

## 基于泡沫图像特征的金锑浮选入矿品位估计

吴佳, 谢永芳, 阳春华<sup>†</sup>, 桂卫华

(中南大学信息科学与工程学院, 湖南长沙 410083)

**摘要:** 入矿品位是金锑浮选加药量控制的重要依据. 针对入矿品位在线检测困难的问题, 提出一种基于泡沫图像特征的入矿品位估计方法. 该估计方法首先针对样本数据中存在的确定性, 提出一种基于核主元分析(KPCA)和模糊C均值聚类-概率支持向量回归(FCM-PSVR)的建模方法, 然后利用泡沫图像特征与加药量等数据建立起金锑入矿品位和精矿品位的估计模型, 最后采用基于专家规则的方法对入矿品位估计结果的可信度进行评价. 该方法在金锑浮选工艺中进行了工业验证, 为指导金锑浮选加药量的控制起到了重要作用.

**关键词:** 浮选; 不确定性; 支持向量机; 模糊聚类

中图分类号: TP273 文献标识码: A

## Mineral concentration estimation of feed ore in gold and stibium flotation based on froth image features

WU Jia, XIE Yong-fang, YANG Chun-hua<sup>†</sup>, GUI Wei-hua

(School of information Science and Engineering, Central South University, Changsha Hunan 410083, China)

**Abstract:** The regulation of reagents heavily depends on the mineral concentration of feed ore in flotation. Since it is difficult to measure the concentration of feed ore online, an estimator based on froth image features is developed in this paper. At first, considering the uncertainty existing in the sample data, we develop a soft-sensor method based on kernel principal component analysis (KPCA) and fuzzy C-means clustering-probabilistic support vector regress (FCM-PSVR). Then, estimators for the gold and stibium concentration of feed ore and concentrate are built. Finally, the output of the estimators of feed ore is evaluated based on expert rules. The estimator has been validated in the gold and stibium froth flotation, and it plays an important role in guiding the reagent addition in gold and stibium flotation.

**Key words:** flotation; uncertainty; support vector machine; fuzzy cluster

### 1 引言(Introduction)

在金锑浮选过程中, 工况变化的主要因素为入矿条件. 其中金和锑的入矿品位变化范围分别为1.2~3.2 g/t和1%~2%. 入矿品位难以在线测量, 人工往往因不能及时发现入矿品位变化而造成加药量调控大滞后, 出现生产指标不达标, 甚至浮选流程长时间不稳定等异常情况. 因此, 在线估计入矿品位对浮选加药量的自动控制具有重要的意义.

随着浮选生产自动化技术的发展, 泡沫图像越来越受到学者们的关注和重视<sup>[1]</sup>. He等<sup>[2]</sup>实现了基于泡沫图像特征的硫浮选监测. Supomo等<sup>[3]</sup>提出一种基于泡沫速度特征的矿浆液位控制方法. Liu等<sup>[4]</sup>提出一种基于泡沫图像特征的串级控制器. 由此可见, 挖

掘泡沫图像所能提供的丰富信息具有重要意义. 浮选泡沫状态是入矿品位与加药量共同作用的结果, 因此, 可利用泡沫状态与加药量等实时数据反推入矿品位. 由于入矿品位与精矿品位的变化趋势相同, 因此还可利用泡沫状态估计精矿品位, 再比对入矿品位的估计结果来评价估计结果的可信度. 综上所述, 本文针对入矿品位在线检测困难的问题, 提出一种基于泡沫图像特征的入矿品位在线估计方法. 实践表明, 该方法能有效估计入矿品位, 并为金锑浮选加药量控制提供更为精确的决策依据.

### 2 金锑浮选工艺及数据不确定性(Gold and stibium flotation and uncertainty of data)

金锑浮选的工艺流程如图1所示. 含金矿物粒子通

收稿日期: 2014-04-07; 录用日期: 2014-09-02.

<sup>†</sup>通信作者. E-mail: ychh@csu.edu.cn; Tel.: +86 731-88836876.

国家自然科学基金项目(61134006, 61473318), 国家杰出青年科学基金(61025015)资助.

Supported by State Key Program of National Natural Science of China (61134006, 61473318) and National Science Foundation for Distinguished Young Scholars of China (61025015).

过金粗选及金精选 I、II 等子过程被优先浮选, 含锑矿物粒子通过锑粗选及锑精选 I、II 等子过程被后浮选, 锑扫选 I、II 用来提高两种矿物回收率. 在金粗选、锑粗选中加入黄药、二号油等多种药剂. 同时采用文献[5-6]中的方法处理图1中监控点的泡沫图像, 得到相应的泡沫图像特征, 包括: 速度、稳定度等动态图像特征和灰度均值、泡沫尺寸大小及尺寸方差、负载率等静态图像特征. 由于存在自相关性等冗余信息, 采用文献[7]中的约简方法得到关键图像特征集. 由于浮选过程中强烈的搅拌及鼓风作用使得矿物粒子与泡沫分布非均匀, 从而造成样本数据存在不确定性, 数据的不确定性严重影响入矿品位在线估计的精准性.

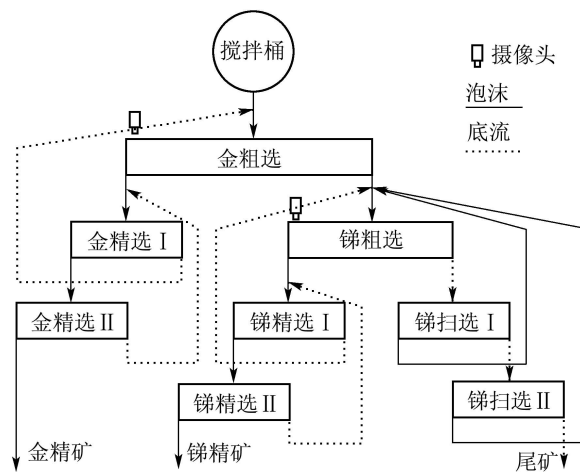


图 1 金锑浮选工艺流程

Fig. 1 Process flow of gold and stibium flotation

对于多变量的复杂系统, 采用多T-S模糊推理系统进行建模是一种行之有效的方法. 如文献[8]中采用多T-S模糊推理方法建立聚氯乙烯聚合转化率的预测模型, 然而该方法未考虑到数据不确定性的影响. 而针对数据的不确定性, 近年来云推理方法<sup>[9]</sup>、二型模糊逻辑推理系统<sup>[10]</sup>和概率支持向量机(probabilistic support vector machine, PSVM)<sup>[11]</sup>等方法相继被提出. 文献[9]提出一种基于云神经网络模型的建模方法, 文献[12]中提出一种基于二型模糊神经网络的建模方法. 但这些方法也存在神经网络模型固有的缺点, 如收敛速度慢. 文献[11]中提出一种PSVM方法, 该方法核心思想是通过随机抽样建立分布支持向量机(support vector machine, SVM)系统, 然后估计SVM参数的分布密度函数, 最后通过优化SVM的参数期望获取最终决策函数. 虽然PSVM算法是针对分类器中的数据不确定性而提出的, 但是该方法的思想也可用于解决回归模型中的不确定性. 综上所述, 本文首先提出一种基于核主元分析(kernel principal component analysis, KPCA)<sup>[13]</sup>与FCM-PSVR(probabilistic support vector regress)的多T-S模糊推理的建模方法.

### 3 基于KPCA与FCM-PSVR的多T-S建模方法(Multiple T-S modeling method based on KPCA and FCM-PSVR)

因为变量间存在自相关性<sup>[7]</sup>, 所以首先采用KPCA方法进行特征转换, 设样本集  $X = [x_1, x_2, \dots, x_M]'$ , 其中:  $x_i$  为第  $i$  个样本数据,  $M$  为样本总数, 设输入数据通过非线性函数  $\phi(\cdot)$  映射到特征空间  $F$ , 且  $F$  中的数据满足中心化条件, 即

$$\sum_{i=1}^M \phi(x_i) = 0, \quad (1)$$

则  $F$  中的协方差矩阵为

$$C = \frac{1}{M} \sum_{i=1}^M \phi(x_i) \phi^T(x_i). \quad (2)$$

现求式(2)中的特征值与特征向量, 即有

$$CV = \lambda V, \quad (3)$$

其中:  $\lambda \geq 0$  为特征值,  $V$  为特征向量. 在式(3)两边左乘  $\phi(x_j)$ , 则有

$$\phi(x_j)CV = \lambda \phi(x_j)V, \quad j = 1, 2, \dots, M, \quad (4)$$

特征向量  $V$  可由  $F$  中的样本张成, 即

$$V = \sum_{v=1}^M \alpha_v \phi(x_v), \quad (5)$$

引入核函数

$$K_{ij} = K(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle, \quad (6)$$

将式(2)和式(5)代入式(4), 并简化得到

$$K\alpha = M\lambda\alpha, \quad (7)$$

其中:  $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_M]'$ , 特征值  $\lambda_k (\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_M)$  和特征向量  $\alpha^k (k = 1, 2, \dots, M)$ ,  $K$  为核函数矩阵. 归一化特征向量  $V^k$ , 样本  $x$  在特征空间  $V^k$  上的映射为

$$t_k = \langle V^k, \phi(x) \rangle = \sum_{i=1}^M \alpha_i^k \langle \phi(x_i), \phi(x) \rangle = \sum_{i=1}^M \alpha_i^k K(x_i, x), \quad (8)$$

其中  $t_k$  为特征空间中的第  $k$  个主元, 特征值较小的主元被认为是噪声引起的. 主元数量的选取一般依据规则

$$\left( \sum_{i=1}^p \lambda_i / \sum_{j=1}^M \lambda_j \right) > E, \quad (9)$$

式(9)表示前  $p$  个  $\lambda_i$  的总和比大于  $E$ . 当特征空间变量均值不满足式(1)时, 式(7)中的  $K$  取

$$K = K - L_{M \times M} K - K L_{M \times M} + L_{M \times M} K L_{M \times M}, \quad (10)$$

其中:  $L_{M \times M}$  为  $M \times M$  矩阵, 矩阵元素为  $1/M$ .

将样本数据进行特征转换后, 基于新特征变量建立多T-S结构的估计模型, 设第  $l$  条规则形式为

$$R_l : \text{if } x' \text{ is } A_l, \text{ then } y = f_l(x'), \quad l = 1, 2, \dots, c,$$

其中:  $x'$  为特征变量,  $A_l$  为前件模糊集,  $f_l(x')$  为第  $l$  条规则的后件子模型,  $c$  为规则数. 采用FCM方法对转换后的数据  $X' = [x'_1 \ x'_2 \ \dots \ x'_M]'$  进行聚类获得模糊集  $\{A_l\}$ . 聚类准则是求得适当的模糊划分矩阵与模糊聚类中心, 使得目标函数达到极小值. 聚类目标函数为

$$J(U, C) = \sum_{i=1}^c \sum_{j=1}^M u_{ij}^m d_{ij}^2, \quad (11)$$

其中:  $U = [\mu_{ij}]$  为隶属度矩阵,  $C = \{c_i\}$  为聚类中心,  $c$  为聚类数,  $m \in [1, \infty)$  为加权指数, 根据经验本文取  $m = 2$ . 通过迭代算法可以近似得到式(11)的最优值,  $u_{ij}$  与  $c_i$  的迭代公式分别为式(12)与式(13),  $d_{ij}^2$  的计算为式(14), 反复迭代直到满足  $|U^{(t)} - U^{(t-1)}| \leq \varepsilon$  或达到最大迭代次数停止迭代, 其中  $t$  为迭代次数:

$$u_{ij} = ((d_{ij})^{2/(m-1)} / \sum_{k=1}^c (d_{kj})^{2/(m-1)}), \quad (12)$$

$$c_i = \sum_{j=1}^M u_{ij}^m x'_j / \sum_{j=1}^M u_{ij}^m, \quad (13)$$

$$d_{ij}^2 = \sum_{k=1}^l \|x'_{ik} - c_{jk}\|^2. \quad (14)$$

对样本进行分类后, 然后建立T-S的后件子模型. 假设第  $l$  子类训练样本集  $\{x'_i, y_i\}_{i=1}^N$ ,  $x'_i \in \mathbb{R}^p$ , 其中:  $p$  为样本空间维数,  $N$  为样本总数,  $y_i \in \mathbb{R}$ . 在特征空间中, 最优超平面需要求解如下优化问题:

$$\begin{cases} \min J(\mathbf{w}, b, \varepsilon_i) = \frac{1}{2} \mathbf{w}^T \mathbf{w} + \frac{\varsigma}{2} \sum_{i=1}^N \varepsilon_i^2, \\ \text{s.t.} : y_i = \mathbf{w}^T \Phi(x'_i) + b + \varepsilon_i, \end{cases} \quad (15)$$

其中:  $\varsigma$  为惩罚系数,  $\varepsilon_i$  为松弛变量,  $\Phi(\cdot)$  为非线性映射函数. 由于数据存在不确定性, 导致最优超平面也存在不确定性, 即式(15)中  $\|\mathbf{w}\|$  取值会在一定范围内随机分布. 因此, 将式(15)转换为求解最优超平面的期望

$$\begin{cases} \min J(\mathbf{w}, b, \varepsilon_i) = \\ \frac{1}{2} \int_{\|\mathbf{w}\|} \|\mathbf{w}\|^2 \rho(\|\mathbf{w}\|) d(\|\mathbf{w}\|) + \frac{\varsigma}{2} \sum_{i=1}^N \varepsilon_i^2, \\ \text{s.t.} : y_i = \mathbf{w}^T \Phi(x'_i) + b + \varepsilon_i, \end{cases} \quad (16)$$

其中  $\rho(\|\mathbf{w}\|)$  为  $\|\mathbf{w}\|$  的概率密度分布函数. 最后第  $l$  条规则的后件子模型为

$$f_l(x') = \sum_{i=1}^N \alpha_i k(x'_i, x') + b, \quad (17)$$

其中:  $k(x'_i, x') = \Phi(x'_i) \cdot \Phi(x')$  为核函数,  $\alpha_i$  为 Lagrange 乘子. 在求解式(16)的过程中, 必须先对样本进行随机抽样后再对  $\rho(\|\mathbf{w}\|)$  进行估计.

#### 4 入矿品位估计 (Mineral concentration estimation of feed ore)

图2所示为本文提出的入矿品位在线估计策略, 其中  $I, P, S$  分别代表入矿品位、加药量、图像特征等变量,  $p$  代表可信用度. 首先利用泡沫图像特征与加药量直接估计出入矿品位, 同时利用泡沫图像特征估计出精

矿品位, 最后通过对比入矿品位与精矿品位估计的结果对入矿品位估计结果的可信用度进行评价.

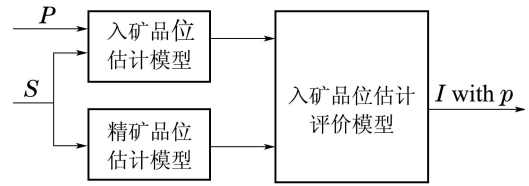


图2 浮选入矿品位估计方法

Fig. 2 Mineral concentration estimation of feed ore

建立入矿品位与精矿品位估计模型. 将第  $k$  个稳定工况状态下特定时段内样本数据记为  $x^k = [x_1^{kT} \ \dots \ x_l^{kT}]^T$  ( $l=5$ ), 将原始样本总集记为  $X = [x^1 \ \dots \ x^k]^T$ , 样本均值集记为  $\bar{X} = \{\bar{x}^k\}$   $\bar{x}^k = \{\bar{x}_j^k\}$ , 其中  $\bar{x}_j^k$  表示  $x^k$  内第  $j$  维的均值. 建立基于FCM-PSVR的估计模型, 其步骤为:

**Step 1** 将  $\bar{X}$  的输入部分经过KPCA处理后与  $\bar{X}$  的输出结合作为聚类样本, 将之记为  $\bar{X}'$ , 其中KPCA中所用核函数为高斯核函数;

**Step 2** 对  $\bar{X}'$  的输入部分数据进行FCM聚类, 将聚类后的第  $i$  个子类的均值样本集记为  $\bar{X}'_i$ , 对应的原始样本集记为  $X_i$ ;

**Step 3** 对  $X_i$  进行随机抽样, 即对所有  $x^k \in X_i$  进行随机删除某一行并重新计算均值集, 得到第  $j$  次随机抽样后的均值集, 将之记为  $\bar{X}_{i,j}$ ;

**Step 4** 将  $\bar{X}_{i,j}$  的输入部分映射到核主元空间后与  $\bar{X}_{i,j}$  的输出结合得到训练样本, 将之记为  $\bar{X}'_{i,j}$ ;

**Step 5** 将  $\bar{X}'_{i,j}$  作为SVR的训练样本, 训练得到  $\|\mathbf{w}\|_{i,j}$ , 并采用密度估计法得到第  $i$  个子类的概率密度分布函数  $\rho_i(\|\mathbf{w}\|)$ , 其中支持向量回归(support vector regress, SVR)中所用核函数为高斯核函数;

**Step 6** 将所有的  $\bar{X}_{i,j}$  与  $\bar{X}'$  进行合并作为式(16)的训练集, 采用遗传优化方法求解式(16), 并得到最终的回归模型式(17).

加药量不变时, 将入矿品位与精矿品位的变化趋势记为  $v_i$  ( $i = 1, 2$ ), 其中  $v_i = -1, 0, 1$  分别表示变化趋势为: 降、平稳、升. 若两者变化趋势一致则表示估计结果可信用度高, 不一致则可信用度低, 因此令  $\Delta v = \|v_1 - v_2\|$ , 则可采用表1中的规则对入矿品位估计结果进行评价.

表1 可信用度评价规则

Table 1 Evaluation rules for the reliability

可信用度	高	中	低
$\Delta v$	0	1	2

#### 5 工业验证 (Industrial verification)

基于机器视觉的浮选泡沫分析系统在2012年第2季度进行了安装调试与功能测试, 之后投入生产实践.

系统采用本文提出的方法对入矿品位进行在线估计. 从历史数据中得到2575个金粗选样本和2740个铈粗选样本, 以及515个金粗选均值样本和548个铈粗选均值样本.

入矿金品位的估计采用泡沫尺寸、负载率、蓝色分量均值与色度等4个图像特征变量外加5个加药量变量, 入矿铈品位的估计采用泡沫尺寸、负载率、蓝色分量均值与色度等4个图像特征变量外加5个加药量变量. 分别采用365个金粗选均值样本和388个铈粗选均值样本作为KPCA的训练样本, 取累积贡献率90%为门限值, 可得金有6个核主元而铈有5个核主元. 对精矿品位进行估计时仅利用4个图像特征, 经过KPCA处理后, 得到金有3个核主元而铈有2个核主元. 然后采用FCM-PSVR方法建立入矿和精矿金、铈品位估计模型. 图3与图4分别为入矿品位估计与精矿品位估计结果. 采用误差均值与标准差作为衡量标准与文献[8]中的方法进行比较, 所得结果如表2所示. 从表2可知本文方法所得结果误差更小且误差分布更均匀.

为能更鲜明体现基于专家规则入矿估计评价模型的作用, 可先忽略加药量的影响. 对图3与图4中的入矿品位估计与精矿品位估计结果归一化, 并采用基于表1的专家规则对其进行评价得到图5. 从图5中可知绝大多数入矿品位估计结果的可信度为高和中, 而在 $\Delta v = 2$ 的时段, 可信度较低的主要原因可能为: 1) 精矿品位还受加药量操作模式的影响; 2) 数据不确定性对建模准确性的影响. 通过可信度可有效提示工人注意当前工况的变化.

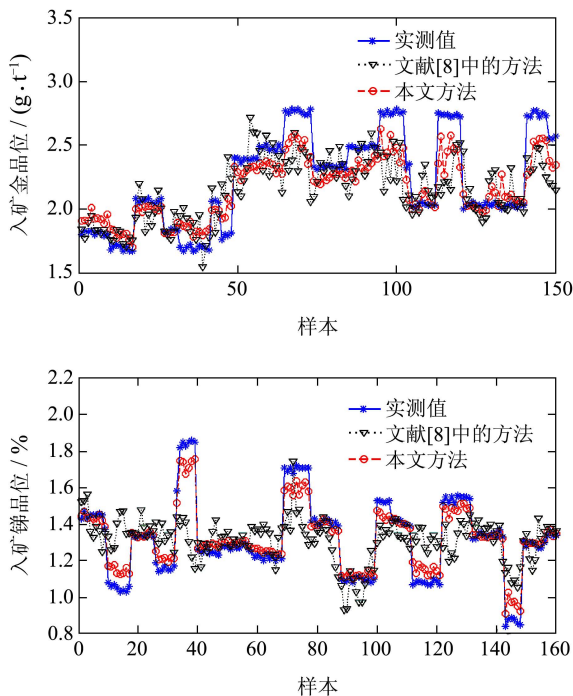


图 3 入矿品位估计结果

Fig. 3 Mineral concentration estimation of feed ore

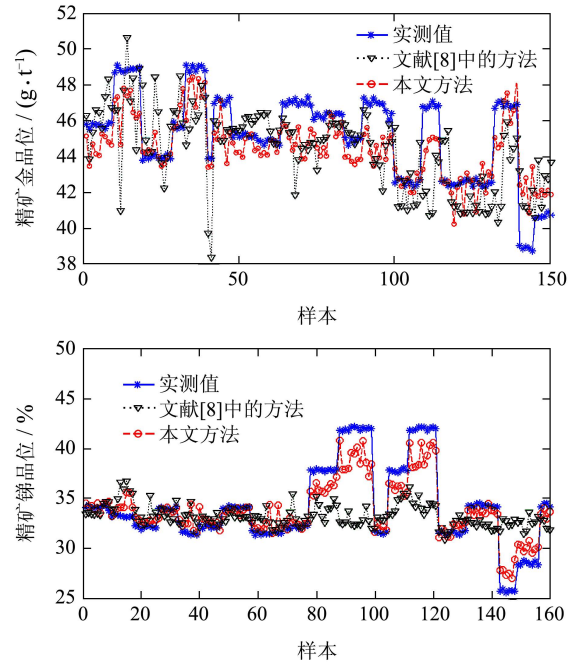


图 4 精矿品位估计结果

Fig. 4 Mineral concentration estimation of concentrate

表 2 估计结果的效果对比

Table 2 Comparison of performance of estimation

品位估计模型	本文方法		文献[8]中的方法	
	误差均值	误差标准差	误差均值	误差标准差
入矿金/(g · t <sup>-1</sup> )	0.196	0.139	0.248	0.159
入矿铈/%	0.15	0.085	0.197	0.114
精矿金/(g · t <sup>-1</sup> )	1.883	1.274	2.291	1.454
精矿铈/%	2.808	2.239	3.844	3.043

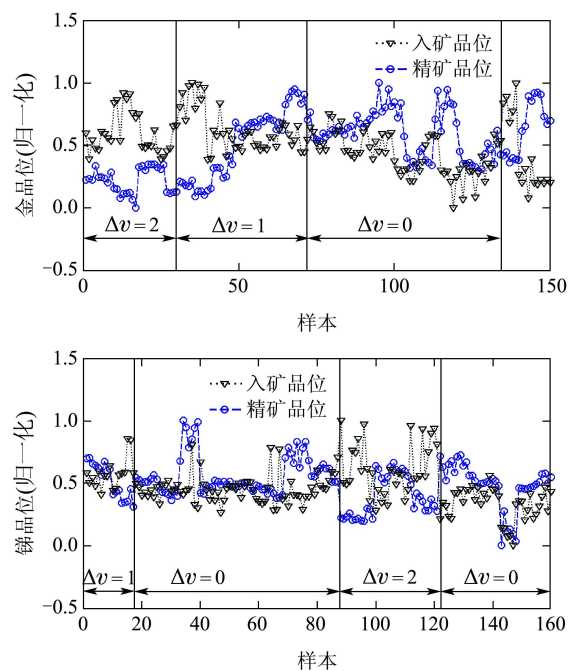


图 5 入矿品位估计评价

Fig. 5 Evaluation of mineral concentration estimation of feed ore

## 6 结语(Conclusion)

通过挖掘图像特征与入矿品位、精矿品位之间的知识和入矿品位与精矿品位之间的知识,提出一种在线估计金梯浮选入矿品位的方法.并针对样本数据中存在的 uncertainty 提出一种基于KPCA与FCM-PSVR的建模方法.然后分别建立基于泡沫图像特征的入矿品位与精矿品位估计模型.最后通过基于专家规则的方法对入矿品位的估计结果进行评价.实践表明,该方法能有效对入矿品位进行在线估计,为后续加药量的调整提供有效指导.

## 参考文献(References):

- [1] 桂卫华, 阳春华, 徐德刚, 等. 基于机器视觉的矿物浮选过程监控技术研究进展 [J]. 自动化学报, 2013, 39(11): 1879 – 1888.  
(GUI Weihua, YANG Chunhua, XU Degang, et al. Machine-vision-based online measuring and controlling technologies for mineral flotation a review [J]. *Acta Automatica Sinica*, 2013, 39(11): 1879 – 1888.)
- [2] HE M, YANG C, WANG X, et al. Nonparametric density estimation of froth colour texture distribution for monitoring sulphur flotation process [J]. *Minerals Engineering*, 2013, 53: 203 – 212.
- [3] SUPOMO A, YAPE, ZHENG X, et al. PT Freeport Indonesia's mass-pull control strategy for rougher flotation [J]. *Minerals Engineering*, 2008, 21: 808 – 816.
- [4] LIU J J, MACGREGOR J F. Froth-based modeling and control of flotation processes [J]. *Minerals Engineering*, 2008, 21: 642 – 651.
- [5] 唐朝晖, 刘金平, 桂卫华, 等. 基于数字图像处理的浮选泡沫速度特征提取及分析 [J]. 中南大学学报(自然科学), 2009, 40(6): 1616 – 1622.  
(TANG Zhaohui, LIU Jinping, GUI Weihua, et al. Froth bubbles speed characteristic extraction and analysis based on digital image processing [J]. *Journal of Central South University (Science and Technology)*, 2009, 40(6): 1616 – 1622.)
- [6] 阳春华, 周开军, 牟学民, 等. 基于计算机视觉的浮选泡沫颜色及尺寸测量方法 [J]. 仪器仪表学报, 2009, 30(4): 717 – 721.  
(YANG Chunhua, ZHOU Kaijun, MOU Xuemin, et al. Froth color and size measurement method for flotation based on computer vision [J]. *Chinese Journal of Scientific Instrument*, 2009, 30(4): 717 – 721.)
- [7] 吴佳, 谢永芳, 阳春华, 等. 一种无监督约简的浮选泡沫图像特征选择方法及应用 [J]. 信息与控制, 2014, 43(3): 314 – 317.  
(WU Jia, XIE Yongfang, YANG Chunhua, et al. An unsupervised reduction method for selection of flotation froth image characters and its application [J]. *Information and Control*, 2014, 43(3): 314 – 317.)
- [8] GAO S Z, WANG J S, GAO X W. Modeling and advanced control method of PVC polymerization process [J]. *Journal of Process Control*, 2013, 23(5): 664 – 681.
- [9] 黄景春, 肖建. 基于小波神经网络的云模型 [J]. 控制理论与应用, 2011, 28(1): 53 – 57.  
(HUANG Jingchun, XIAO Jian. Cloud model based on wavelet neural networks [J]. *Control Theory & Applications*, 2011, 28(1): 53 – 57.)
- [10] 潘永平, 黄道平, 孙宗海. II型模糊控制综述 [J]. 控制理论与应用, 2011, 28(1): 13 – 23.  
(PAN Yongping, HUANG Daoping, SUN Zonghai. Overview of type-2 fuzzy logic control [J]. *Control Theory & Applications*, 2011, 28(1): 13 – 23.)
- [11] LI H X, YANG J L, ZHANG G, et al. Probabilistic support vector machines for classification of noise affected data [J]. *Information Sciences*, 2013, 221: 60 – 71.
- [12] ABIYEV R H, KAYNAK O. Type 2 Fuzzy neural structure for identification and control of time-varying plants [J]. *IEEE Transactions on Industrial Electronics*, 2010, 57(12): 4147 – 4159.
- [13] LEE J M, YOO C K, CHOI S W, et al. Nonlinear process monitoring using kernel principal component analysis [J]. *Chemical Engineering Science*, 2004, 59(1): 223 – 234.

## 作者简介:

吴佳 (1982–), 男, 博士研究生, 目前研究方向为复杂有色冶金过程建模、数据挖掘, E-mail: star0723@163.com;

谢永芳 (1972–), 男, 教授, 博士生导师, 目前研究方向为分散鲁棒控制、故障诊断等, E-mail: yfxie@mail.csu.edu.cn;

阳春华 (1965–), 女, 教授, 博士生导师, 目前研究方向为复杂工业过程建模及优化控制, E-mail: ychh@csu.edu.cn;

桂卫华 (1950–), 男, 教授, 博士生导师, 中国工程院院士, 目前研究方向为复杂工业过程建模及优化控制, E-mail: gwh@mail.csu.edu.cn.