

基于振动频谱的磨矿过程球磨机负荷参数集成建模方法

汤 健¹, 柴天佑^{1,2}, 赵立杰^{1,3}, 岳 恒², 郑秀萍²

(1. 流程工业综合自动化国家重点实验室, 东北大学, 辽宁 沈阳 110004;

2. 东北大学 自动化研究中心, 辽宁 沈阳 110004; 3. 沈阳化工大学 信息工程学院, 辽宁 沈阳 110142)

摘要: 以磨矿过程的湿式球磨机为背景, 针对传统磨机负荷(ML)检测方法只能依靠灵敏度较低的轴承振动、筒体振声和磨机功率等信号监督判断ML状态, 难以检测磨机内部负荷参数的问题, 提出了一种基于高灵敏度的筒体振动频谱的集成建模方法. 首先, 依据磨矿过程的研磨机理, 将振动频谱采用波峰聚类方法自动划分具有不同物理意义的分频段; 然后利用核偏最小二乘(KPLS)算法分别建立各分频段的ML参数子模型; 最后, 依据子模型训练数据预测误差的信息熵获得初始权重, 加权得到最终的ML参数集成预测模型; 在线使用中则根据子模型预测误差的变化进行权值的在线自适应更新. 仿真结果证明了该方法的有效性.

关键词: 磨矿过程; 球磨机负荷; 频谱聚类; 核偏最小二乘; 信息熵; 集成建模

中图分类号: TP29 **文献标识码:** A

Ensemble modeling for parameters of ball-mill load in grinding process based on frequency spectrum of shell vibration

TANG Jian¹, CHAI Tian-you^{1,2}, ZHAO Li-jie^{1,3}, YUE Heng², ZHENG Xiu-ping²

(1. State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang Liaoning 110004, China;

2. Research Center of Automation, Northeast University, Shenyang Liaoning 110004, China;

3. College Of Information Engineering, Shenyang Institute of Chemical Technology, Shenyang Liaoning 110142, China)

Abstract: The mill load (ML) of a wet ball-mill is generally determined by using the low-sensitivity shaft vibration signal, acoustical signal from mill shell and mill power during the grinding process. The result doesn't reflect the status of the load parameters inside the ball-mill. To deal with this problem, we propose the ensemble modeling method based on the high-sensitivity frequency-spectrum signal from the shell vibration. According to the grinding mechanism, the vibration frequency spectrum is automatically partitioned into several different spectral segments with different physical meanings based on the frequency-spectrum clustering method. The sub-model of ML parameters in each spectral segment is built by using the kernel partial least-squares (KPLS) algorithm. The final ML parameters ensemble prediction model is obtained as the weighed combination of sub-models. Each initial weighting coefficient is determined from the information entropy of the training data prediction error of the corresponding sub-model. The weighting coefficient can be adaptively updated on-line based on the variation of prediction error of the sub-model. Simulation results demonstrate the validity of the proposed method.

Key words: grinding process; ball-mill load; frequency-spectrum clustering; kernel partial least squares; information entropy; ensemble model

1 引言(Introduction)

球磨机是保证磨矿过程生产稳定、提高磨矿生产率和产品质量、实现节能降耗的关键设备^[1]. 实际生产过程中, 由于矿石属性及磨矿运行状态的频繁波动, 难以保持磨机的稳定优化运行状态. 磨机负荷(mill load, ML)的准确检测是实现实时调整磨机运行参数, 优化当前运行工况的必要手段. 虽然针对ML的检测已进行了大量研究^[2], 但球磨机的研磨机理仍不清晰^[3], ML的实时检测也未得到完

全解决^[4]. 基于磨机筒体振声信号针对干式球磨机的ML检测仪表已经产品化, 并在氧化铝回转窑制粉系统的智能负荷控制系统中得到了成功应用^[5]. 磨机内部矿浆的存在导致ML难以描述^[6]是湿式球磨机负荷检测的研究与应用远落后于干式球磨机的原因之一. 选矿过程中湿式球磨机的负荷检测主要是结合领域专家知识、规则推理和统计过程控制及融合多源信号估计ML状态^[7-8]. 针对能够准确反映磨机内部负荷并影响磨矿过程产品质量及产

量的磨机内部负荷参数的测量,逐渐成为研究热点^[9-10].文献[11]指出,矿浆浓度、矿石粒度等关键参数与轴承振动及筒体振声频谱的某些特征频段相关,并且筒体振声比轴承振动包含更多的磨机操作参数信息,却受到临近磨机和噪声的交叉干扰.文献[12]则建立基于偏最小二乘(PLS)和主元回归(PCR)的特征频段与磨机操作参数间的回归模型,但其涉及到的内部负荷参数只有矿浆浓度.

近几年,基于灵敏度高、抗干扰性强的磨机筒体振动信号的ML检测方法在半自磨机(SAG)^[13]及干式球磨机上取得了成功^[14-15].研究表明,筒体振动的灵敏度更高,并且能够反映磨机内部的矿浆浓度(PD)和粘度.基于此,澳大利亚CSIRO公司开发了在线筒体振动检测系统,用于监视SAG内部物料的运动,降低钢耗.针对湿式球磨机,文献[16]表明,ML参数与筒体振动直接相关;并建立了基于遗传算法—偏最小二乘(GA-PLS)的ML参数软测量模型,但是该方法基于全部频谱选择特征频段,不但未选频段会导致部分信息缺失,而且所选子频段不具备清晰的物理意义.结合筒体振动的产生机理及空磨、干磨、水磨及湿磨等不同研磨条件下振动频谱的特点,并针对频谱的超高维和共线性问题,文献[17]提出了基于主元分析(PCA)提取分频段特征,串行组合后建立支持向量机(SVM)模型的方法.但该方法存在分频段硬性划分、PCA只能提取线性特征、SVM需要解决二次规划问题等缺点.依据振动、振声等不同的信号及分频段频谱间存在冗余与互补信息的问题,文献[18-19]从不同的角度分别提出了基于PCA/KPCA的融合多源数据特征的ML参数软测量方法.但是,PCA/KPCA用于建模时,提取的主元虽能够解释原始输入数据,但并没有考虑对输出数据的影响^[20];而且采用方差变化率很小的主元用于建模,会导致模型预测性能的不稳定^[21].

从振动系统分析的角度讲,基于振动频谱的ML检测的实质是针对由磨机筒体与ML组成的机械振动系统的物理参数及其冲击载荷的识别问题.当振动频谱的频率分辨率较高时,分频段的载荷识别方法可以提高识别模型的稳定性和可靠性^[22].偏最小二乘方法(PLS)能够同时提取输入与输出数据中的变化信息,采用潜变量建立多元线性回归模型,能够对受噪声污染,存在共线性、高维、病态等特征的数据进行有效回归分析^[23].因此,可用PLS算法建立不同分频段且具有较清晰物理含义的子模型.但是PLS难以提取频谱中的非线性信息.研究表明,集成多个子模型的方法,可以提高模型的泛化性、有效性及可信度^[24-26].针对回归问题,常用的子模型集成方法有简单平均集成、基于多元线性回归的集成、贝叶斯模型平均集成及加权或非加权的集成等

方法^[27].文献[28]采用基于最小均方差的自适应加权融合方法进行多传感器信息的融合.文献[29]利用信息熵^[30]的概念,提出了基于熵值的加权系数确定方法.工业过程多具有时变特性,高灵敏度的筒体振动频谱在不同的研磨条件下变化显著.因此,有必要根据分频段子模型的预测误差在线更新加权系数.

本文针对上述问题,结合实验磨机的磨矿过程,在对研磨过程机理分析的基础上,建立了ML参数的集成模型.该方法结合先验知识,首先用频谱聚类的方法自动划分振动频谱为若干个分频段,然后建立基于核偏最小二乘(KPLS)算法的各分频段ML参数子模型,最后依据子模型训练预测误差的信息熵加权得到最终的集成预测模型.在线使用时则根据子模型预测误差的变化进行权值的在线自适应更新.本文最后与不同的建模方法进行了比较.

2 球磨机研磨过程的机理分析(Mechanism analysis of the ball mill grinding process)

磨矿过程是一个复杂的物理、化学及物理化学过程,对ML的影响因素很多,并且这些因素间彼此相互作用,互相制约^[31].钢球在磨机中的运动极为复杂,其轨迹存在圆周、抛落和泻落3种形态.筒体振动由分层排列、数量众多、不同运动形态的钢球与矿石、钢球与衬板、钢球与钢球及矿石与衬板间的多种冲击力的叠加所致.但是,冲击力以做抛物运动的钢球为主,且在任一时刻,最外层的钢球具有最大的功能.因此,取磨机筒体的任一截面上最外层的某一个钢球为研究对象,其研磨过程可分为抛落、冲击、研磨、滑动4个过程,如图1所示.

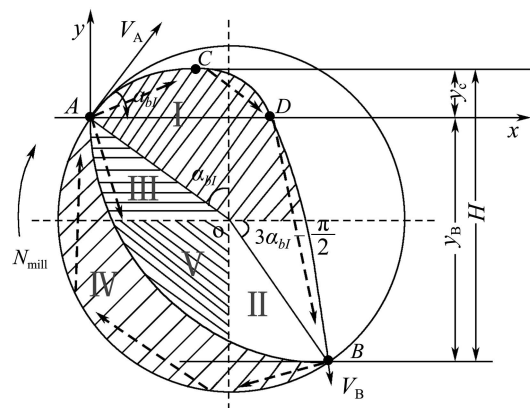


图1 磨机研磨过程示意图

Fig. 1 Motion analysis of the mill grinding process

磨机内各层钢球被筒体衬板带动的运动轨迹可分为上升段和下落段.在上升段,钢球从落回点B到脱离点A绕圆形轨迹运动;在下落段,钢球到达顶点后作抛物线运动撞击衬板后并反弹.按是否产生对物料的破碎作用,可分五区: I区为抛落区,钢球被提升到A点后以初速度 V_A 向下抛落; II区为冲击破碎

区, 磨球从A点落到B点产生最大的冲击作用; III区为滑动区; IV区为研磨区, 基本无冲击作用; V区为死角区, 既无冲击又无摩擦。

在实际的湿式磨矿过程中, 矿浆浮升作用和粘滞作用阻碍了钢球的运动, 从而影响了在B点的径向分速度和切向分速度, 最终影响了钢球的冲击和磨剥力^[31]。其中, 料球比(MBVR)、矿浆浓度(PD)和粘度影响对钢球的浮升与粘滞作用及钢球表面的罩盖层厚度; 磨机内充填率(CVR)的高低则影响钢球在矿浆内的运动时间。因此, 单个钢球下落时对磨机筒体的冲击和磨剥能量与磨机的转速、料球比、矿浆浓度、充填率及矿浆粘度均相关。基于单个钢球的离散元分析表明, 钢球和衬板间的摩擦系数决定钢球负荷脱离衬板的“肩部”区域和下落的“底角”区域; 钢球与衬板间的反弹系数, 决定了冲击后的反弹速度^[32]。这两个系数均受ML参数影响, 进而影响在冲击点的冲击能量。

以上分析只是针对最外层的钢球的一个定性分析。实际上, 磨机内部钢球数量众多, 分层抛落, 不同层的冲击力是不同的; 钢球配比不同, 不同直径大小的钢球的冲击力也不同; 不同性质的矿石, 其硬度和粒度分布的不同, 对冲击的影响不同; 不同的矿浆粘度下, 对钢球的浮升和粘滞作用不同。实际生产中, 还有泻落状态的钢球, 钢球之间相互碰撞产生的冲击力更加难以描述。因此, 不同来源、不同频率的冲击力相互叠加, 实现矿石的粉碎。同时, 这些冲击力周期性的作用于磨机筒体, 产生了组成复杂的筒体振动信号。

众所周知, 任何结构体在频域内的振动波形都是该结构体固有模态或外部冲击力引起的模态的体现^[33], 并且振动频谱中每个大的波峰均表征振动的一个模态^[34]。因此, 筒体振动的加速度频域信号应该包含至少3个模态即磨机与负荷组成的新机械结构体的固有模态、直接冲击磨机衬板引起的冲击模态和其他高频冲击力引起的次冲击模态等。显然, 不同的振动模态中包含不同的ML信息。

3 球磨机负荷参数集成模型(Ensemble model for parameters of ball mill load)

3.1 建模策略(Modeling strategy)

结合小型球磨机的磨矿过程, 提出了由数据预处理模块、频谱自动分割模块、ML参数子模型模块及在线加权集成算法模块共4部分组成的ML参数软测量策略, 如图2所示。

数据预处理模块主要是针对原始信号进行重采样、处理离群点以及滤除振动信号中的低频和高频干扰, 同时将特征难以提取的时域信号转换为频域信号, 详见文献[34]; 频谱自动分割模块结合先验知

识采用频谱聚类方法实现频谱自动分割并选择频段; ML参数子模型模块则建立基于KPLS算法的分频段ML参数子模型; 在线加权集成算法模块则加权子模型的预测输出获得最终ML参数估计值, 并在线实时更新权值。图2中, 上标t及f分别表示时域及频域信号; x_0^t 表示未经信号预处理的时域信号; x^t 表示预处理后的时域信号; x^f 表示振动频谱; x_j^f 表示为分频段频谱, 其中 $j = 1, 2, 3$ 时分别表示频谱的低(LF)、中(MF)和高(HF)频段; \hat{y}_{ji} 表示ML参数子模型的预测输出, 其中 $i = 1, 2, 3$ 分别表示MVBR, PD, CVR; \hat{w}_{ji} 表示子模型的加权系数; \hat{y}_i 表示集成预测模型的输出。

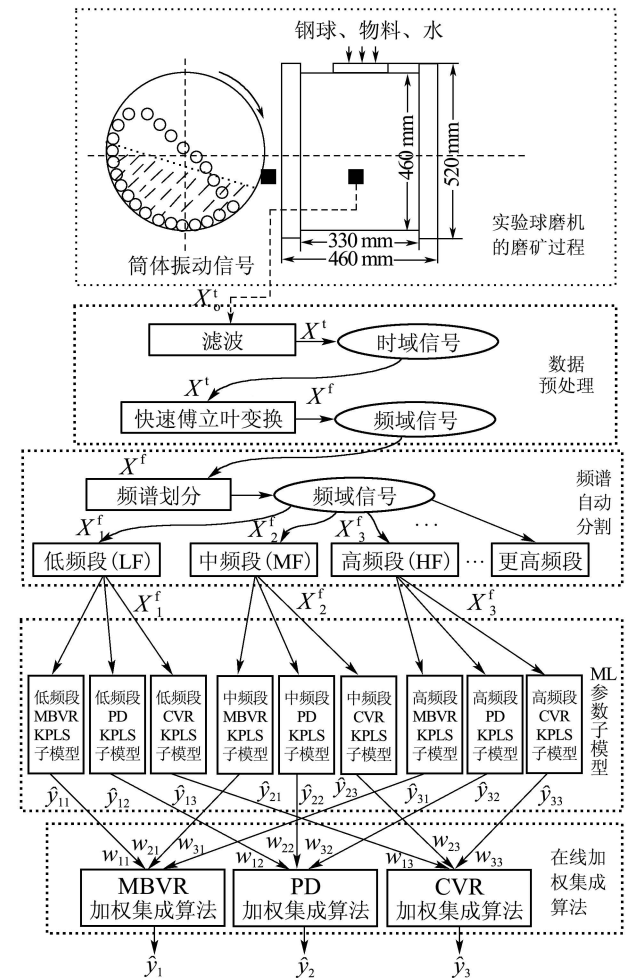


图 2 磨机负荷参数软测量策略

Fig. 2 Soft sensor strategy of the mill load parameters

3.2 基于频谱聚类的分频段自动识别算法(Spectral segments auto partition algorithm based on the frequency-spectrum clustering)

结合先验知识, 通过改进文献[34]提出的波峰聚类的方法实现振动频谱分频段的自动分割, 克服人工硬性划分的随意性。振动频谱是很多局部波峰的集合, 首先定义频谱局部波峰的质量 B_m 和质心 B_c :

$$B_m = \sum_{n=n1}^{n2} x_n^f, \quad (1)$$

$$B_c = \sum_{n=n1}^{n2} n(x_n^f)^2 / \sum_{n=n1}^{n2} (x_n^f)^2, \quad (2)$$

其中: $n1 < n2$, $n1$ 和 $n2$ 是局部波峰的频率范围; 将 n_c 记为局部波峰的中心频率. 将局部波峰表示为 $\langle n1, n2, n_c, B_m, B_c \rangle$, 则可根据某种准则将某些局部波峰聚为一个波峰类. 一个波峰类就是振动频谱的一个模态即要划分的分频段. 分频段可表示为 $\langle C_{n1}, C_{n2}, C_{B_m}, C_{B_c} \rangle$, 局部波峰和分频段间的引力定义如下^[34]:

$$D(B, C) = B_m C_{B_m} / (B_c - C_{B_c})^2, \quad (3)$$

本文提出的频谱自动分割算法步骤如下:

Step 1 给定分频段的数量 D_v 和局部波峰的集合 $B = \{B_1, \dots, B_z, \dots, B_{NB}\}$, 其中 NB 是局部波峰的数量, $B_z = \langle n_{z1}, n_{z2}, n_{zc}, B_{zm}, B_{zc} \rangle$ 是第 z 个局部波峰.

Step 2 根据知识给定分频段的范围, 在每个分频段的范围内搜索最大质量的局部波峰, 并将这些局部波峰作为初始的分频段. 所有的波峰类按频率的升序排列, 可以表示为 $C = C = \{C_1, \dots, C_r, \dots, C_{D_v}\}$, 其中, $C_r = \langle C_{rn1}, C_{rn2}, C_{rB_m}, C_{rB_c} \rangle$ 是第 r 个分频段.

Step 3 计算第 z 个局部波峰的质心与每个初始分频段的质心间的引力($B_{zc} - C_{rB_c}$), 从而判断第 z 个波峰与分频段间的相对位置.

Step 4 如果第 z 个局部波峰在第一分频段的左面, 则将该局部波峰合并到第一个分频段.

Step 5 如果第 z 个局部波峰位于两个初始分频段的中间, 则采用公式(3)分别计算该局部波峰与相邻两个分频段的引力, 并将局部波峰合并到引力较大的相邻分频段.

Step 6 如果第 z 个局部波峰在最后一个分频段的右面, 则将该局部波峰合并到最后一个分频段.

Step 7 重新计算每个分频段的质量 C_{B_m} 和质心 C_{B_c} . 重复Step 4–Step 6直到所有局部波峰均聚到某一分频段.

通过以上算法, 将分段后的振动频谱记为

$$\mathbf{x}^f = [\mathbf{x}_1^f \ \dots \ \mathbf{x}_d^f \ \dots \ \mathbf{x}_{D_v}^f], \quad (4)$$

其中: \mathbf{x}_d^f 表示振动频谱的第 d 段; 本文中 $DV = 4$, 结合先验知识, 只取前3个分频段建立子模型.

3.3 基于KPLS算法的预测子模型(Prediction sub-models based on KPLS algorithm)

划分后的分频段分别对应着筒体振动的3个不同模态: \mathbf{x}_1^f 表示由磨机筒体和ML组成的新结构的固有振动模态; \mathbf{x}_2^f 表示钢球周期性的直接撞击磨机衬板引起的冲击模态; \mathbf{x}_3^f 表示钢球之间的磨剥或冲击作用等其他原因引起的次冲击模态. 显

然, 不同频段与磨机负荷参数间的映射不同. 振动频谱具有超高维共线性的特点, 难以建立有效模型. PLS算法^[35]能够通过提取频谱中与ML参数相关的潜变量实现降维及消除共线性, 并保持频谱中尽可能多的变化信息. 假设训练样本数量是 k , 分频段频谱包含的频率个数为 p_j . KPLS算法^[36]就是将分频段频谱 $(\mathbf{x}_j^f)_{l=1}^k$ 非线性映射到高维特征空间. 即映射为 $\Phi: (\mathbf{x}_j^f)_l = \Phi((\mathbf{x}_j^f)_l)$, 从而在这个高维特征空间中执行线性的PLS算法, 从而得到原始输入空间的非线性模型. 为避免显示的非线性映射, 采用核技巧 $\tilde{\mathbf{K}}_j = (\Phi((\mathbf{x}_j^f)_l))^T \Phi((\mathbf{x}_j^f)_l)$, $l, m = 1, 2, \dots, k$ 将分频段频谱映射到高维特征空间. 对分频段的核矩阵 $\tilde{\mathbf{K}}_j$ 按下式进行中心化处理:

$$\tilde{\tilde{\mathbf{K}}}_j = (\mathbf{I} - (1/k)\mathbf{1}_k\mathbf{1}_k^T)\tilde{\mathbf{K}}_j(\mathbf{I} - (1/k)\mathbf{1}_k\mathbf{1}_k^T), \quad (5)$$

其中: \mathbf{I} 是 k 维的单位阵; $\mathbf{1}_k$ 是值为1, 长度为 k 的向量. 则基于非线性迭代偏最小二乘算法(NIPALS)和再生核希尔伯特空间(reproducing kernel Hilbert space, RKHS)理论, 可以得到如下形式的算法(以第 j 个子特征建立的第 i 个ML参数的子模型为例):

Step 1 随机初始化 \mathbf{u}_j .

Step 2 $\mathbf{t}_j = \tilde{\tilde{\mathbf{K}}}_j^T \mathbf{u}_j, \mathbf{t}_j \leftarrow \mathbf{t}_j / \|\mathbf{t}_j\|$.

Step 3 $\mathbf{c}_j = \mathbf{y}_j^T \mathbf{t}_j$.

Step 4 $\mathbf{u}_j = \mathbf{y}_j \mathbf{c}_j, \mathbf{c}_j \leftarrow \mathbf{c}_j / \|\mathbf{c}_j\|$.

Step 5 重复Step 2–Step 4, 直至收敛.

Step 6 计算残差: $\tilde{\tilde{\mathbf{K}}}_j \leftarrow (\mathbf{I} - \mathbf{t}_j \mathbf{t}_j^T)\tilde{\tilde{\mathbf{K}}}_j(\mathbf{I} - \mathbf{t}_j \mathbf{t}_j^T), \mathbf{y}_j \leftarrow \mathbf{y}_j - \mathbf{t}_j \mathbf{t}_j^T \mathbf{y}_j$.

Step 7 如果残差满足收敛条件, 则结束, 否则转Step 2.

最终训练数据 $(\mathbf{x}_j^f)_{l=1}^k$ 基于KPLS回归的ML预测子模型可表示为

$$\hat{y}_{ji} = \Phi_j \mathbf{B}_{ji} = \tilde{\tilde{\mathbf{K}}}_j \mathbf{U}_j (\mathbf{T}_j^T \tilde{\tilde{\mathbf{K}}}_j \mathbf{U}_j)^{-1} \mathbf{T}_j^T \mathbf{y}_i. \quad (6)$$

对于测试样本 $(\mathbf{x}_j^f)_{l=1}^{k_t}$, 则要首先对测试样本按下式进行标度处理:

$$\tilde{\tilde{\mathbf{K}}}_{t,j} = (\mathbf{K}_{t,j} \mathbf{I} - (1/k)\mathbf{1}_{k,t}\mathbf{1}_{k,t}^T)\mathbf{K}_j(\mathbf{I} - (1/k)\mathbf{1}_k\mathbf{1}_k^T), \quad (7)$$

其中: $\mathbf{K}_{t,j}$ 是测试样本的核矩阵, $\mathbf{K}_{t,j} = \mathbf{K}_j((\mathbf{x}_{t,j}^f)_l, (\mathbf{x}_j^f)_m)$; k_t 是测试样本的个数; $\mathbf{1}_{k,t}$ 是值为1, 长度为 k_t 的向量. 则测试样本基于KPLS回归的ML参数预测子模型为

$$\hat{y}_{t,j,i} = \Phi_{t,j} \mathbf{B}_{ji} = \tilde{\tilde{\mathbf{K}}}_{t,j} \mathbf{U}_j (\mathbf{T}_j^T \tilde{\tilde{\mathbf{K}}}_j \mathbf{U}_j)^{-1} \mathbf{T}_j^T \mathbf{y}_i. \quad (8)$$

3.4 基于信息熵的在线集成模型(On-line ensemble model based on information entropy)

首先根据训练数据的预测值采用文献[34]提出

的熵值法确定各子模型的初始加权系数; 然后再针对每个测试样本, 进行集成预测和依据预报误差在线自适应更新权重系数。

1) 依据训练样本数据计算初始权值。

设 y_{il} 为训练样本中第 i 个ML参数在时刻 l 的实际值, \hat{y}_{jil} 为第 j 个预测子模型对第 i 个ML参数在时刻 l 的预报值, 则计算初始加权系数的步骤如下:

Step 1 计算第 j 个预测子模型在每个时刻的预测相对误差为

$$e_{jil} = |(y_{il} - \hat{y}_{jil})/y_{il}|. \quad (9)$$

如果 $|(y_{il} - \hat{y}_{jil})/y_{il}| \geq 1$, 则 $e_{jil} = 1$. 其中, $j = 1, 2, 3$ 分别表示LF, MF和HF分频段的预测模型; $i = 1, 2, 3$ 分别表示ML参数: 料球比、矿浆浓度和充填率; $l = 1, \dots, k$, k 为训练样本的个数。

Step 2 计算第 j 个预测模型的预测相对误差的比重 p_{jil}^k :

$$p_{jil}^k = e_{jil} / (\sum_{l=1}^k e_{jil}). \quad (10)$$

Step 3 计算第 j 个预测模型的预测相对误差的熵值 E_{jil}^k :

$$E_{jil}^k = (1/\ln(k)) \sum_{l=1}^k p_{jil}^k \cdot \ln(p_{jil}^k). \quad (11)$$

Step 4 计算第 j 个预测模型的加权系数 W_{ji}^k :

$$W_{ji}^k = (1/(z-1)) [1 - (1 - E_{ji}^k) / \sum_{j=1}^z (1 - E_{ji}^k)], \quad (12)$$

其中: $\sum_{j=1}^z W_{ji}^k = 1$, z 是针对第 i 个ML参数预测模型的个数, 本文中 $z = 3$.

2) 针对测试样本数据的集成预测和权值的在线更新。

在模型的在线使用中, 首先根据初始权值进行集成预测, 然后根据新样本的预测相对误差的变化更新子模型权值, 步骤如下:

Step 1 对新样本进行预测: 将新样本记为 $\mathbf{x}_{t,j}^f$,

其对应的KPLS子模型的预测值记为 $\hat{y}_{t,j}^{k+1}$, 则加权集成模型的预测输出为

$$\hat{y}_{t,i}^{k+1} = \sum_{j=1}^z W_{ji}^k \hat{y}_{t,j}^{k+1}. \quad (13)$$

Step 2 计算新样本的预测相对误差

$$e_{ji}^{k+1} = |(y_{t,i}^{k+1} - \hat{y}_{t,j}^{k+1})/y_{t,i}^{k+1}|. \quad (14)$$

并记 $e_{jil}^{k+1} = [e_{jil} e_{ji}^{k+1}]$.

Step 3 更新第 j 个预测模型的预测相对误差的比重 p_{jil}^{k+1} :

$$p_{jil}^{k+1} = e_{jil}^{k+1} / (\sum_{l=1}^{k+1} e_{jil}^{k+1}). \quad (15)$$

Step 4 更新第 j 个预测模型的预测相对误差的熵值 E_{jil}^k :

$$E_{jil}^{k+1} = (1/\ln(k+1)) \sum_{l=1}^{k+1} p_{jil}^{k+1} \cdot \ln(p_{jil}^{k+1}). \quad (16)$$

Step 5 更新第 j 个预测模型的加权系数 W_{ji}^{k+1} , 并替代 W_{ji}^k :

$$W_{ji}^{k+1} = (1/(z-1)) [1 - (1 - E_{ji}^{k+1}) / \sum_{j=1}^z (1 - E_{ji}^{k+1})]. \quad (17)$$

4 仿真结果分析(Simulation results analysis)

针对实验室内小型球磨机筒体振动的数据, 进行了数据重采样、离群点剔除、中心化处理后采用welch法计算了功率谱密度(PSD), 其时频曲线详见文献[12]. 由振动频谱的曲线可知, 频谱至少可以被分割为3个频段. 采用自动分割模块将频谱准确划分为4个分频段, 其频率范围是102~2385 Hz(LF), 2385~4122 Hz(MF), 4122~7227 Hz(HF), 和7227~11000 Hz(HHF). 本文取前3个频段, 采用13个样本基于KPLS算法建立MBVR, PD和CVR的子模型, 其中核函数采用径向基函数(RBF), 核半径均为10. 分频段频谱和ML参数对应的各个潜变量的方差变化率如图3-5所示。

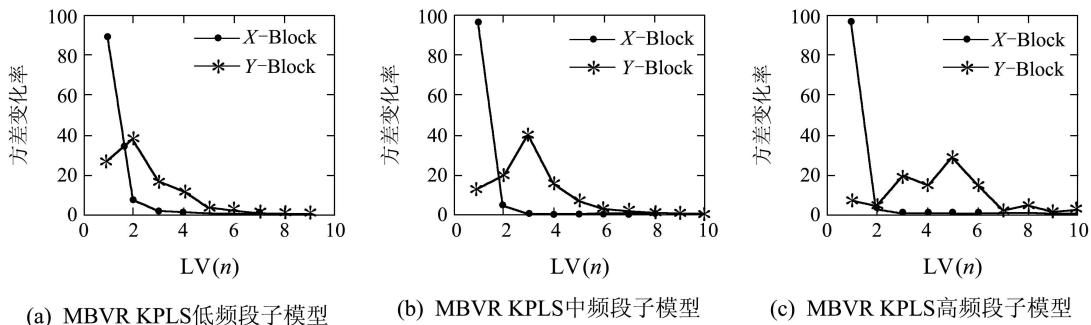


图 3 料球比子模型的方差变化率

Fig. 3 Variance percent ratio of the MBVR sub-models

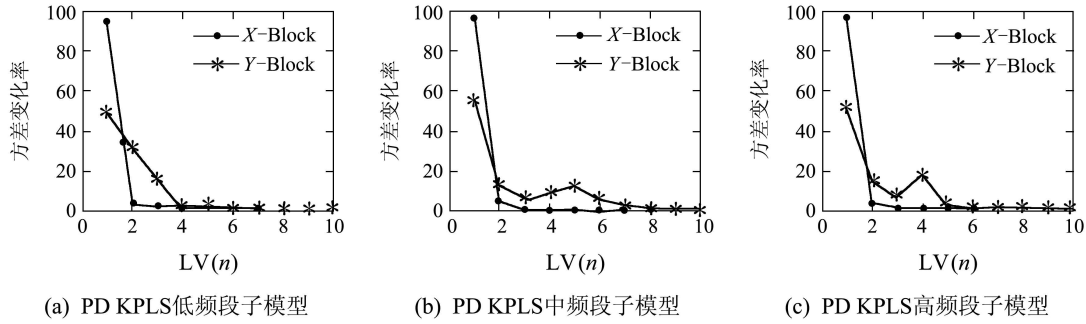


图4 矿浆浓度子模型的分差变化率

Fig. 4 Variance percent ratio of the PD sub-models

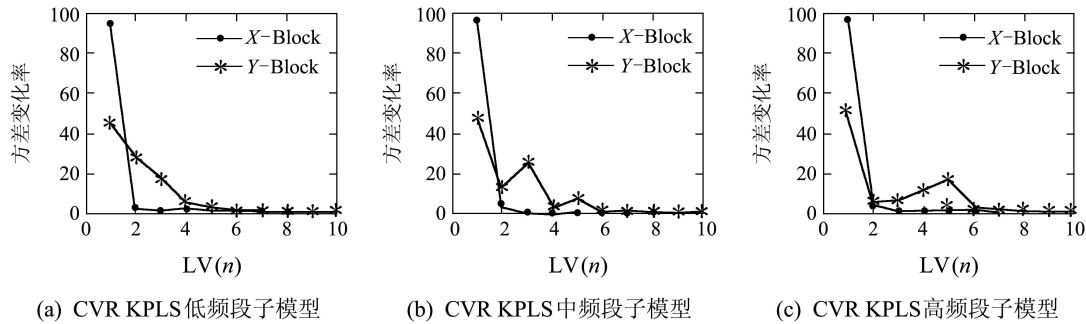


图5 充填率子模型的分差变化率

Fig. 5 Variance percent ratio of the CVR sub-models

由图3-5可知: ① 不同频段的振动频谱的变化率不同, 以低频段最低, 中频段次之, 高频段最高, 这与研磨机理相符合: 低频段是磨机筒体和负荷组成的机械结构体的固有振动模态, 并且实验过程中, 球负荷保持不变, 因水和矿石负荷导致的振动模态的变化相对较弱; 中频和高频则是负荷的冲击引起的, 矿浆浓度和充填率对冲击力的影响比较明显. ② 不同的ML负荷参数与各个频段的相关性也不同: 料球比的方差变化率与频谱方差变化率并不一致, 这也是料球比模型精度较低的原因之一; 矿浆浓度和充填率则是与中高频段的相关性稍高于低频段, 其方差变化率与频谱的方差变化率一致.

在建立分频段KPLS子模型后, 按文中方法得到初始训练权值. 因实验在批次磨矿条件下进行, 测试样本与训练样本的采集方式相同, 分批获得后进行模型的测试. 测试样本更新后的权值如图6-8所示. 由图6-8可知, 料球比集成预测模型的权系数均在0.31和0.36间波动, 变化范围较小. 充填率集成预测的权系数在0.30和0.35间波动, 中频段和低频段的权系数略高于高频段, 而且权值更新的波动也比较的平缓. 矿浆浓度的权系数波动比较大, 其范围是0.25-0.43, 根据训练数据得到的初始权值和更新后的权值差异较大, 这与建模选择的样本覆盖的范围相关, 但是更新后中频段的权值最大, 这与研磨过程的机理分析相符. 图6-8同时表明需要更多接近实

际工业过程的实验数据进行深入验证.

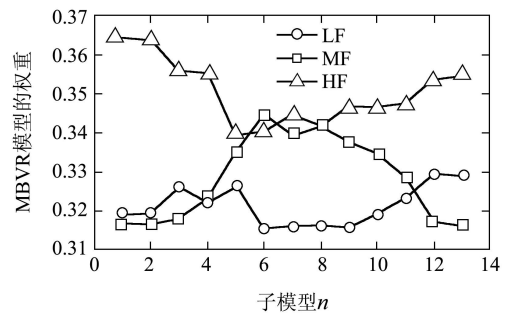


图6 料球比不同频段子模型的权值

Fig. 6 Weighting values of different spectral segments of the MBVR sub-models

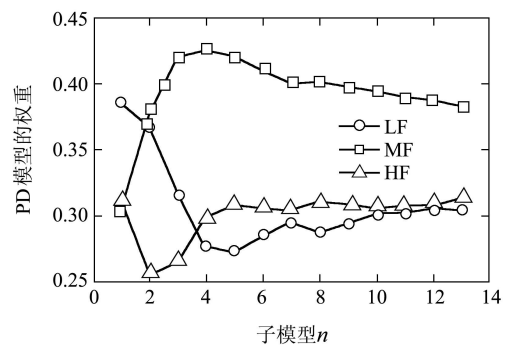


图7 矿浆浓度不同频段子模型的权值

Fig. 7 Weighting values of different spectral segments of the PD sub-models

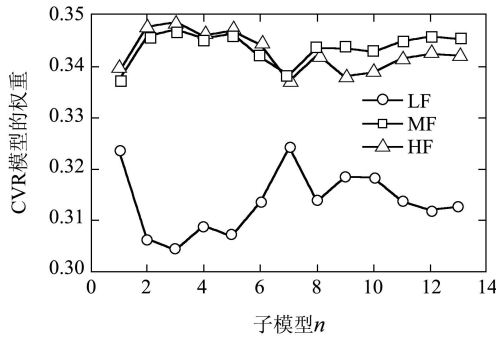


图 8 充填率不同频段子模型的权值

Fig. 8 Weighting values of different spectral segments of the CVR sub-models

基于本文所提方法建立的ML集成预测模型与分频段预测模型的结果如图9-11所示。

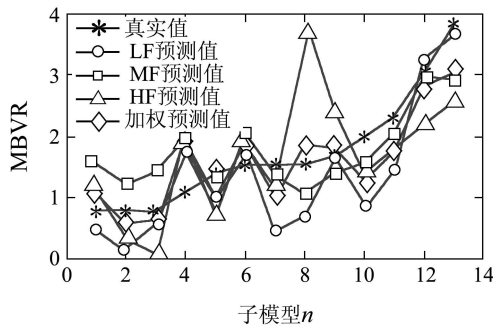


图 9 料球比模型预测曲线

Fig. 9 Prediction curves of the MBVR models

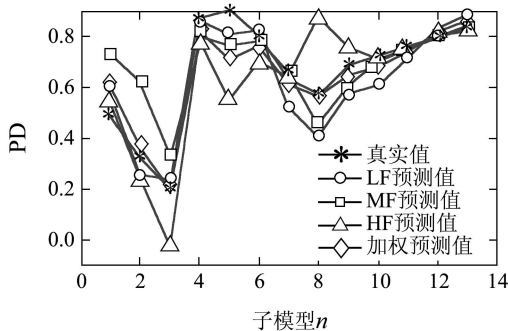


图 10 矿浆浓度模型预测曲线

Fig. 10 Prediction curves of the PD models

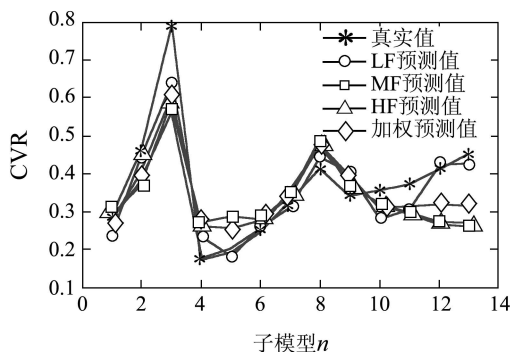


图 11 充填率模型预测曲线

Fig. 11 Prediction curves of the CVR models

为了与本文提出的方法相比较,建立了基于全谱的主元回归(PCR), PLS, GA-PLS^[16]和KPLS模型;基于以PCA提取分频段的特征,并串行组合后的特征变量为模型输入的PCA-SVM^[17], PCA-LSSVM^[18]模型;基于在线自适应加权融合算法^[28]的PLS, KPLS集成模型,其统计结果见表1。

在表1中,“V-Weighted”和“V-O-Weighted”分别表示不在线和在线的自适应加权融合方法;“E-Weighted”和“E-O-Weighted”分别表示不在线和在线的基于信息熵的加权融合方法;“LVs/PCs”表示潜变量和主元个数;“LF-MF-HF”和“LF+MF+HF”分别表示建模的数据集是分频段和全谱.基于PCA的模型,在低、中和高频段选择的主元个数分别为5, 3和1个.从仿真结果看,集成预测模型的预测精度高于单一预测模型;基于信息熵集成算法的预测精度高于基于自适应加权融合算法;核PLS方法的精度高于PLS方法,表明了振动频谱与ML参数间存在非线性关系;基于PCA算法模型精度低于集成预测方法,说明了PLS/KPLS算法用于频谱建模的合理性,但是,通过适当的选择主元个数和模型参数,基于结构风险最小的PCA-SVM/PCA-LSSVM可以具有更好的泛化能力。

但是,在线集成预测方法的精度却稍微低于非权值更新的预测算法,原因一是样本点较少且多数都是异常工况,难以充分说明问题;二是进行了基于每个样本的权值更新,恶化了预测性能.为适应工业过程的时变特性,需要更深入的实验验证在线的权值更新算法,同时进行KPLS更新算法的研究。

注 1 本文的建模精度与文献 [16-19]相比相差不大,其原因之一是建模数据的差别:本文中的ML参数料球比的计算方法与振动频谱的划分范围与文献 [16-19]不同.如本文中料球比的计算采用了文献 [16]中的公式(3)基于经验给定介质空隙率的方法;而 [16-19]则是基于文献 [31]给出的钢球在空间随机排列的介质孔隙率的公式,采用文献 [16]中的公式(3)计算空隙体积的方法;本文采用分频段自动识别方法划分的分频段范围是“102~2385 Hz, 2385~4122 Hz, 4122~7227 Hz”,而文献 [17-18]手动划分的频段范围是“100~1800 Hz, 1800~4000 Hz, 4000~7500 Hz”;原因之二是本文未进行KPLS算法的核类型和核参数的优化选择,如前文中所提“核函数采用径向基函数(RBF),核半径均为10”;原因之三是本文中基于PCA的方法未进行主元的优化选择,而是为了与KPLS算法选择的潜变量具有相近的方差变化率,设定了主元的贡献率的阈值为99%.同时,由表 1 可知,本文建立的充填率集成模型的预测精度

低于低频段的子模型;同时考虑到筒体振动和振声频谱及磨机电流所包含信息的冗余与互补性,需要深入的进行优

化KPLS子模型的核类型及核参数、建立基于信息融合及选择性集成的KPLS建模方法的研究。

表1 软磨机负荷软测量建模统计结果

Table 2 Errors statistic results of soft sensor models

模型	数据集	MBVR	MBVR	PD	PD	CVR	CVR	平均
		LVs/PCs	RMSSEs	LVs/PCs	RMSSEs	LVs/PCs	RMSSEs	
PCR	LF+MF+HF	10	0.4747	10	0.2662	10	0.2494	0.3301
PLS	LF+MF+HF	7	0.5626	4	0.2140	2	0.2434	0.34001
GAPLS	LF+MF+HF	2	0.3279	1	0.3406	2	0.2847	0.3177
PCA-SVM	LF+MF+HF	(5, 3, 1)	0.4303	(5, 3, 1)	0.1483	(5, 3, 1)	0.2732	0.2839
PCA-LSSVM	LF+MF+HF	(5, 3, 1)	0.4458	(5, 3, 1)	0.3783	(5, 3, 1)	0.1675	0.3305
PLS	LF	9	0.6755	11	0.4346	11	0.2460	0.4520
PLS	MF	4	0.7650	3	0.3134	3	0.2795	0.4526
PLS	HF	5	2.2921	5	0.9190	6	0.6811	1.2974
PLS+V-Weighted	LF-MF-HF	(9, 4, 5)	0.7063	(11, 3, 5)	0.1878	(11, 3, 6)	0.2789	0.3910
PLS+V-O-Weighted	LF-MF-HF	(9, 4, 5)	0.8191	(11, 3, 5)	0.1500	(11, 3, 6)	0.3335	0.4342
PLS+E-Weighted	LF-MF-HF	(9, 4, 5)	0.5286	(11, 3, 5)	0.1491	(11, 3, 6)	0.2643	0.3140
PLS+E-O-Weighted	LF-MF-HF	(9, 4, 5)	0.6782	(11, 3, 5)	0.1752	(11, 3, 6)	0.2607	0.3713
KPLS	LF+MF+HF	8	0.4725	3	0.4641	11	0.1736	0.3701
KPLS	LF	9	0.4529	4	0.1553	12	0.1579	0.2554
KPLS	MF	3	0.5132	1	0.3399	1	0.2906	0.3812
KPLS	HF	4	0.6201	7	0.3750	1	0.2866	0.4272
KPLS+V-Weighted	LF-MF-HF	(9, 3, 4)	0.3483	(4, 1, 7)	0.1220	(12, 1, 1)	0.2501	0.2401
KPLS+V-O-Weighted	LF-MF-HF	(9, 3, 4)	0.3339	(4, 1, 7)	0.1355	(12, 1, 1)	0.2293	0.2329
KPLS+E-Weighted	LF-MF-HF	(9, 3, 4)	0.3344	(4, 1, 7)	0.1133	(12, 1, 1)	0.2084	0.2187
KPLS+E-O-Weighted	LF-MF-HF	(9, 3, 4)	0.3345	(4, 1, 7)	0.0937	(12, 1, 1)	0.2294	0.2192

5 结论(Conclusion)

本文基于磨机研磨过程的研磨机理、谱频段的自动分割方法、KPLS的分频段子模型及基于信息熵值的集成加权融合方法,针对实验球磨机磨矿过程的ML参数的软测量问题开展研究,论文的主要工作和结论如下: 1)定性分析了球磨机的研磨机理,指出了ML参数与筒体振动频谱间的映射关系,并通过分频段KPLS子模型的预测结果,表明了不同分频段频谱与ML参数间的非线性映射关系,验证了机理分析的正确性; 2)通过频谱分频段的自动分割算法,实现了频段的精确划分; 3)基于信息熵的集成预测模型的平均精度高于单一模型,验证了所提方法的有效性,但是在线自适应加权方法需要进一步的验证.该方法为湿式球磨机的ML参数的软测量提供了新的思路,应用前景广阔.但是本方法仍然需要进行接近工业实际工况的实验磨机和工业磨机的验证实验.进一步的研究方向是融合筒体振声、磨机电流等信号进行多传感器信息融合及选择性集成算法和KPLS在线更新算法的研究,提高模型的预测性能及适应工业过程的时变特性.

参考文献(References):

- [1] ZHOU P, CHAI T Y, WANG H. Intelligent optimal-setting control for grinding circuits of mineral processing[J]. *IEEE Transactions on Automation Science and Engineering*, 2009, 6(4): 730 - 743.
- [2] 汤健, 赵立杰, 岳恒, 等. 磨机负荷检测方法研究综述[J]. *控制工程*, 2010, 17(5): 565 - 570.
(TANG Jian, ZHAO Lijie, YUE Heng, et al. Present status and future developments of detection method for mill load[J]. *Control Engineering of China*, 2010, 17(5): 565 - 570.)
- [3] HU G, OTAKI H, WATANUKI K. Motion analysis of a tumbling ball mill based on non-linear optimization[J]. *Minerals Engineering*, 2000, 13(8/9): 933 - 947.
- [4] BEHERA B, MISHRA B K, MURTY C V R. Experimental analysis of charge dynamics in tumbling mills by vibration signature technique[J]. *Minerals Engineering*, 2007, 20(1): 84 - 91.
- [5] 张立岩, 柴天佑. 氧化铝回转窑制粉系统磨机负荷的智能控制[J]. *控制理论与应用*, 2010, 27(11): 1471 - 1478.
(ZHANG Liyan, CHAI Tianyou. Intelligent mill-load control for pulverizing system in alumina sintering process[J]. *Control Theory & Applications*, 2010, 27(11): 1471 - 1478.)
- [6] MORI H, MIO H, KANO J, et al. Ball mill simulation in wet grinding using a tumbling mill and its correlation to grinding rate[J]. *Powder Technology*, 2004, 143/144(25): 230 - 239.
- [7] 周平, 柴天佑. 磨矿过程磨机负荷的智能监测与控制[J]. *控制理论与应用*, 2008, 25(6): 1095 - 1099.
(ZHOU Ping, CHAI Tianyou. Intelligent monitoring and control of mill load for grinding processes[J]. *Control Theory & Applications*, 2008, 25(6): 1095 - 1099.)
- [8] 白锐, 柴天佑. 基于数据融合与案例推理的球磨机负荷优化控制[J]. *化工学报*, 2009, 60(7): 1746-1751.

- (BAI Rui, CHAI Tianyou. Optimization control of ball mill load in blending process with data fusion and case-based reasoning[J]. *Journal of Chemical Industry and Engineering*, 2009, 60(7): 1746 – 1751.)
- [9] 李勇, 邵诚. 灰色软测量在介质填充率检测中的应用研究[J]. 中国矿业大学学报, 2006, 35(4): 549 – 555.
(LI Yong, SHAO Cheng. Application research of grey soft sensor for charge ratio of media[J]. *Journal of China University of Mining Technology*, 2006, 35(4): 549 – 555.)
- [10] 李勇, 邵诚. 一种新的灰关联分析算法在软测量中的应用[J]. 自动化学报, 2006, 32(2): 311 – 317.
(LI Yong, SHAO Cheng. Application of a novel grey relation analysis algorithm to soft sensor[J]. *Acta Automatica Sinica*, 2008, 32(2): 311 – 317.)
- [11] ZENG Y, FORSSBERG E. Monitoring grinding parameters by signal measurements for an industrial ball mill[J]. *International Journal of Mineral Processing*, 1993, 40(1): 1 – 16.
- [12] ZENG Y, FORSSBERG E. Monitoring grinding parameters by vibration signal measurement – a primary application[J]. *Minerals Engineering*, 1994, 7(4): 495 – 501.
- [13] SPENCER S J, CAMPBELL J J, WELLER K R, et al. Acoustic emissions monitoring of SAG mill performance[C] // *Proceedings of the Second International Conference on Intelligent Processing and Manufacturing of Materials*. New York: IEEE, 1999, 2: 939 – 946.
- [14] GUGEL K, MOON R M. Automated mill control using vibration signal processing[C] // *Proceedings of IEEE Cement Industry Technical Conference Record*. Charleston, SC: IEEE, 2007, 17 – 25.
- [15] 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[J]. *Minerals Engineering*, 2009, 22(14): 1200 – 1208.
- [16] TANG J, ZHAO L J, ZHOU J W, et al. Experimental analysis of wet mill load based on vibration signals of laboratory-scale ball mill shell[J]. *Minerals Engineering*, 2010, 23(9): 720 – 730.
- [17] TANG J, ZHAO L J, YU W, et al. Soft sensor modeling of ball mill load via principal component analysis and support vector machines[C] // *Advances in Neural Network Research and Applications, Lecture Notes in Electrical Engineering*, 2010, 67: 803 – 810.
- [18] 汤健, 赵立杰, 岳恒, 等. 基于多源数据特征融合的球磨机负荷软测量方法[J]. 浙江大学学报(工学版), 2010, 44(7): 1406 – 1413.
(TANG Jian, ZHAO Lijie, YUE Heng, et al. Soft sensor method of ball mill load based on multi-source data feature fusion[J]. *Journal of Zhejiang University (Engineering Science)*, 2010, 44(7): 1406 – 1413.)
- [19] 汤健, 郑秀萍, 赵立杰, 等. 基于频域特征提取与信息融合的磨机负荷软测量[J]. 仪器仪表学报, 2010, 31(10): 2161 – 2167.
(TANG Jian, ZHENG Xiuping, ZHAO Lijie, et al. Mill load soft-sensing based on feature extraction of frequency domain and information fusion[J]. *Chinese Journal of Scientific Instrument*, 2010, 31(10): 2161 – 2167.)
- [20] LIU J L. On-line soft sensor for polyethylene process with multiple production grades[J]. *Control Engineering Practice*, 2007, 15(7): 769 – 778.
- [21] JOLLIFFE I T. *Principal Component Analysis*[M]. Berlin, Germany: Springer, 2002.
- [22] 刘利军, 樊江玲, 张志谊, 等. 密频系统模态参数辨识及其振动控制的研究进展[J]. 振动与冲击, 2007, 26(4): 109 – 115.
(LIU Lijun, FAN Jiangling, ZHANG Zhiyi, et al. Study progress in model parameters identification and vibration control of systems with crowded modes[J]. *Journal of Vibration and Shock*, 2007, 26(4): 109 – 115.)
- [23] QIN S J. Statistical process monitoring: basics and beyond[J]. *Journal of Chemometrics*, 2003, 17(8/9): 480 – 502.
- [24] HANSEN L K, SALAMON P. Neural network ensembles[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1990, 12(10): 993 – 1001.
- [25] NIU D P, WANG F L, ZHANG L L, et al. Neural network ensemble modeling for nosiheptide fermentation process based on partial least squares regression[J]. *Chemometrics and Intelligent Laboratory Systems*, 2011, 105(1): 125 – 130.
- [26] BREUER L, HUISMAN J A, WILLEMS P, et al. Assessing the impact of land use change on hydrology by ensemble modeling (LUCHEM), I: model intercomparison with current land use[J]. *Advances in Water Resources*, 2009, 32(2): 129 – 146.
- [27] BREUER L, HUISMAN J A, WILLEMS P, et al. Assessing the impact of land use change on hydrology by ensemble modeling (LUCHEM), II: ensemble combinations and predictions[J]. *Advances in Water Resources*, 2009, 32(2): 147 – 158.
- [28] XU L J, ZHANG J Q, YAN Y. A wavelet-based multisensor data fusion algorithm[J]. *IEEE Transactions on Instrumentation and Measurement*, 2004, 53(6): 1539 – 1544.
- [29] 王春生, 吴敏, 曹卫华, 等. 铅锌烧结配料过程的智能集成建模与综合优化方法[J]. 自动化学报, 2009, 35(5): 605 – 612.
(WANG Chunsheng, WU Min, CAO Weihua, et al. Intelligent integrated modeling and synthetic optimization for blending process in lead-zinc sintering[J]. *Acta Automatica Sinica*, 2009, 35(5): 605 – 612.)
- [30] MOHAMED S. Estimating market shares in each market segment using the information entropy concept[J]. *Applied Mathematics and Computation*, 2007, 190(2): 1735 – 1739.
- [31] CHEN B C. *Grinding Principle*[M]. Beijing: Metallurgical Industry Press, 1980.
- [32] DONG H, MOYS M H. Assessment of discrete element method for one ball bouncing in a grinding mill original research article[J]. *International Journal of Mineral Processing*, 2002, 65(3/4): 213 – 226.
- [33] RICHARDSON M H, FORMENTI D L. Global curve fitting of frequency response measurements using the rational fraction polynomial method[C] // *Proceedings of the 3rd International Modal Analysis Conference*. Orlando, FL: Union Coll, Schenectady, 1985: 390 – 397.
- [34] 叶庆卫, 汪同庆. 基于幅谱分割的粒子群最优模态分解研究与应用[J]. 仪器仪表学报, 2009, 30(8): 547 – 588.
(YE Qingwei, WANG Tongqing. Optimization modal analysis with PSO based on spectrum segmentation[J]. *Chinese Journal of Scientific Instrument*, 2009, 30(8): 547 – 588.)
- [35] WOLD S, SJOSTORM M, ERIKSSON L. PLS-regression: a basic tool of chemometrics[J]. *Chemometrics and Intelligent Laboratory Systems*, 2001, 58(2): 109 – 130.
- [36] ROSIPAL R, TREJO L J. Kernel partial least squares regression in reproducing kernel Hilbert space[J]. *Journal of Machine Learning Research*, 2002, 2(2): 97 – 123.

作者简介:

汤健 (1974—), 男, 目前研究方向为综合自动化系统、基于数据的软测量建模等, E-mail: tjian001@126.com;

柴天佑 (1947—), 男, 博士生导师, 中国工程院院士, 主要研究方向为自适应控制、多变量智能解耦控制与流程工业综合自动化等, E-mail: tychai@mail.neu.edu.cn;

赵立杰 (1972—), 女, 硕士生导师, 主要研究方向为建模、故障诊断及复杂工业过程优化等, E-mail: zlj_lunlun@163.com;

岳恒 (1968—), 男, 硕士生导师, 主要研究方向为多变量智能解耦、复杂工业过程的建模与优化等, E-mail: hengyue@mail.neu.edu.cn;

郑秀萍 (1966—), 女, 主要研究方向为流程工业综合自动化研究与应用等, E-mail: xiupingzheng@263.net.