National Natural Science Foundation of China (61034008, 61004051, 61203102, 61020106003, 61134006), the 111 Project (B08015), National Key Technology Support Program Project (2012-BAF19G00), National Science Foundation for Post-doctoral Scientists of China (2013M532118, 2013M530953, 2013M541820)

本文责任编辑 刘德荣

Recommended by Associate Editor LIU De-Rong

1. 中国人民解放军 92941 部队 葫芦岛 125001, 中国 2. 东北大学

磨矿是选矿工艺流程的主要能耗、物耗单元,如何控制该过程的优化运行一直是业界研究重点<sup>[1-2]</sup>.

自动化研究中心 沈阳 110004, 中国 3. 中国科学院沈阳自动化研究所信息服务与智能控制技术研究室 沈阳 110016, 中国 4. 沈阳化工大学信息工程学院 沈阳 110142, 中国 5. 墨西哥国立理工大学高级研究中心 墨西哥 07360, 墨西哥

1. Unit 92941, PLA, Huludao 125001, China 2. Research Center of Automation, Northeastern University, Shenyang 110004, China 3. Department of Information Service & Intelligent Control, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, China 4. College of Information Engineering, Shenyang University of Chemical Technology, Shenyang 110142, China 5. Departamento de Control Automatico, CINVESTAV-IPN (National Polytechnic Institute), Av. IPN 2508, México City 07360, México

实时准确检测磨机负荷是实现这一复杂过程优化运行的关键因素之一<sup>[3]</sup>。

磨机旋转运行使得钢球能对原矿进行周期性的冲击和磨剥,实现矿石研磨,同时产生蕴含着丰富磨机负荷信息的振动和振声信号。该领域专家常用这些信号判断能够表征磨机负荷的料球比 (Material to ball volume ratio, MBVR)、磨矿浓度 (Pulp density, PD)、充填率 (Charge volume ratio, CVR) 等磨机负荷参数,调节给矿、给水等过程变量保证选矿生产过程的稳定运行<sup>[4]</sup>。近年来,基于高灵敏度的磨机筒体振动信号开发磨机负荷在线检测仪表成为新的研究热点<sup>[5-7]</sup>,并在基于干式球磨机、半自磨机生产工艺的选矿过程中成功应用,取得了较好的经济效益<sup>[8]</sup>。

由于研磨机理差异,目前基于筒体振动信号进行湿式球磨机负荷的研究多在小型实验磨机上进行。研究表明,不同磨机负荷参数与筒体振动、振声、磨机电流等传感器信号的不同特征相关。文献 [9] 通过提取和选择筒体振动信号的多种频谱特征建立了磨机负荷参数软测量模型,但这种基于单一筒体振动信号的单一软测量模型存在泛化性差、精度低的问题。集成模型可以有效提高工业过程软测量模型的泛化能力<sup>[10]</sup>。文献 [11] 提出了基于筒体振动频谱的磨机负荷参数集成模型,但仍然存在信息融合不充分、泛化性能差等问题。基于上述研究,文献 [12] 提出了选择性融合多传感器信息的磨机负荷参数选择性集成建模,该方法首先确定子模型加权系数计算方式 (自适应加权融合算法),然后采用分支定界算法 (Branch and band, BB) 优化选择最佳特征子集,缺点是虽然可以获得最优子模型,但加权算法却不一定最优,所提方法的普适性有待提高。提高集成模型对特性漂移的适应能力是当前建模难点<sup>[13]</sup>。针对磨矿过程的时变特性,文献 [14-15] 提出了磨机负荷参数在线集成建模方法。这些软测量方法均以传统的基于傅里叶变换获得的单尺度筒体振动和振声频谱为基础,对磨机采取多种不同方式进行了特征提取和特征选择,获得了具有不同优缺点的候选特征子集,但是每个特征子集的具体物理含义却是难以解释,难以获得对磨机负荷检测问题的更深入理解。

磨矿过程是涉及破碎力学、矿浆流变学、机械振动与噪声学、导致金属磨损和腐蚀的“物理-力学”与“物理-化学”等多个学科的复杂过程。磨机内物料和钢球粒径大小及分布的变化、钢球和磨机衬板磨损及腐蚀的不确定性、与钢球冲击破碎直接相关的矿浆粘度的复杂多变等多种因素导致磨机筒体受到大量的不同强度、频率的冲击力,由此产生的筒体振动和振声信号具有较强的非线性、非平稳性和多

尺度特性。如何从产生机理上分析这些信号组成,以及如何将它们有效分解和进行系统解释是目前基于这些信号进行磨机负荷参数软测量面临的挑战之一。

Huang 等提出的经验模态分解 (Empirical mode decomposition, EMD) 是一种基于信号局部特征的自适应分解方法<sup>[16]</sup>,可以有效地将原始时域信号分解为具有多尺度时频特性的本征模态函数 (Intrinsic mode functions, IMFs)。该方法在旋转机械故障诊断领域、高层建筑和桥梁健康状态监测得到了广泛应用<sup>[17-19]</sup>,分解得到的不同 IMFs 具有不同的物理含义,如文献 [19] 给出了高层建筑健康状态监测信号 EMD 分解后的物理含义。受文献 [20] 提出的基于 EMD 和功率谱密度 (Power spectral density, PSD) 的故障诊断方法的启发,本文作者结合 EMD、PSD 和偏最小二乘 (Partial least squares, PLS) 算法分析了筒体振动信号<sup>[21]</sup>,并提出了基于 PLS 潜变量的方差贡献率度量 IMF 信号蕴含信息量的准则;基于这一准则,Zhao 等详细分析了不同研磨工况下 IMF 分量及其频谱的变化,提出了基于 EMD 和 PLS 的选择性集成建模方法<sup>[22]</sup>;因这两种方法仅仅是基于 PLS 模型的第 1 个潜变量方差贡献率选择子模型,其合理性有待探讨,建模精度也较低。本文采用文献 [12] 提出的选择性集成建模方法建立了基于 EMD、核 PLS (Kernel partial least squares, KPLS)、分支定界和误差信息熵加权算法的磨机负荷参数选择性集成模型<sup>[23]</sup>。上述基于 EMD 的建模方法均未对信号组成复杂机理进行分析,未对 IMF 频谱特征进行分析及选择,未对磨机负荷参数与 IMF 间的关系进行深入分析,并且没有考虑选择性融合振声 IMF 频谱,未能结合 EMD 技术、机理分析、领域专家的操作经验对 IMF 频谱与磨机负荷参数间的关系进行深入探讨。

综上,本文在对筒体振动和振声信号组成的复杂性进行定性分析的基础上,提出了基于 EMD 和选择性集成学习算法的磨机负荷参数软测量方法。该方法首先基于 EMD 自适应分解筒体振动及振声信号为不同时间尺度的 IMF,然后基于互信息 (Mutual information, MI) 方法选择多尺度 IMF 频谱特征,接着基于核偏最小二乘 (KPLS) 提取潜在变量建立磨机负荷参数子模型,最后基于分支定界 (BB) 算法为不同的磨机负荷参数选择性集成不同的磨机负荷参数子模型,进而实现选择性融合多源信息特征。不同于之前研究,选择性集成算法中的子模型加权算法不固定为自适应加权融合算法 (Adaptive weighting fusion, AWF),而是扩展为各种加权算法,提高了该算法的普适性。基于实验球磨机的实际运行数据仿真验证了该方法的有效性。

## 1 筒体振动和振声信号组成的定性分析

### 1.1 筒体振动信号组成分析

磨机筒体可近似看作板壳结构体, 通过对磨机筒体的任意径向截面最外层单个钢球运动过程的分析表明, 筒体上任意点  $B_{shell}$  处可检测的振动信号(即筒体径向振动加速度信号)可由下式表示<sup>[24]</sup>:

$$x_V^t = \Theta_a [M_{wet}(B_{shell}), C_{wet}(B_{shell}), K_{wet}(B_{shell}), F_{bmw}] \quad (1)$$

其中,  $x_V^t$  表示筒体振动信号;  $\Theta_a$  表示未知非线性函数;  $M_{wet}(B_{shell})$ 、 $C_{wet}(B_{shell})$  和  $K_{wet}(B_{shell})$  分别表示磨机筒体上点  $B_{shell}$  处的质量、阻尼和刚度特性, 均与磨机负荷相关;  $F_{bmw}$  表示磨机筒体上点  $B_{shell}$  在磨机旋转一周过程中不同时刻受到的冲击力, 即:

$$F_{bmw} = \{F_{i1}, \dots, F_{in}, F_{u1}, \dots, F_{um}\} \quad (2)$$

其中,  $F_{i1}, \dots, F_{in}$  和  $F_{u1}, \dots, F_{um}$  分别表示磨机筒体上点  $B_{shell}$  处在钢球冲击和研磨滑动阶段某时刻受到的冲击力. 这些力在磨机旋转一周的任意时刻都不相同, 可用图 1 表示. 图 1 中,  $T_{mill}$  表示磨机旋转周期, 箭头的长短示意磨机筒体受到的冲击力的大小. 由图 1 可知, 单独分析固定在筒体上的振动加速度传感器在某一时刻采集的筒体振动信号不能包含磨机旋转一周内全部筒体振动信号的特性, 这也是在之前研究中以磨机旋转整周期的信号为单位长度进行分析的原因<sup>[6]</sup>.

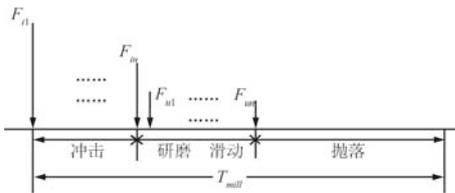


图 1 磨机旋转一周内筒体任意点所受冲击力示意图

Fig. 1 Impact force diagram of any point on mill shell within one mill rotating period

由于磨机内钢球数以十万计、分层排列, 某一时刻对筒体的冲击力实质上是由幅度和频率各不同的力合成的, 由此造成筒体振动信号的组成异常复杂. 此外, 磨机筒体质量不平衡、安装偏心等原因也会造成筒体振动. 这些振动相互耦合、叠加后形成我们通常采集得到的筒体振动信号, 可表示为:

$$x_V^t = \sum_{j_V=1}^{J_V} x_{j_V}^t \quad (3)$$

其中,  $x_{j_V}^t$  表示筒体振动信号的第  $j_V$  个组成成分,  $J_V$  表示筒体振动信号组成成分个数.

### 1.2 振声信号组成分析

磨机振声信号由振动辐射噪声即筒体结构噪声、磨机内部混合声场传输至磨机外部的空气噪声、与磨机负荷无关的环境噪声等三部分组成, 其主要来源是筒体振动, 如图 2 所示:

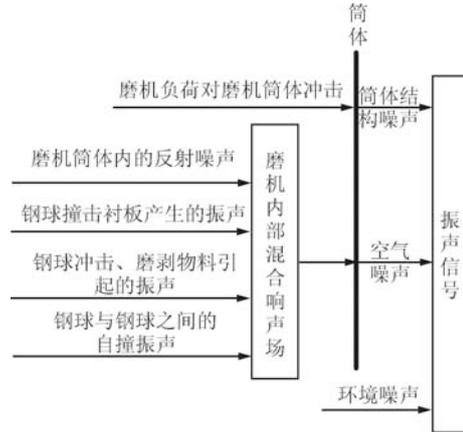


图 2 磨机振声信号的组成示意图

Fig. 2 Mill acoustical signal composition diagram

由图 2 可知, 筒体结构噪声主要是由磨机负荷对磨机筒体冲击造成的筒体振动辐射产生的; 磨机内部混合声场主要是由钢球撞击衬板与磨机筒体产生的噪声、钢球冲击与磨剥物料引起的噪声、钢球与钢球之间的自撞噪声, 以及以上声音在磨机筒内产生的反射噪声等组成. 环境噪声主要是由磨机系统及其它设备产生的噪声、相邻磨机产生的背景噪声等组成.

基于上述定性分析可知, 振声信号是由多种噪声信号叠加而成, 并且不同组成部分蕴含的磨机负荷参数信息明显不同, 其组成较筒体振动信号更为复杂, 可用下式表示:

$$x_A^t = \sum_{j_A=1}^{J_A} x_{j_A}^t \quad (4)$$

其中,  $x_{j_A}^t$  表示振声信号的第  $j_A$  个组成成分,  $J_A$  表示振声信号组成成分个数.

## 2 基于 EMD 和选择性集成学习算法的磨机负荷参数软测量

综合前面分析可知, 对筒体振动和振声信号进行有效分解并提取和选择信号主要组成成分的特征, 是深入理解和建立寓意明确的磨机参数软测量模型的关键. 基于之前的研究成果, 本文提出了由信号分解模块、频谱特征选择模块、子模型模块、选择性集成学习模块 4 部分组成的软测量策略, 其中: 信号分解模块采用 EMD 算法将预处理后的筒体振动和

振声信号自适应分解为若干个具有不同时间尺度的 IMF; 频谱特征选择模块将这些时域 IMF 信号变换为多尺度频谱, 并采用 MI 方法选择频谱特征; 子模型模块建立基于 IMF 频谱特征的 KPLS 磨机负荷参数子模型; 选择性集成学习模块采用 BB 和子模型加权算法优化选择加权 KPLS 子模型并获得最终的磨机负荷参数选择性集成模型, 结构如图 3 所示。

图 3 中,  $x_V^t$ 、 $x_A^t$  和  $x^t$  分别表示时域筒体振动、振声和磨机电流信号;  $x_{VIMF1}^t$  和  $x_{VIMFJ_V}^t$  表示第 1 和第  $J_V$  个筒体振动 IMF 信号;  $x_{AIMF1}^t$  和  $x_{AIMFJ_A}^t$  表示第 1 和第  $J_A$  个振声 IMF 信号;  $z_{ji}$  表示为第  $i$  个磨机负荷参数选择的第  $j$  个特征子集;  $\hat{y}_{ji}$  表示第  $i$  个磨机负荷参数的第  $j$  个子模型的输出;  $\hat{y}_i$  表示第  $i$  个磨机负荷参数选择性集成模型的输出;  $i = 1, 2, 3$  时分别表示 MBVR、PD 和 CVR;  $j = 1, 2, \dots, J_{sel}$  表示频谱特征子集的编号;  $J_{sel}i$  表示针对  $i$  个磨机负荷参数选择的频谱特征子集的数量。

本文对基于 IMF 频谱特征的单一 PLS/KPLS、集成 PLS/KPLS 和选择集成 PLS/KPLS 建模算法的简写如表 1 所示。

### 2.1 基于 EMD 的筒体振动及振声信号分解

筒体振动和振声信号是由周期性作用于磨机筒体的不同振幅和频率的冲击力形成的振动相互叠加产生, 并耦合其它与磨机负荷无关的信号, 构成复杂并且难以解释。研究表明, 这种具有强非线性、非平稳性的信号适合于采用 EMD 技术进行自适应分解和分析。采用 EMD 算法将筒体振动和振声信号自适应分解为不同时间尺度 IMF 信号的步骤如下 (以筒体振动为例):

#### 算法 1. 筒体振动信号自适应分解算法

输入. 磨机旋转若干周的筒体振动信号。

输出. 不同时间尺度的 IMF 信号。

步骤 1. 寻找筒体旋转若干周筒体振动信号  $x_V^t$  极值点;

步骤 2. 连接筒体振动信号最大点和最小点获得上下包络线;

步骤 3. 计算上下包络线均值  $m_{V1}$ , 将原始信号  $x_V^t(t)$  与  $m_{V1}$  的差值作为第一成分, 记为  $h_{V1}$ :

$$h_{V1} = x_V^t(t) - m_{V1} \quad (5)$$

步骤 4. 检查  $h_{V1}$  是否满足 IMF 准则, 即极值点和过零点的个数必须相等或最多相差 1 个; 在任何点上, 局部最大包络和局部最小包络的均值是 0。如果  $h_{V1}$  是 IMF, 则  $h_{V1}$

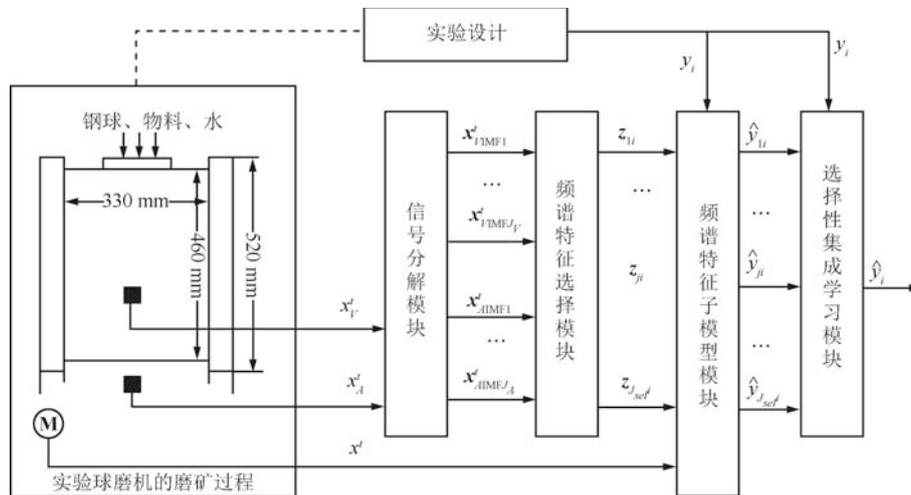


图 3 基于 EMD 和选择性集成学习算法的磨机负荷参数软测量策略

Fig. 3 Mill load parameters soft sensor strategy based on EMD and selective ensemble learning algorithm

表 1 本文所用算法的中英文说明

Table 1 Chinese/English description of the algorithm used in this paper

算法简写	算法英文解释	算法中文解释
IMFPLS	Intrinsic mode functions based partial least squares	基于本征模式函数的偏最小二乘算法
IMFEPLS	Intrinsic mode functions based ensemble partial least squares	基于本征模式函数的集成偏最小二乘算法
IMFSEPLS	Intrinsic mode functions based selective ensemble partial least squares	基于本征模式函数的选择性集成偏最小二乘算法
IMFKPLS	Intrinsic mode functions based kernel partial least squares	基于本征模式函数的核偏最小二乘算法
IMFEKPLS	Intrinsic mode functions based ensemble kernel partial least squares	基于本征模式函数的集成核偏最小二乘算法
IMFSEKPLS	Intrinsic mode functions based selective ensemble kernel partial least squares	基于本征模式函数的选择性集成核偏最小二乘算法

是  $\mathbf{x}_V^t(t)$  的第一个成分;

**步骤 5.** 如果不是 IMF, 重复步骤 1 到步骤 3, 此时,  $\mathbf{h}_{V1}$  作为原始信号:

$$\mathbf{h}_{V11} = \mathbf{h}_{V1} - \mathbf{m}_{V11} \quad (6)$$

其中,  $\mathbf{m}_{V11}$  是  $\mathbf{h}_{V1}$  上下包络的均值. 这个过程重复  $k_V$  次直到  $\mathbf{h}_{V1k_V}$  满足 IMF 准则:

$$\mathbf{h}_{V1k} = \mathbf{h}_{V1(k_V-1)} - \mathbf{m}_{V1k_V} \quad (7)$$

每次都要检查  $\mathbf{h}_{V1k_V}$  的过 0 次数是否与极值点个数相等; 最后得到的成分即第一个 IMF, 并记为:

$$\mathbf{x}_{VIMF1_V}^t = \mathbf{h}_{V1k_V} \quad (8)$$

其中  $\mathbf{x}_{VIMF1_V}^t$  包含筒体振动信号的最小时间尺度;

**步骤 6.** 从原始信号  $\mathbf{x}_V^t(t)$  中剥离  $\mathbf{x}_{VIMF1_V}^t$  得到:

$$\mathbf{r}_{V1}^t = \mathbf{x}_V^t - \mathbf{x}_{VIMF1_V}^t \quad (9)$$

**步骤 7.** 判断是否满足 EMD 分解终止条件: 若不满足, 令  $\mathbf{x}_V^t = \mathbf{r}_{V1}^t$ , 并转至步骤 1; 若满足, 则分解结束.

按上述步骤, 筒体振动和振声信号可以分解为若干个 IMFs 和 1 个残差之和. 分解得到的 IMF 信号按照频率从高到低的顺序进行排列.

EMD 分解得到的各 IMF 信号与原始筒体振动和振声信号的关系可用如下公式表示:

$$\mathbf{x}_V^t = \sum_{j_V=1}^{J_V} \mathbf{x}_{VIMFj_V}^t + \mathbf{r}_{J_V} \quad (10)$$

$$\mathbf{x}_A^t = \sum_{j_A=1}^{J_A} \mathbf{x}_{AIMFj_A}^t + \mathbf{r}_{J_A} \quad (11)$$

其中,  $\mathbf{r}_{J_V}$  和  $\mathbf{r}_{J_A}$  分别表示筒体振动和振声信号分解后的残差.

### 2.2 基于 MI 的 IMF 频谱特征选择

筒体振动和振声信号分解得到的 IMF 时域信号中蕴含着与磨机负荷参数直接相关的信息, 但仍然难以提取有益信息, 并且建模需要关注的是磨机筒体上任意点旋转周期内蕴含的信息; 频谱与磨机负荷参数虽然直接相关<sup>[6]</sup>, 但不同频谱与不同磨机负荷参数的映射关系差异性很大, 需要一种能够对 IMF 频谱特征进行分析和选择的方法. 研究表明, 互信息 (MI) 能有效地描述输入和输出数据间的映射关系. 基于 MI 选择 IMF 特征的算法步骤如下 (以筒体振动为例):

**算法 2. 基于 MI 的筒体振动 IMF 频谱选择算法**

输入. 筒体振动 IMF 信号, MI 阈值

输出. 筒体振动 IMF 频谱特征

**步骤 1.** 采用 Welch's 方法计算 IMF 频谱, 并将振动信号的第  $j$  个 IMF 频谱表示为  $\mathbf{x}_{VIMFj_V}^f$ ;

**步骤 2.** 计算频谱  $\mathbf{x}_{VIMFj_V}^f$  的第  $p_{j_V m_V}$  个变量与第  $i$  个磨机负荷参数间的 MI 值:

$$MI_i((\mathbf{x}_{VIMFj_V}^f)_{p_{j_V m_V}}; \mathbf{y}_i) = \iint \sum \sum p((\mathbf{x}_{VIMFj_V}^f)_{p_{j_V m_V}}, \mathbf{y}_i) \times \log \frac{P((\mathbf{x}_{VIMFj_V}^f)_{p_{j_V m_V}}, \mathbf{y}_i)}{P((\mathbf{x}_{VIMFj_V}^f)_{p_{j_V m_V}})P(\mathbf{y}_i)} d((\mathbf{x}_{VIMFj_V}^f)_{p_{j_V m_V}}) d\mathbf{y}_i \quad (12)$$

其中,  $p((\mathbf{x}_{VIMFj_V}^f)_{p_{j_V m_V}})$  和  $p(\mathbf{y}_i)$  是  $(\mathbf{x}_{VIMFj_V}^f)_{p_{j_V m_V}}$  和  $\mathbf{y}_i$  的边缘概率密度;  $p((\mathbf{x}_{VIMFj_V}^f)_{p_{j_V m_V}}, \mathbf{y}_i)$  是联合概率密度;  $MI_i((\mathbf{x}_{VIMFj_V}^f)_{p_{j_V m_V}}; \mathbf{y}_i)$  采用密度估计方法 (Parzen 窗法) 近似计算<sup>[25]</sup>;

**步骤 3.** 依据经验设定 MI 阈值  $\theta_{MIi}$ ;

**步骤 4.** 若  $MI_i((\mathbf{x}_{VIMFj_V}^f)_{p_{j_V m_V}}; \mathbf{y}_i) \geq \theta_{MIi}$ , 保留该谱变量; 否则丢弃该特征;

**步骤 5.** 重复步骤 2 和 4, 直到选择全部频谱变量, 记为  $\mathbf{x}_{IMFj_{sel\_V}^f}$ , 并简写为  $\mathbf{z}_{Vj_i}$ .

采用上述算法为第  $i$  个磨机负荷参数选择的筒体振动和振声的 IMF 频谱特征的集合  $\mathbf{Z}_i$  采用下式表示:

$$\mathbf{Z}_i = \{\mathbf{z}_{V1i}, \dots, \mathbf{z}_{Vj_i}, \dots, \mathbf{z}_{VJ_{sel\_V}i}, \mathbf{z}_{A1i}, \dots, \mathbf{z}_{Aj_i}, \dots, \mathbf{z}_{AJ_{sel\_A}i}\} = \{\mathbf{x}_{VIMF1_{sel\_V}i}^f, \dots, \mathbf{x}_{VIMFj_{sel\_V}i}^f, \dots, \mathbf{x}_{VIMFJ_{sel\_V}i}^f, \mathbf{x}_{AIMF1_{sel\_A}i}^f, \dots, \mathbf{x}_{AIMFj_{sel\_A}i}^f, \dots, \mathbf{x}_{AIMFJ_{sel\_A}i}^f\} \quad (13)$$

其中,  $\mathbf{x}_{VIMFj_{sel\_V}i}^f$  和  $\mathbf{x}_{AIMFj_{sel\_A}i}^f$  是为第  $i$  个磨机负荷参数选择的筒体振动和振声 IMF 频谱特征变量.

由式 (13) 可知, 为第  $i$  个磨机负荷参数选择得到的频谱特征子集为  $(J_{sel\_V} + J_{sel\_A})i$  个. 本文此处记  $(J_{sel} = J_{sel\_V} + J_{sel\_A})$ , 并将频谱特征子集重新编号, 采用下式表示:

$$\mathbf{Z}_i = \{\mathbf{z}_{1i}, \dots, \mathbf{z}_{j_i}, \dots, \mathbf{z}_{J_{sel}i}\} \quad (14)$$

其中,  $\mathbf{z}_{j_i}$  表示为第  $i$  个磨机负荷参数选择的第  $j$  个频谱特征子集.

### 2.3 基于 KPLS 频谱特征子模型

虽然采用基于 MI 的 IMF 频谱特征选择并没有考虑频谱变量间的相关性, 但是与能够消除输入变量间的相关性并提取与输入输出数据均相关的潜在变量建模的 KPLS 算法结合后克服了这一缺点. 以频谱特征子集  $\mathbf{z}_{j_i}$  为例, 假设训练样本数量是  $k$ ,  $\mathbf{z}_{j_i}$  包含的频谱变量个数为  $p_j$ , 基于 KPLS 频谱特征子模型算法步骤如下所示.

**算法 3. 基于 KPLS 的频谱特征子模型建模算法**

输入. 筒体振动/振声 IMF 的频谱特征.

输出. 筒体振动/振声 IMF 频谱特征子模型.

步骤 1. 将  $\{(\mathbf{z}_{ji})_l\}_{l=1}^k$  映射到高维特征空间, 获得建模样本的核矩阵:

$$K_{ji} = \Phi((\mathbf{z}_{ji})_l)^T \Phi((\mathbf{z}_{ji})_m), l, m = 1, 2, \dots, k \quad (15)$$

步骤 2. 对核矩阵  $K_{ji}$  进行中心化处理:

$$\tilde{K}_{ji} = (I - \frac{1}{k} \mathbf{1}_k \mathbf{1}_k^T) K_{ji} (I - \frac{1}{k} \mathbf{1}_k \mathbf{1}_k^T) \quad (16)$$

步骤 3. 基于非线性迭代偏最小二乘算法 (Non-linear iterative partial least squares, NIPALS), 运用留一交叉验证方法建立频谱特征子模型; 计算交叉验证模型对训练样本  $\{(\mathbf{z}_{ji})_l\}_{l=1}^k$  的输出:

$$\hat{Y}_{ji} = \tilde{K}_{ji} U_{ji} (T_{ji}^T \tilde{K}_{ji} U_{ji})^{-1} T_{ji}^T Y_i \quad (17)$$

其中,  $T_{ji}$  和  $U_{ji}$  是建模样本的得分矩阵;

步骤 4. 计算测试样本  $\{(\mathbf{z}_{t,ji})_l\}_{l=1}^{k_t}$  的核矩阵:

$$K_{t,ji} = \Phi((\mathbf{z}_{t,ji})_l) \Phi((\mathbf{z}_{ji})_m) \quad (18)$$

步骤 5. 对测试样本核矩阵进行中心化处理:

$$\tilde{K}_{t,ji} = (K_{t,ji} I - \frac{1}{k} \mathbf{1}_{k_t} \mathbf{1}_k^T K_{ji}) (I - \frac{1}{k} \mathbf{1}_k \mathbf{1}_k^T) \quad (19)$$

步骤 6. 计算测试样本的输出:

$$\hat{y}_{t,ji} = \tilde{K}_{t,ji} U_{ji} (T_{ji}^T \tilde{K}_{ji} U_{ji})^{-1} T_{ji}^T Y_i \quad (20)$$

其中,  $Y_i$  是建模样本第  $i$  个磨机负荷参数的输出矩阵.

采用上述算法, 针对第  $i$  个磨机负荷参数共建立  $J_{sel}^i$  个频谱特征子模型, 将训练样本的输出记为:

$$\hat{Y}_i = \{\hat{Y}_{1i}, \dots, \hat{Y}_{ji}, \dots, \hat{Y}_{J_{sel}^i i}\} \quad (21)$$

其中,  $\hat{Y}_{ji}$  为  $\mathbf{z}_{ji}$  频谱特征子集的输出.

## 2.4 选择性集成学习模块

筒体振动和振声信号的 IMF 分量具有不同时间尺度和难以描述的物理意义. 基于多尺度 IMF 频谱特征子模型映射了这些 IMF 分量与磨机负荷参数间的函数关系. 基于选择性集成学习算法的磨机负荷参数软测量模型就是如何选择最佳子模型及它们之间的最佳组合方式得到最佳映射关系描述. 在某种意义上, 可以将该过程看作是最优特征选择过程. 基于文献 [12] 的研究成果, 本文提出首先确定加权系数 (不限定加权系数的确定方法), 再采用 BB 优化算法选择子模型及子模型加权系数的选择集成建模方法, 算法步骤如下所示:

**算法 4. 基于频谱特征子模型的选择性集成算法**

输入. 基于频谱特征子集的候选子模型.

输出. 最优选择性集成模型.

步骤 1. 按如下准则排序候选子模型:

$$J_{rmsre\_sub}^i = \theta_{th} - \sqrt{\frac{1}{k} \sum_{l=1}^k \left( \frac{y_{ji}^l - \hat{y}_{ji}^l}{y_{ji}^l} \right)^2} \quad (22)$$

其中,  $k$  为样本个数;  $y_{ji}^l$  为第  $l$  个样本第  $i$  磨机负荷参数的真值;  $\hat{y}_{ji}^l$  为基于  $\mathbf{z}_{ji}$  建立的磨机负荷参数子模型对第  $l$  样本的估计值;  $\theta_{th}$  为依据经验设定的阈值.

步骤 2. 设定第  $i$  个磨机负荷参数集成子模型的数量  $J_{selopt}^i = 2$ ;

步骤 3. 基于子模型排序, 结合分支定界优化算法和子模型加权系数计算方法, 按如下准则选择最优的集成子模型:

$$\text{Max} J_{rmsre\_ens}^i = \theta_{th} - \sqrt{\frac{1}{k} \sum_{l=1}^k \left( \frac{y_{ji}^l - \sum_{J_{sel}^i=1}^{J_{selopt}^i} w_{J_{sel}^i} \hat{y}_{J_{sel}^i}^l}{y_{ji}^l} \right)^2} \quad (23)$$

其中,  $J_{rmsre\_ens}^i$  为选择性集成模型的子模型数量为  $J_{selopt}^i$  时的预测精度;  $w_{J_{sel}^i}$  为子模型加权系数. 若采用自适应加权融合 (AWF)<sup>[12]</sup>、基于误差信息熵的加权<sup>[11]</sup> 等方法计算加权系数, 对  $w_{J_{sel}^i}$  的约束条件为:

$$\sum_{J_{sel}^i=1}^{J_{selopt}^i} w_{J_{sel}^i} = 1, 0 \leq w_{J_{sel}^i} \leq 1 \quad (24)$$

若采用各种线性、非线性回归方法计算加权系数,  $w_{J_{sel}^i}$  用下式表示:

$$w_{J_{sel}^i} = f_w(\hat{y}_{J_{sel}^i}) = f_w(f_{J_{sel}^i}(\mathbf{z}_{J_{sel}^i})) \quad (25)$$

其中,  $f_w(\cdot)$  表示选择的频谱特征子模型输出矩阵  $\hat{Y}_{sel,i}$  与真值  $Y_i$  间的映射关系;  $f_{J_{sel}^i}(\cdot)$  表示频谱特征子集  $\mathbf{z}_{J_{sel}^i}$  与磨机负荷参数  $Y_i$  间的映射关系;

步骤 4. 令  $J_{selopt}^i = J_{selopt}^i + 1$ ;

步骤 5. 若  $J_{selopt}^i = J_{sel}^i - 1$ , 转至步骤 6; 否则转至步骤 3;

步骤 6. 从大到小排序 ( $J_{sel}^i - 2$ ) 个选择性集成模型, 确定  $\text{Max}(J_{rmsre\_ens})$  的选择集成模型为最终磨机负荷参数模型.

采用上述算法, 针对第  $i$  个磨机负荷参数建立的选择性集成模型记为:

$$y_i = \sum_{J_{sel}^i=1}^{J_{selopt}^i} w_{J_{sel}^i} \hat{y}_{J_{sel}^i} = \sum_{J_{sel}^i=1}^{J_{selopt}^i} w_{J_{sel}^i} f_{J_{sel}^i}(\mathbf{z}_{J_{sel}^i}) \quad (26)$$

其中,  $J_{selopt}^i$ 、 $\mathbf{z}_{J_{sel}^i}$ 、 $f_{J_{sel}^i}(\mathbf{z}_{J_{sel}^i})$  为第  $i$  个磨机负荷参数选择集成模型的子模型数量、频谱特征子集和子模型表达式.

## 3 仿真研究与实验研究

实验在 XMQL420×450 球磨机上进行, 该磨机的最大钢球装载量为 80 kg, 设计磨粉能力为 10 kg/h, 转速为 57 r/min. 实验选择的铜矿石直

径小于 6 mm, 钢球尺寸为  $\Phi 30$  mm、 $\Phi 20$  mm 和  $\Phi 15$  mm 三种, 球径配比为 3:4:3. 为保证球做抛落运动, 球负荷由经验丰富的操作人员确定. 为便于研究筒体振动、振声等信号与磨机负荷间的关系, 实验工况的覆盖范围较宽, 详见文献 [6].

### 3.1 EMD 分解结果

筒体振动信号采用固定在筒体表面的加速度传感器直接采集. 考虑到数据量较大, 取磨机旋转 4 周期长度的筒体振动和振声信号采用 EMD 方法自适应分解为 13 和 14 个 IMF 信号, 这些 IMF 的 2 个

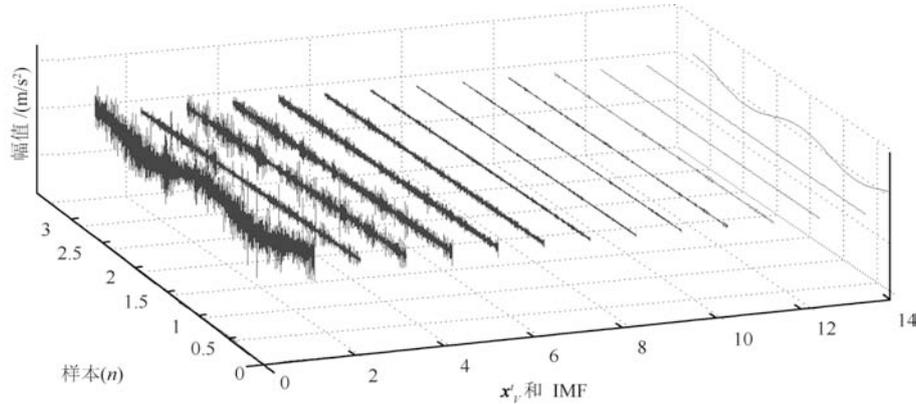


图 4 筒体振动信号 IMF 曲线 (球负荷 40 kg, 物料负荷 10 kg, 水负荷 5 kg)

Fig. 4 IMF curves of vibration signal (Ball load 40 kg, material load 10 kg, water load 5 kg)

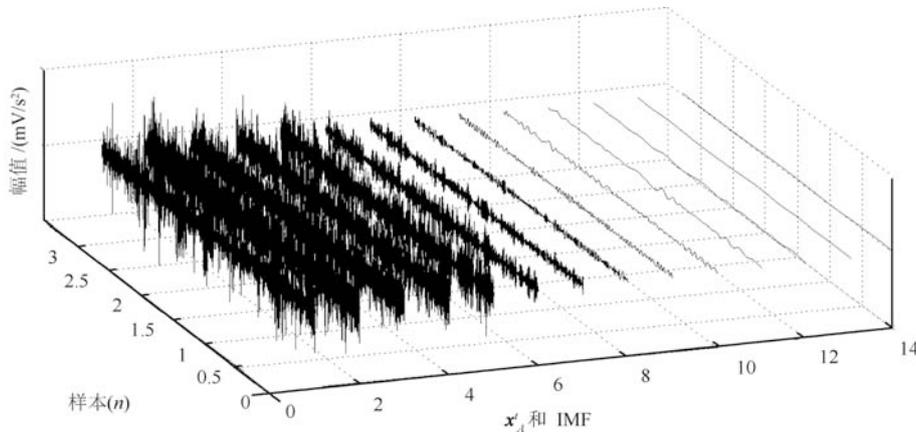


图 5 振声信号 IMF 曲线 (球负荷 40 kg, 物料负荷 10 kg, 水负荷 5 kg)

Fig. 5 IMF curves of acoustical signal (Ball load 40 kg, material load 10 kg, water load 5 kg)

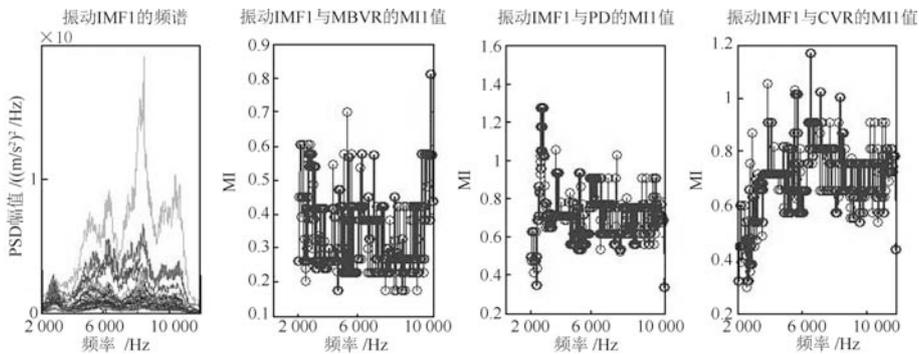


图 6 VIMF1 频谱及其与磨机负荷参数间的 MI 值

Fig. 6 VIMF1 frequency spectrum and its MI values with mill load parameters

周期时域曲线如图 4 和图 5 所示, 图中包含了原始的筒体振动和振声信号.

由图 4 和图 5 可知: 采用 EMD 分解得到的 IMF 时间尺度是递减的, 本文中对筒体振动和振声信号的系列 IMF 分别标记为 VIMF1~VIMF13, AIMF1~AIMF14, 其中图 4 中的 IMF13 明显是一个周期性的正弦信号, 文献 [21] 的研究表明该信号是磨机筒体自身的旋转周期. 可见, EMD 分解后的 IMF 应该是具有明确物理含义的. 由于磨机研磨机理、振动和振声信号产生机理的认识目前还不够透彻, 大部分 IMF 信号物理含义难以合理解释. 但

是通过选择性集成建模, 我们可以确定哪些 IMF 蕴含丰富磨机负荷参数信息.

### 3.2 频谱特征选择结果

对每个 IMF 信号采用同文献 [6] 相同的参数变换到频域, 基于文献 [9] 的方法计算磨机负荷参数与这些 IMF 频谱间的互信息, 部分结果如图 6~10 所示, 其中, 图 6~8 表示筒体振动信号的 VIMF1、VIMF4、VIMF7 频谱及其与不同磨机负荷参数的 MI 值, 图 9 和图 10 表示振声信号的 AIMF2 和 AIMF3 频谱及其与不同磨机负荷参数的

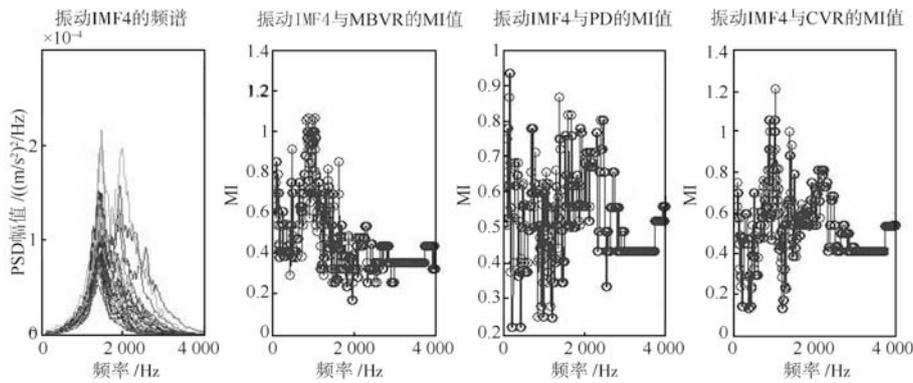


图 7 VIMF4 频谱及其与磨机负荷参数间的 MI 值

Fig. 7 VIMF4 frequency spectrum and its MI values with mill load parameters

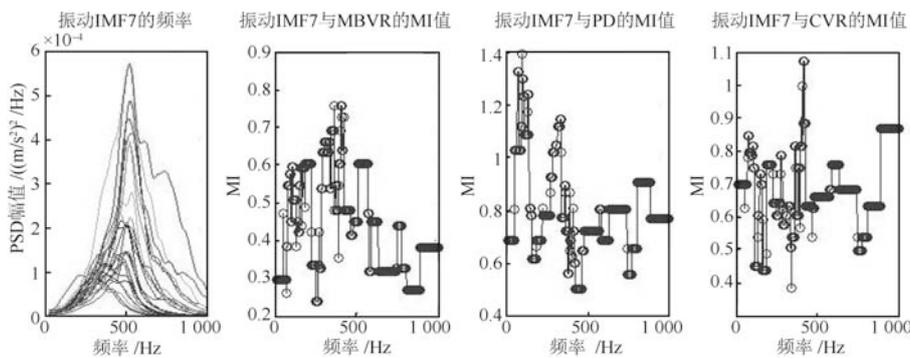


图 8 VIMF7 频谱及其与磨机负荷参数间的 MI 值

Fig. 8 VIMF7 frequency spectrum and its MI values with mill load parameters

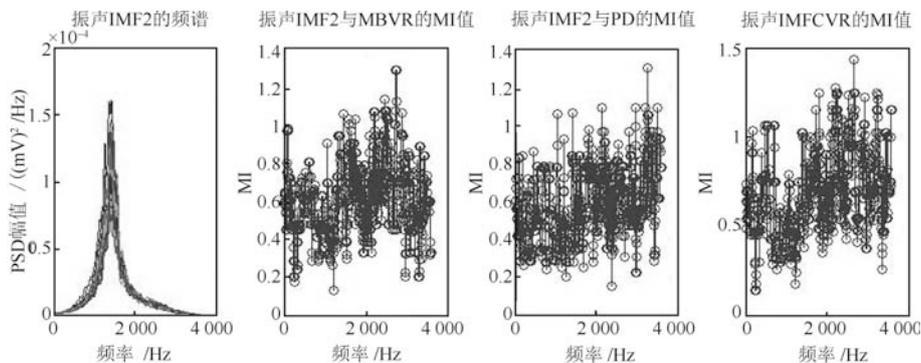


图 9 AIMF2 频谱及其与磨机负荷参数间的 MI 值

Fig. 9 AIMF2 frequency spectrum and its MI values with mill load parameters

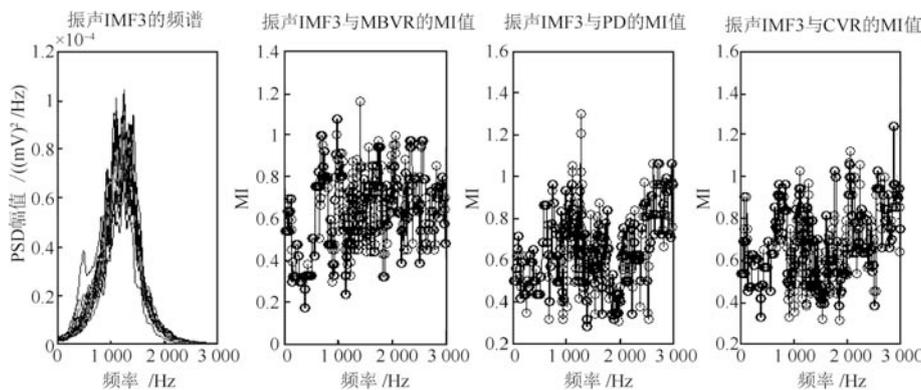


图 10 AIMF3 频谱及其与磨机负荷参数间的 MI 值

Fig. 10 AIMF3 frequency spectrum and its MI values with mill load parameters

MI 值. 为减少参与建模的频谱数量, 文中依据经验采用阈值 0.6, 基于第 2.2 节方法选择了不同频谱特征, 最终为不同磨机负荷参数选择的频谱变量个数如表 2 所示, 其中 VIMF1~12 和 AIMF1~12 分别表示由原始筒体振动和振声信号采用 EMD 算法获得的频率由高到底排列的 IMF 信号的频谱.

上述结果表明: 1) 不同 IMF 频谱的范围不同, 时间尺度不同, EMD 分解是有效的; 2) 不同 IMF 频谱包含的磨机负荷参数信息不同, 频谱选择是必要的; 3) 不同磨机负荷参数与 IMF 频谱的相关性不同, 进行选择性信息融合是合理的. 因此, 有效选择 IMF 频谱特征, 建立磨机负荷参数选择性集成模型是必要的. 本文计算互信息的样本数量是 13 个, 采用更多样本可提高互信息估计的准确度.

### 3.3 子模型比较结果

基于 IMF 频谱特征, 采用第 3.3 节方法建立 KPLS 子模型, 其中子模型均采用统一的径向基函数 (Radical basic function, RBF), 核半径采用网格法搜索, 潜变量个数采用留一交叉验证法确定. 基于筒体振动及振声 IMF 频谱的 KPLS 子模型测试误差如图 11.

从图中可知: 1) 最佳频谱子模型是 AIMF2、VIMF7 和 VIMF2, 与文献 [12] 的研究结论: MBVR 主要和振声频谱、PD 和 CVR 主和振动频谱的结论相符合; 2) 磨机负荷参数子模型的测试误差并不完全与 MI 最大值和选择的变量个数相对应, 表明了映射关系的复杂性.

为对比, 图 12 和图 13 给出了与不进行筒体振动 IMF 频谱特征选择的 PLS 子模型<sup>[22]</sup>、KPLS 子模型<sup>[23]</sup> 测试误差比较. 图 12 和图 13 结果表明, 基于 MI 进行特征选择的策略是有效的. 值得提出的是, 考虑到计算量较大, 本文并未结合预测误差选择 MI 阈值, 也未考虑频谱数据自身特点, 特征选择方法有进一步深入研究的必要.

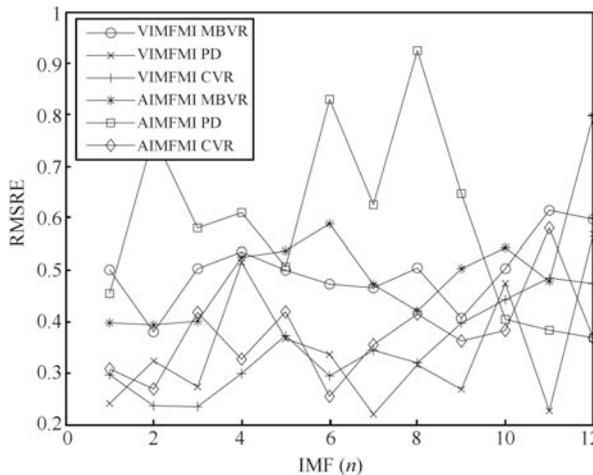


图 11 振动/振声 IMF KPLS 子模型测试误差

Fig. 11 Prediction errors of vibration/acuostical IMFKPLS sub-models

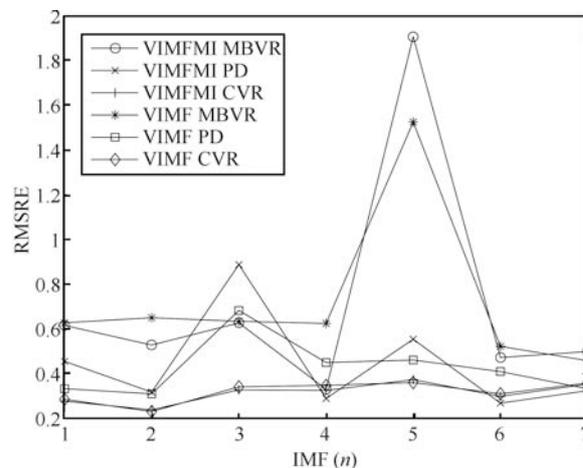


图 12 振动 IMFPLS 与文献 [22] 方法的对比误差

Fig. 12 Comparison of prediction errors of vibration IMFPLS sub-models with [22]

### 3.4 选择性集成结果

采用第 3.4 节所提算法对 IMF 子模型进行选择

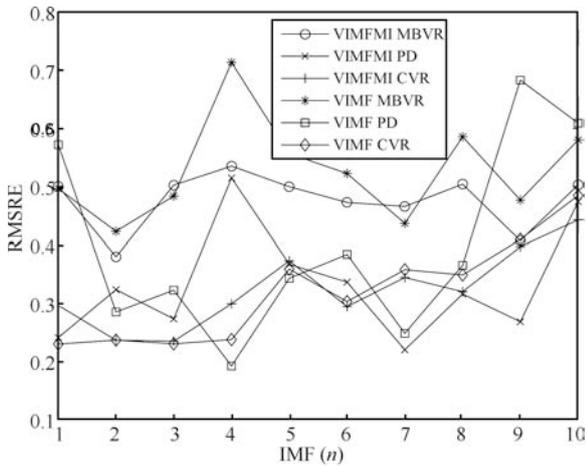


图 13 振动 IMFKPLS 与文献 [23] 方法的对比误差  
Fig. 13 Comparison of prediction errors of vibration IMFPLS sub-models with [23]

性集成, 子模型加权算法分别选择了自适应加权融合算法 (AWF)、基于预测误差信息熵、基于偏最小二乘回归 (Partial least squares regression, PLSR) 的子模型加权算法. 为了与文献 [10] 进行比较, 图 14 给出了基于 AWF 算法、子模型数量为 2~10 时的磨机负荷参数选择性集成 PLS/KPLS 模型测试误差曲线.

本文中, 采用 IMFKPLS 和 IMFPLS 表示最

佳 IMF 子模型; IMFEKPLS 和 IMFEPLS 表示集成全部 IMF 子模型的集成模型; IMFSEKPLS 和 IMFSEPLS 表示采用选择性集成 IMF 子模型的 KPLS 和 PLS 选择性集成模型. 基于 AWF 加权算法的磨机负荷参数选择性集成和集成模型子模型、子模型权系数、预测误差和预测曲线如图 15 所示. 表 3~5 给出了采用不同的子模型加权算法的选择

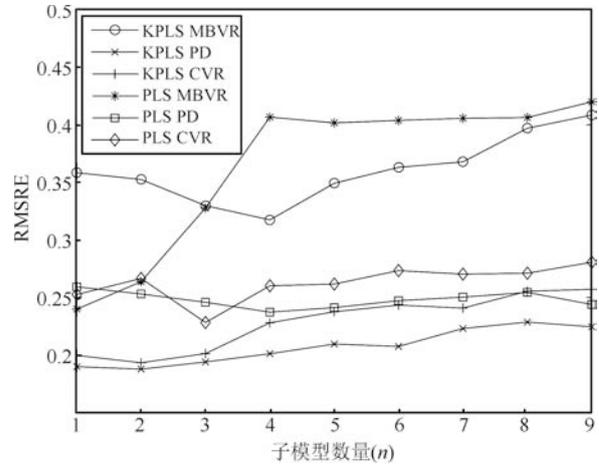


图 14 集成子模型数量与预测误差的关系  
Fig. 14 Relationship between ensemble sub-models' size and prediction errors

表 2 不同 IMF 频谱的范围、MI 值及选择的频谱数量

Table 2 Range, MI value and selected number for different IMF frequency spectrum

频谱	频谱范围	MBVR			PD			CVR		
		MI 最小值	MI 最大值	选择变量个数	MI 最小值	MI 最大值	选择变量个数	MI 最小值	MI 最大值	选择变量个数
VIMF1	2 000:12 000	0.1716	0.8118	228	0.3351	1.2709	8 768	0.2954	1.1699	8 239
VIMF2	1 000:8 500	0.1716	0.9989	1 506	0.3351	1.0870	6 848	0.3254	1.0570	6 063
VIMF3	500:5 500	0.1716	0.8731	1 083	0.2820	1.2485	2 939	0.1716	1.0269	2 750
VIMF4	100:4 000	0.1716	1.0646	705	0.2220	0.9331	992	0.1307	1.2108	960
VIMF5	100:3 000	0.09123	1.102	527	0.2373	1.3289	1 443	0.2373	1.3066	1 645
VIMF6	10:2 000	0.09123	0.9108	144	0.2373	1.0513	1 185	0.2373	1.0570	1 028
VIMF7	10:1 000	0.2373	0.7570	201	0.5016	<b>1.394</b>	9 109	0.3835	1.0723	832
VIMF8	10:1 000	0.09123	1.0646	181	0.2373	1.0212	3 680	0.2264	0.9957	3 240
VIMF9	1:500	0.2674	<b>1.1150</b>	133	0.3555	1.1247	222	0.3254	1.1527	259
VIMF10	1:300	0.1716	0.9485	39	0.2488	0.8751	22	0.2642	1.1024	37
VIMF11	1:200	0.1955	0.7461	35	0.2520	1.2709	30	0.2877	1.2485	54
VIMF12	1:150	0.3178	0.8150	17	0.2597	1.0493	22	0.2373	<b>1.4604</b>	26
AIMF1	1:4 000	0.2297	<b>1.3723</b>	1 795	0.1813	1.1751	1 831	0.2297	1.2485	2 705
AIMF2	1:3 600	0.1307	1.2842	1 840	0.1530	<b>1.2989</b>	1 541	0.1416	<b>1.4381</b>	2 181
AIMF3	1:3 000	0.1716	1.1604	1 703	0.2820	1.2989	1 728	0.3128	1.2389	1 877
AIMF4	1:2 000	0.2751	1.0570	907	0.2220	1.0870	1 130	0.1996	1.0646	3 370
AIMF5	1:1 200	0.2877	1.0646	567	0.2974	1.1093	401	0.2597	1.1828	315
AIMF6	1:800	0.2597	0.8731	565	0.2016	1.0493	583	0.3254	1.1808	236
AIMF7	1:500	0.2877	1.0346	159	0.3478	1.0870	399	0.4103	1.0269	345
AIMF8	1:300	0.1416	0.9689	123	0.2373	1.0212	96	0.09123	0.9957	86
AIMF9	1:200	0.2297	0.8150	87	0.3478	0.8954	48	0.3254	0.7083	53
AIMF10	1:100	0.2954	0.9408	18	0.5016	1.0870	77	0.3254	0.7570	26
AIMF11	1:80	0.2297	0.6989	4	0.3478	0.8297	48	0.3835	0.9408	50
AIMF12	1:60	0.5450	0.6835	8	0.4940	0.6255	52	0.4716	0.6989	54

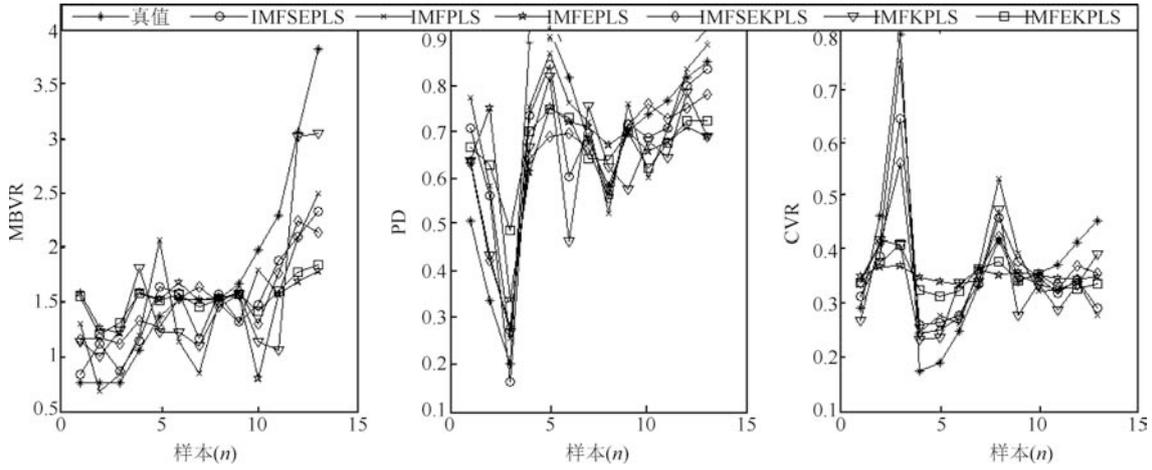


图 15 基于 AWF 加权子模型的磨机负荷参数软测量模型测试曲线

Fig. 15 Testing curves of mill load parameters soft sensor model based on AWF weighting sub-models

表 3 基于 AWF 加权子模型的不同建模方法的预测精度比较

Table 3 Prediction accuracy comparion of different modeling approaches based on AWF weighting sub-models

建模方法	子模型与测试误差						RMSRE (均值)
	MBVR		PD		CVR		
	子模型及权系数	RMSRE	子模型及权系数	RMSRE	子模型及权系数	RMSRE	
IMFEPLS	{1:16}	0.4951	{1:16}	0.4129	{1:16}	0.4127	0.4402
IMFPLS	{VIMF4}	0.3306	{VIMF6}	0.2690	{VIMF2}	0.2390	0.2795
IMFSEPLS	{AIMF3(0.4624); VIMF4(0.5376)}	0.2398	{VIMF7(0.2072); VIMF2(0.1481); VIMF8(0.2425); VIMF4(0.1302); VIMF6(0.2721)}	0.2371	{AIMF2(0.1465); VIMF6(0.4621); VIMF1(0.1970); VIMF2(0.1943)}	0.2281	0.2350
IMFEKPLS	{1:16}	0.4659	{1:16}	0.4757	{1:16}	0.3590	0.4335
IMFKPLS	{VIMF2}	0.3802	{VIMF7}	0.2204	{VIMF2}	0.2352	0.2786
IMFSEKPLS	{VIMF9(0.1293); AIMF3(0.1238); AIMF1(0.4796); AIMF2(0.1359); VIMF2(0.1314)}	0.3173	{VIMF1(0.2158); VIMF11(0.4557); VIMF7(0.3285);}	0.1876	{AVIMF7(0.3748); VIMF2(0.3485); VIMF3(0.2767) }	0.1932	0.2327

表 4 基于不同子模型加权方法的选择性集成 PLS 模型测试误差 (RMSRE)

Table 4 Testing errors (RMSRE) of selective ensemble PLS based on different sub-model weighting methods

	MBVR		PD		CVR	
	IMFEPLS	IMFSEPLS	IMFEPLS	IMFSEPLS	IMFEPLS	IMFSEPLS
AWF	0.4951	<b>0.2398</b>	0.4129	0.2371	0.4127	0.2281
信息熵	0.3545	0.3931	0.4770	<b>0.2370</b>	0.3181	0.2994
PLSR	1.2518	0.2460	1.9227	0.3042	0.3416	<b>0.2148</b>
	0.7004	0.2929	0.9375	0.2594	0.3574	0.2474

性集成模型间的精度比较. 其中, 不同算法的解释说明如表 5 所示.

由图 14~15 及表 3~5 可知:

1) 针对 MBVR 模型: 由表 3 知, IMFSEKPLS 方法选择了 VIMF9、AIMF3、AIMF1、AIMF2、VIMF2 共 5 个子模型, 但其预测误差却大于选择了 AIMF3、VIMF4 共 2 个子模型的 IMFSEPLS 方法, 表明 MBVR 与 IMF 频谱具有线性关系; 两种选择性集成模型中都选择了筒体振动信号的 IMF 子模型, 这与文献 [12] 的选择不一致; 但是从振声产生机理的角度看, 振声信号的主要来源是筒体振动, 通过 EMD 分解得到与 MBVR 更相关的 IMF 是合理的, 也说明本文方法的合理性; 从表 4 和表 5 可知, 不同子模型加权方法的 IMFSEPLS 模型的平均预测误差低于 IMFSEKPLS 方法, 同样表明了 MBVR 与 IMF 频谱间的线性映射关系, 这与文献研究中可通过振声信号直接判断 MBVR 的结论相符.

2) 针对 PD 模型: 由表 3 知, IMFSEKPLS 方法选择了 VIMF1、VIMF11、VIMF7 共 3 个子模型, 预测误差为 0.1876, 小于选择了 VIMF7、VIMF2、VIMF8、VIMF4、VIMF6 共 5 个子模型的 IMFSEPLS 方法, 表明 PD 与 IMF 频谱非线性关系的存在; 表 4 和表 5 的结果也表明 PD 与 IMF 频谱间的非线性映射关系; 两种选择集成模型都只选择了筒体振动信号, 该选择与文献 [12] 和国外关于半自磨机的研究相符合, 也与筒体振动信号的产生机理的定性分析一致.

3) 针对 CVR 模型: 由表 3 知, IMFSEKPLS 方法选择了 AVIMF7、VIMF2、VIMF3 共 3 个子模型, 预测误差为 0.1932, 小于选择了 AIMF2、VIMF6、VIMF1、VIMF2 共 4 个子模型的 IMFSEPLS 方法, 表明 CVR 与筒体振动和振声 IMF 频谱均存在非线性映射关系; 表 4 和表 5 的结果也表明 CVR 与 IMF 频谱间的非线性映射关系; 两种选择集成模型都选择了筒体振动和振声频谱,

这与文献 [12] 只选择了筒体振动频谱特征的结论不相符合; 但本文方法显然更为合理, 因为工业现场操作人员往往根据振声信号沉闷与否判断是否“堵磨”, 显然是靠人耳将振声信号“分解”, 这与本文采用 EMD 方法分解信号相类似; 文献 [10] 未分解振声信号, 难以从振声频谱中提取到有效特征子集; 同时, 由于振声的源是筒体振动, 选择振动 IMF 也是合理的.

4) 从表 4 和表 5 可知: 集成全部子模型的 IMFSEKPLS 和 IMFSEPLS 模型的平均预测误差最大, 说明简单的融合全部 IMF 信息的集成模型并不能获得最佳的建模性能, 进行选择性集成建模是合理的; 不同子模型加权方法针对不同的磨机负荷参数选择性集成模型的误差不同, 表明加权方法还需要进一步的寻优.

## 4 结论

本文从建立筒体振动和振声信号的主要组成成分特征与磨机负荷参数间的非线性映射关系角度出发, 提出了基于经验模态分解和选择性集成学习算法的软测量方法, 建立了基于不同时间尺度的本征模态函数频谱特征的选择性集成模型. 该方法将经验模态分解技术与磨机筒体振动和振声信号的产生机理、工业现场领域专家识别磨机负荷的经验相结合, 为深入理解筒体振动和振声蕴含磨机负荷参数信息的机理提供了更为有效的分析手段, 较之前的研究方法更为合理. 该方法可以推广到具有类似信号特征的工业过程关键变量软测量建模中.

本文对筒体振动及振声信号仅是进行 4 个旋转周期自适应分解, 需要结合机理和工业实际选择更有效预处理方式; 目前对集成子模型的选择准则只考虑了选择性集成模型的均方根相对误差小, 目前对集成子模型的选择准则只考虑了选择性集成模型的均方根相对误差最小, 下一步需要深入研究如何综合考虑子模型间的差异进行选择集成; 互信息阈值与集成模型预测精度间的关系有待于结合更多

表 5 基于不同子模型加权方法的选择性集成 KPLS 模型测试误差 (RMSRE)

Table 5 Testing errors (RMSRE) of selective ensemble KPLS based on different sub-model weighting methods

	MBVR		PD		CVR	
	IMFEKPLS	IMFSEKPLS	IMFEKPLS	IMFSEKPLS	IMFEKPLS	IMFSEKPLS
AWF	0.4659	0.3173	0.4757	<b>0.1876</b>	0.3590	0.1932
信息熵	<b>0.2964</b>	0.4008	0.3838	0.2350	0.2441	0.2645
PLSR	0.3935	0.3717	0.3169	0.2580	0.2980	<b>0.1857</b>
	0.3852	0.3632	0.3921	0.2268	0.3003	0.2144

样本进行深入分析; 考虑如何采用更有效的多尺度分解方法获取蕴含更多有价值信息的可靠频谱和如何结合频谱特点研究更有效的非线性特征选择方法; 子模型的加权算法需要进一步深入研究. 本课题的长远研究方向是建立磨机负荷对磨机筒体冲击过程的数值仿真模型, 结合经验模态分解及其改进算法深入分析和测量磨机负荷参数.

## References

- Wei D H, Craig I K. Grinding mill circuits — a survey of control and economic concerns. *International Journal of Mineral Process*, 2009, **90**(1–4): 56–66
- Chai Tian-You. Operational optimization and feedback control for complex industrial processes. *Acta Automatica Sinica*, 2013, **39**(11): 1744–1757  
(柴天佑. 复杂工业过程运行优化与反馈控制. *自动化学报*, 2013, **39**(11): 1744–1757)
- Tang Jian, Zhao Li-Jie, Yue Heng, Chai Tian-You. Present status and future developments of detection method for mill load. *Control Engineering of China*, 2010, **17**(5): 565–570  
(汤健, 赵立杰, 岳恒, 柴天佑. 磨机负荷检测方法研究综述. *控制工程*, 2010, **17**(5): 565–570)
- Zhou P, Chai T Y, Wang H. Intelligent optimal-setting control for grinding circuits of mineral processing. *IEEE Transactions on Automation Science and Engineering*, 2009, **6**(4): 730–743
- Huang P, Jia M P, Zhong B L. Investigation on measuring the fill level of an industrial ball mill based on the vibration characteristics of the mill shell. *Minerals Engineering*, 2009, **14**(22): 1200–1208
- Tang J, Zhao L J, Zhou J W, Yue H, Chai T Y. Experimental analysis of wet mill load based on vibration signals of laboratory-scale ball mill shell. *Minerals Engineering*, 2010, **23**(9): 720–730
- Das S P, Das D P, Behera S K, Mishra B K. Interpretation of mill vibration signal via wireless sensing. *Minerals Engineering*, 2011, **24**(3–4): 245–251
- Feng Tian-Jing, Wang Huan-Gang, Xu Wen-Li, Xu Ning. An on-line mill load monitoring system based on shell vibration signals. *Mining & Metallurgy*, 2012, **19**(2): 66–69  
(冯天晶, 王焕钢, 徐文立, 徐宁. 基于筒壁振动信号的磨机工况检测系统. *选矿*, 2012, **19**(2): 66–69)
- Tang J, Chai T Y, Yu W, Zhao L J. Feature extraction and selection based on vibration spectrum with application to estimating the load parameters of ball mill in grinding process. *Control Engineering Practice*, 2012, **20**(10): 991–1004
- Gui Wei-Hua, Yang Chun-Hua, Chen Xiao-Fang, Wang Ya-Lin. Modeling and optimization problems and challenges arising in nonferrous metallurgical processes. *Acta Automatica Sinica*, 2013, **39**(3): 197–207  
(桂卫华, 阳春华, 陈晓方, 王雅琳. 有色冶金过程建模与优化的若干问题及挑战. *自动化学报*, 2013, **39**(3): 197–207)
- Tang Jian, Chai Tian-You, Zhao Li-Jie, Yue Heng, Zheng Xiu-Ping. Ensemble modeling for parameters of ballmill load in grinding process based on frequency spectrum of shell vibration. *Control Theory & Applications*, 2012, **29**(2): 183–191  
(汤健, 柴天佑, 赵立杰, 岳恒, 郑秀萍. 基于振动频谱的磨矿过程球磨机负荷参数集成建模方法. *控制理论与应用*, 2012, **29**(2): 183–191)
- Tang J, Chai T Y, Yu W, Zhao L J. Modeling load parameters of ball mill in grinding process based on selective ensemble multisensor information. *IEEE Transactions on Automation Science and Engineering*, 2013, **10**(3): 726–740
- Hao Hong-Wei, Wang Zhi-Bin, Yin Xu-Cheng, Chen Zhi-Qiang. Dynamic selection and circulating combination for multiple classifier systems. *Acta Automatica Sinica*, 2013, **39**(11): 1290–1295  
(郝红卫, 王志彬, 殷绪成, 陈志强. 分类器的动态选择与循环集成方法. *自动化学报*, 2013, **39**(11): 1290–1295)
- Tang Jian, Zhao Li-Jie, Chai Tian-You, Yue Heng. On-line soft-sensing modeling of mill load based on vibration spectrum. *Information and Control*, 2012, **41**(1): 123–128  
(汤健, 赵立杰, 柴天佑, 岳恒. 基于振动频谱的磨机负荷在线软测量建模. *信息与控制*, 2012, **41**(1): 123–128)
- Tang Jian, Chai Tian-You, Yu Wen, Zhao Li-Jie. On-line KPLS algorithm with application to ensemble modeling parameters of mill load. *Acta Automatica Sinica*, 2013, **39**(5): 471–486  
(汤健, 柴天佑, 余文, 赵立杰. 在线 KPLS 建模方法及在磨机负荷参数集成建模中的应用. *自动化学报*, 2013, **39**(5): 471–486)
- Huang N E, Long S R, Shen Z. The mechanism for frequency downshift in nonlinear wave evolution. *Advances in Applied Mechanics*, 1996, **32**: 59–117
- Yang J N, Lei Y, Pan S W, Huang N. System identification of linear structure based on Hilbert-Huang spectral analysis. Part 1: Normal Modes. *Earthquake Engineering & Structure Dynamics*, 2003, **32**(9): 1443–1467
- Yan R Q, Gao R X. Rotary machine health diagnosis based on empirical mode decomposition. *Journal of Vibration and Acoustics*, 2008, **130**(2): 1–12
- Chen J. Application of empirical mode decomposition in structural health monitoring: some experience. *Advances in Adaptive Data Analysis*, 2009, **1**(4): 601–621
- Huang P, Pan Z W, Qi X L, Lei J P. Bearing fault diagnosis based on EMD and PSD. In: *Proceedings of the 8th World Congress on Intelligent Control and Automation*, WCICA 2010. Ji'nan, China: IEEE, 2010. 1300–1304
- Tang J, Zhao L J, Yue H, Yu W, Chai T Y. Vibration analysis based on empirical mode decomposition and partial least square. *Procedia Engineering*, 2011, **16**: 646–652
- Zhao L J, Tang J, Zheng W R. Ensemble modeling of mill load based on empirical mode decomposition and partial least squares. *Journal of Theoretical and Applied Information Technology*, 2012, **45**(1): 179–191
- Tang J, Zhao L J, Long J, Chai T Y, Yu W. Selective ensemble modeling parameters of mill load based on shell vibration signal. *Lecture Notes in Computer Science*, 2012, **7367**: 489–497
- Tang Jian. *Soft Sensing of Ball Mill Load for Grinding Process* [Ph.D. dissertation], Northeastern University, China, 2012  
(汤健. 磨矿过程磨机负荷软测量方法研究 [博士学位论文], 东北大学, 中国, 2012)
- Peng H C, Long F H, Ding C. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, **27**(8): 1226–1238



**汤 健** 北京交通大学计算技术研究所博士后. 1998 年获得海军工程学院工学学士学位, 2006 年和 2012 年获得东北大学控制理论与控制工程专业硕士和博士学位. 主要研究方向为工业过程综合自动化系统, 基于数据驱动的软测量和复杂系统建模与仿真.

E-mail: tjian001@126.com

(**TANG Jian** Postdoctor at Research Institute of Computing Technology, Beijing Jiaotong University. He received bachelor degree from Naval College of engineering in 1998, M. S. and Ph. D. degrees from Northeastern University in 2006 and 2012 in control theory and control engineering, respectively. His research interest covers integrated automation of industrial processes, soft sensor based on data-driven, modeling and simulation of complex system.)



**柴天佑** 东北大学教授. 主要研究方向为自适应控制, 智能解耦控制和流程工业综合自动化理论、方法与技术. 本文通信作者.

E-mail: tychai@mail.neu.edu.cn

(**CHAI Tian-You** Professor at Northeastern University. His research interest covers adaptive control, intelligent decoupling control, integrated automation theory, method and technology of industrial process. Corresponding author of this paper.)



**丛秋梅** 中国科学院沈阳自动化研究所博士后, 2000 年和 2004 年获得沈阳化工大学学士学位和硕士学位, 2011 年获得东北大学博士学位. 主要研究方向为复杂工业过程的建模、控制与能效优化.

E-mail: congqiumei@sia.cn

(**CONG Qiu-Mei** Postdoctor at Shenyang Institute of Automation Chinese Academy of Sciences. She received bachelor degree and master degree at Shenyang University of Chemical Technology in 2000 and 2004, respectively, received Ph. D. degree at Northeastern University in 2011. Her research interest covers modeling, control and process optimization of complex process.)



**苑明哲** 博士, 研究员. 1993 年毕业于哈尔滨船舶工程学院自动控制专业, 获学士学位, 分别于 2000 年和 2006 年在中国科学院沈阳自动化研究所机械电子工程专业获得硕士学位和博士学位, 主要研究方向为分布式控制系统和工业过程先进控制与优化.

E-mail: mzyuan@sia.cn

(**YUAN Ming-Zhe** Ph. D. and professor. He received his bachelor degree in automatics from Harbin Shipment Engineering Institute in 1993, master and Ph. D. degrees in mechatronic engineering from Chinese Academy of Sciences respectively in 2000 and 2006. His research interest covers distributed control system, advanced process control and process optimization.)



**赵立杰** 沈阳化工大学副教授. 1996 年和 1999 年分别在沈阳化工学院获得学士和硕士学位, 2006 年在东北大学获得控制理论与控制工程专业博士学位. 主要研究方向为复杂工业过程的建模、故障诊断与优化等.

E-mail: zlj1unlun@163.com

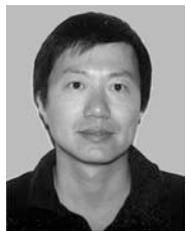
(**ZHAO Li-Jie** Associate professor at Shenyang University of Chemical Technology. She received bachelor degree and MS degree from Shenyang Institute of Chemical Technology in 1996 and 1999 respectively, and Ph. D. degree in control theory and control engineering from Northeastern University in 2006. Her research interest covers modeling, fault diagnosis and optimization of complex industrial processes.)



**刘卓** 东北大学博士研究生. 主要研究方向为复杂工业过程建模.

E-mail: liuzhuo@ise.neu.edu.cn

(**LIU Zhuo** Ph. D. candidate at Northeastern University. Her research interest is complex industrial process modeling.)



**余文** 墨西哥国立理工大学高级研究中心自动化部教授. 1990 年在清华大学获学士学位, 1992 年和 1995 年在东北大学分别获得电子工程专业的硕士和博士学位. 主要研究方向为复杂工业过程建模与控制, 机器学习.

E-mail: yuw@ctrl.cinvestav.mx

(**YU Wen** Professor at the Departamento de Control Automatico of the Centro de Investigaci3n de Estudios Avanzados, National Polytechnic Institute (CINVESTAV-IPN), Mexico City, Mexico. He received the B. S. degree from Tsinghua University, China in 1990 and the M. S. and Ph. D. degrees both in electrical engineering from Northeastern University, Shenyang, China, in 1992 and 1995, respectively. His research interest covers modeling and control of the complex industrial process, machine learning.)