

基于向量空间模型的浮选泡沫图像分类方法研究^①

王雅琳^② 张润钦 谢永芳 桂卫华

(中南大学信息科学与工程学院 长沙 410083)

摘 要 针对基于图像底层特征的泡沫图像分类识别正确率不高、存在语义鸿沟问题,提出一种基于向量空间模型(VSM)的浮选泡沫图像分类方法。该方法借鉴文本分类方法,对工业摄像机获取的大量泡沫图像通过分块、底层特征提取和聚类,构造泡沫状态词汇表;在此基础上,经词汇相似度和词频计算,用词袋向量描述泡沫图像;最后,采用 VSM 实现实时泡沫图像的有监督分类识别。用某金属浮选过程工业现场泡沫图像数据对该方法进行了实验验证,实验结果表明,该方法的工况识别平均准确率近 90%,明显优于基于底层特征的分类方法,并在一定程度上解决了语义鸿沟问题,具有很好的应用价值。

关键字 泡沫图像,分类识别,向量空间模型(VSM),泡沫状态词汇,纹理

0 引 言

浮选是一种广泛应用的矿物分选技术。实际浮选过程一般由有经验的工人观察浮选泡沫状态来进行控制。随着机器视觉和人工智能技术的发展,利用数字图像处理技术^[1]进行浮选泡沫图像分类、浮选工况识别及控制优化^[2]成为浮选过程控制的发展趋势。国内外学者对基于机器视觉的浮选泡沫图像分类以及工况识别开展了大量的研究工作^[3-6]。目前,这方面的研究主要集中在基于图像底层特征的分类方法研究,即采用提取浮选泡沫图像的纹理、颜色、泡沫尺寸分布等底层特征参数对泡沫图像进行描述,然后利用神经网络、支持向量机等方法进行分类和识别。然而,将图像作为一个整体进行处理,用图像底层特征加以描述的泡沫图像不便于人们理解,其分类结果存在语义鸿沟问题,即底层特征和高层语义特征之间的不统一性。另外,由于工业现场存在光照、粉尘等噪声干扰,两类完全不相似的图像提取到的全局底层特征描述可能很接近,这就导致基于泡沫图像底层特征的神经网络、支持向量机等分类识别准确度降低。

向量空间模型^[7,8](vector space model, VSM)是一个表示文本文件的代数模型,它把对文本内容的

处理简化为向量空间中的向量运算。图像和文本有很大不同,图像是信息的最原始直观描述,而文本是图像的语义抽象。若能从语义层面去理解图像,再在语义层面上对泡沫图像进行分类识别,则分类结果更接近人的理解。为此,本文提出了一种基于向量空间模型(VSM)的泡沫图像分类方法,该方法充分利用图像的局部信息,对泡沫图像分块处理后,进行泡沫状态词汇和词袋两次抽象处理,得到泡沫图像语义层面的描述,再基于向量空间模型方法,像处理文本一样对泡沫图像进行分类识别。经验证,该方法不仅提高了分类的准确率,而且也解决了语义鸿沟问题,为泡沫图像分类和工况识别提供了一种新思路。

1 向量空间模型

向量空间模型(VSM)由 Salton 等人于 20 世纪 70 年代提出,在文本分类、检索等领域得到了广泛的应用^[9,10]。VSM 把对文档内容的处理简化为向量空间中的向量运算,通过计算向量之间的相似性来度量文档间的相似性。文档被视为索引词(关键字)形成的多元向量空间,索引词的集合通常为文档中至少出现过一次的词组。检索时,检索词被转换成类似于文档的向量,比较每个文档(向量)和检

① 国家自然科学基金(61273187,61134006),国家科技支撑计划(2012BAF03B05)和湖南省科技计划(2012CK4018)资助项目。

② 女,1973 年生,博士生,教授;研究方向:复杂过程建模、优化与控制;联系人,E-mail:ylwang@mail.csu.edu.cn
(收稿日期:2012-12-05)

索词(向量)的相似程度就可得知文档和搜寻词的相关程度,从而达到分类与检索的目的。

利用 VSM 对文档进行相似度匹配,首先要将每个文档用一个向量表示。其做法如下:对于文档集中每一个不同的词项(或概念),记录为向量中的一个分量;对于某一特定文档而言,当词项出现一次时,就将对应向量的分量加 1,若不出现,则对应向量的分量为 0。得到文档的向量后,就可以通过内积法、夹角余弦法等方法度量向量的相似度。

2 基于 VSM 的泡沫图像分类方法

基于 VSM 的泡沫图像分类方法利用数字图像处理技术提取泡沫图像的纹理特征量和颜色特征量来刻画图像,然后用词袋的方法对泡沫图像进行描述,最后利用 VSM 对泡沫图像进行分类。

2.1 泡沫图像的底层特征选择和提取

浮选过程不同工况的泡沫图像在纹理粗细度、纹沟的深浅等方面有很大差别。此外,浮选泡沫的颜色在很大程度上也能反映泡沫所携带的矿物类型和矿物含量。因此,选择纹理和颜色特征对泡沫图像进行描述。

灰度共生矩阵^[11](GLCM)是由 Haralick 提出的一种基于统计的纹理特征提取算法,对纹理图像的描述精确度高。GLCM 是由图像灰度级之间的联合概率密度 $P(i,j|d,\theta)$ 所构成的矩阵,从统计的角度反映了图像中任意两点间灰度的空间相关性。其中, $P(i,j|d,\theta)$ 表示方向为 $\mu_y = \sum_j j \sum_i p(i,j|d,\theta)$ 、间距为 d 的灰度共生矩阵 P 中第 i 行 j 列元素的值, θ 取 $0^\circ, 45^\circ, 90^\circ$ 和 135° 共 4 个方向。设 $I(x,y)$ 表示一幅像素大小为 $U \times V$ 的二维数字图像中坐标分别为 x,y 所对应像素点的灰度值。则对于不同的 θ , $P(i,j|d,\theta)$ 计算方法如下:

$$P(i,j|d,0^\circ) = \text{Num}\{(x_1,y_1)(x_2,y_2) \in U \times V | x_1 - x_2 = 0, |y_1 - y_2| = d; I(x_1,y_1) = i, I(x_2,y_2) = j\}$$

$$P(i,j|d,45^\circ) = \text{Num}\{(x_1,y_1)(x_2,y_2) \in U \times V | (x_1 - x_2 = d, y_1 - y_2 = d) \text{ 或 } (x_1 - x_2 = -d, y_1 - y_2 = -d); I(x_1,y_1) = i, I(x_2,y_2) = j\}$$

$$P(i,j|d,90^\circ) = \text{Num}\{(x_1,y_1)(x_2,y_2) \in U \times V | |x_1 - x_2| = d, y_1 - y_2 = 0; I(x_1,y_1) = i, I(x_2,y_2) = j\}$$

$$P(i,j|d,135^\circ) = \text{Num}\{(x_1,y_1)(x_2,y_2) \in U \times V | (x_1 - x_2 = d, y_1 - y_2 = -d) \text{ 或 } (x_1 - x_2 = -d, y_1 - y_2 = d); I(x_1,y_1) = i, I(x_2,y_2) = j\}$$

其中, $\text{Num}\{X\}$ 表示集合 X 中的元素个数。

选取纹理常用的 5 个特征值(ASM、ENT、CON、IDM、COR)和泡沫图像的相对红色分量 R_{relative} 来描述泡沫图像。具体描述如下:

(1) ASM(角二阶矩)

$$ASM = \sum_{i,j} \{P(i,j|d,\theta)\}^2 \quad (1)$$

描述图像的灰度均匀性,纹理粗 ASM 较大,反之则较小。

(2) ENT(熵)

$$ENT = - \sum_{i,j} \{p(i,j|d,\theta)\} \log \{p(i,j|d,\theta)\} \quad (2)$$

描述图像所具有的信息量,纹理越多熵值越大。

(3) CON(对比度)

$$CON = \sum (i-j)^2 P(i,j|d,\theta) \quad (3)$$

描述图像的清晰度和纹理沟纹深浅程度,沟纹越深,对比度越大,图像也越清晰。

(4) IDM(逆差矩)

$$IDM = \sum_{i,j} P(i,j|d,\theta) / [1 + (i-j)^2] \quad (4)$$

描述图像纹理的同质性,度量图像纹理局部变化的多少。共生矩阵沿对角线集中,则逆差矩值越大。

(5) COR(相关性)

$$COR = \sum_{i,j} (i - \mu_x)(j - \mu_y) P(i,j|d,\theta) / \sigma_x \sigma_y \quad (5)$$

$$\text{其中, } \mu_x = \sum_i i \sum_j p(i,j|d,\theta)$$

$$\mu_y = \sum_j j \sum_i p(i,j|d,\theta)$$

$$\sigma_x = \sum_i (i - \mu_x)^2 \sum_j p(i,j|d,\theta)$$

$$\sigma_y = \sum_j (j - \mu_y)^2 \sum_i p(i,j|d,\theta)$$

描述灰度共生矩阵中的元素在行或列的方向上的相似程度,反映某种颜色沿某个方向的延伸长度,延伸的越长, COR 越大。

(6) R_{relative} (相对红色分量)

$$R_{\text{relative}} = \frac{R_{\text{red}}}{R_{\text{gray}}} \quad (6)$$

式中, R_{red} 和 R_{gray} 分别代表红色分量均值和灰度均

值。

2.2 泡沫图像的泡沫状态词汇表生成

文档分类是通过匹配文档中所含关键词的相关性来表征文档相关性的。利用 VSM 对图像进行分类,首先要从图像中构建出与文档中的词汇相对应的能够表征泡沫工况信息的泡沫状态词汇。

在文档中,每个关键词出现的频率能够表征文档的主题。同理,在图像分类中,每个泡沫状态词汇出现的频率也能够表征图像的主题,即图像所属的类别。

针对浮选泡沫图像,利用 K-means 聚类方法获取泡沫状态词汇表,具体步骤如下:

步骤 1:从图像库中选取 N 幅图像(这 N 幅图像要尽可能广的覆盖所有图像类别),每幅图像都截取到某一相同的像素大小 $L_x \times L_y$ 。

步骤 2:对于截取后的 N 幅图像,每幅图像均匀划分成 $m \times m$ 块。

步骤 3:对于每一个分块,按式(1) - (6)获取其纹理特征值和相对红色分量值,构成一个 1×6 维的底层特征向量描述。

步骤 4:对所有的底层特征向量描述进行 K-means 聚类,得到的 D 个聚类中心即为泡沫状态词汇,由此组成泡沫状态词汇表。

2.3 泡沫图像的词袋描述

得到泡沫图像的泡沫状态词汇表后,就可以用

词袋的方法对每一幅泡沫图像进行描述,得到图像的一个向量表示,即词袋描述。

获取泡沫图像的词袋描述示意图如图 1 所示。泡沫图像词袋描述的具体步骤如下:

步骤 1:将图像像素截取为 $L_x \times L_y$ 的像素大小。

步骤 2:将图像均匀划分成 $m \times m$ 块。

步骤 3:对于每一个分块,获取其纹理特征和相对红色分量值,构成一个 1×6 维底层特征向量描述。

步骤 4:计算底层特征向量与泡沫状态词汇表中各个状态词汇的欧式距离,度量每个分块底层特征向量与泡沫状态词汇表中所有词汇的相似性,该分块与哪个泡沫状态词汇最相似,就将其标定为哪个词汇。

步骤 5:统计每个泡沫状态词汇出现频次,得到图像的词袋向量表示。

图 1 中得到的泡沫图像词袋向量表示为 $\langle 0, 0, 3, 3, 3, 1, 6, 0 \rangle$,其语义层的含义为:6 表示词汇 7 出现了 6 次,即图像中含有大泡沫的面积较多,其次就是含有较多的细小泡沫(对应值为 3)。值为 0 的元素,表示图像中没有出现对应状态词汇,也即图像中不含虚泡沫和光滑的纹理较浅的区域等,这种描述与人的理解相似,在一定程度上解决了语义鸿沟问题。

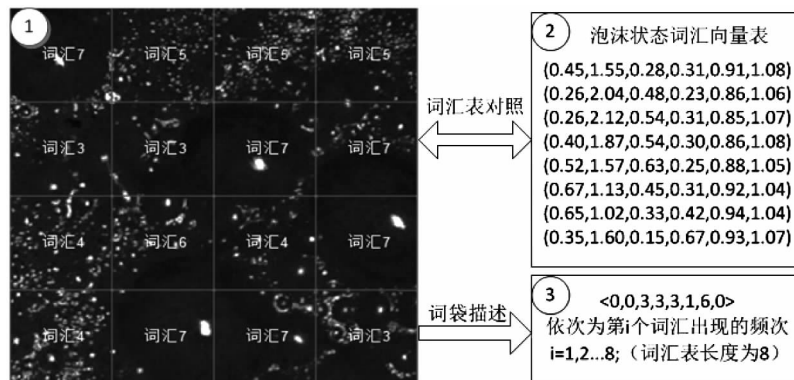


图 1 泡沫图像的词袋描述

2.4 基于 VSM 的泡沫图像分类

将泡沫图像用词袋方法描述成一个向量后,就可以用 VSM 的方法对其进行有监督分类。基于 VSM 的分类方法有向量内积法、贝叶斯方法、K 近邻方法、支持向量机、神经网络方法等。本文仅描述基于向量内积法的分类方法。

为了分类效果稳定,本文并没有用每个类别的中心向量来表征该类图像,而是通过计算待分类图像与训练集内每个类别中所有图像的相似度之和来进行分类。分类思路为,对于训练集内每个类别中的图像,用词袋的方法加以描述,得到该类的词袋向量集合 $C_i, i = 1, 2, \dots, N_s$ (N_s 为训练集中分类图像的

类别数)。当获取到新的实时图像,将该图像用词袋的方法加以描述,得到图像的向量表示,通过计算该词袋向量与各类图像的词袋向量集合 C_i 的相似度来进行分类。具体过程如下:

步骤 1:将训练图库中各个类别的图像集合用词袋模型加以描述得到该类图像的词袋向量集合,每一类别取 M 幅图像,则第 i 类图像的词袋向量集合为

$$C_i = [c_{i1} c_{i2} \cdots c_{iM}]^T, i = 1, 2, \cdots, N_s \quad (7)$$

式中, c_{ij} 表示第 i 类第 j 个图像的词袋描述,为 $1 \times D$ 的行向量。

步骤 2:对于待分类图像,得到其词袋描述 $Q, Q = [q_1 q_2 \cdots q_D]$ 为 $1 \times D$ 的行向量。

步骤 3:将 Q 与 C_i 类中的每个向量 c_{ij} 进行相似度计算(这里用向量的内积来度量其相似度),得到 Q 与每个类别集合 C_i 中所有向量的相似度之和 SC_i :

$$SC_i = \sum_{j=1}^M Sim(c_{ij}, Q), i = 1, 2, \cdots, N_s \quad (8)$$

式中, $Sim(W, Q) = W \cdot Q = \sum_{i=1}^D w_i \times q_i$, $W = [w_1 w_2 \cdots w_D]$ 为 $1 \times D$ 的行向量。

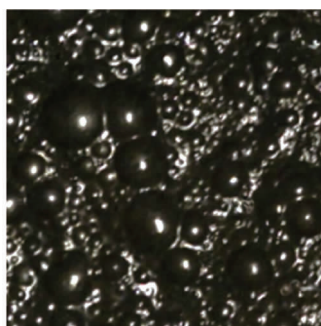
步骤 4:将 Q 归为相似度最大的第 k 类, $k = \arg\max_i SC_i, i = 1, 2, \cdots, N_s$ 。

3 工业数据验证

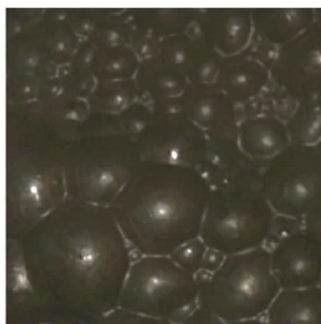
收集某金属浮选过程工业现场获取的 400 幅泡沫图像,用其中的 200 幅作为训练样本,另外 200 幅作为测试样本。根据工业现场的生产经验,将浮选工况分为 4 类(图 2 所示为 4 类典型工况下获得的泡沫图像)。通过仿真实验,对提出的基于 VSM 的泡沫图像分类方法与现有基于泡沫底层特征描述的 BP 神经网络分类方法的分类效果进行比较分析,验证所提方法的有效性。



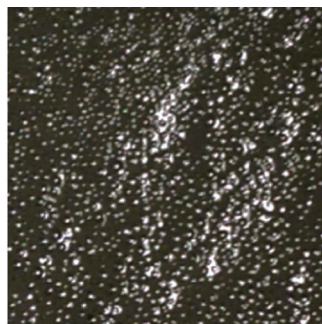
第一类



第二类



第三类



第四类

图 2 四类典型工况下的泡沫图像

泡沫图像分类识别正确率是以人工分类结果为参照。对于每个测试样本图像,先人工标定其类别,然后用本文的分类方法进行分类,如果算法的分类结果与人工标定结果一致,则视为分类正确。分类

准确率公式为:

$$Rate_c = \frac{C_{right}}{C_{sum}} \quad (9)$$

其中: $Rate_c$ 为第 C 类测试样本的分类准确率。 C_{right}

为算法将人工标定为第 C 类的图像正确归类为 C 类的数目, C_{sum} 为测试样本中人工标定为 C 类的图像总数目。

3.1 BP 神经网络分类结果

将泡沫图像截取为 960×960 像素大小, 得到每幅图像的纹理和颜色特征向量表示。取这 6 个底层

特征作为 BP 神经网络的输入节点, 令 4 个类别作为神经网络的输出节点, 并分别取中间隐层节点数为 5—8, 进行分类的训练和预测。预测的分类结果如表 1 所示(表中 T 代表测试样本总个数, Y 代表被正确分类的测试样本个数, R 代表准确率)。

表 1 BP 神经网络分类结果

中间隐层 节点数目	第一类			第二类			第三类			第四类		
	T	Y	R	T	Y	R	T	Y	R	T	Y	R
5	50	39	78%	50	41	82%	50	42	84%	50	40	80%
6	50	39	78%	50	43	86%	50	41	82%	50	41	82%
7	50	40	80%	50	41	82%	50	42	84%	50	41	82%
8	50	39	78%	50	40	80%	50	42	84%	50	40	80%

3.2 基于 VSM 的泡沫图像分类结果

基于 VSM 的泡沫图像分类方法中, 泡沫状态词汇表的长度 D 以及待分类图像的分块数目 $m \times m$, 对分类的准确性都有影响。

用本文提出的方法选取 D 个泡沫状态词汇后, 将泡沫图像截取为 960×960 像素大小, 经处理后, 用词袋方法对图像进行描述, 并分别用向量内积法进行分类。表 2 给出了不同类别图像词袋描述数据, 其中每类图像随机从样本集中选取了 3 幅图像, 词汇表长度取为 8, 相对应的 8 个泡沫状态词汇见

图 1 中的②; 词频 i 指第 i 个词汇出现的次数。表 3 给出了采用不同分块 ($m \times m = 8 \times 8, 10 \times 10, 12 \times 12$) 和不同长度 ($D = 8, 12, 16$) 的泡沫状态词汇表情情况下的分类预测结果(分类准确率以人工分类结果为参照, 表中 T 代表测试样本总个数, Y 代表被正确分类的测试样本个数, R 代表准确率)。由表 2 可以看出, 第 1 类图像出现词汇 4 和 5 的频率较高, 第 2 类主要集中在词汇 1 到词汇 3, 第 4 类集中在词汇 7 和词汇 8。

表 2 泡沫图像词袋描述

图像类别	图像 序号	词 袋 描 述							
		词频 1	词频 2	词频 3	词频 4	词频 5	词频 6	词频 7	词频 8
1	1	0	1	3	38	20	0	2	0
	2	0	0	12	18	34	0	0	0
	3	0	0	0	10	51	2	1	0
2	1	16	0	48	0	0	0	0	0
	2	18	40	6	0	0	0	0	0
	3	14	5	41	4	0	0	0	0
3	1	6	0	3	0	47	8	0	0
	2	0	0	0	0	0	55	9	0
	3	0	0	5	0	8	50	1	0
4	1	0	0	0	0	0	0	16	48
	2	0	0	0	0	0	0	42	22
	3	0	9	0	0	0	0	0	55

表 3 向量内积法的分类结果

状态词汇 表长度	测试图像 分块数目	第一类			第二类			第三类			第四类		
		<i>T</i>	<i>Y</i>	<i>R</i>	<i>T</i>	<i>Y</i>	<i>R</i>	<i>T</i>	<i>Y</i>	<i>R</i>	<i>T</i>	<i>Y</i>	<i>R</i>
8	8 × 8	50	43	86%	50	44	88%	50	45	90%	50	41	82%
8	10 × 10	50	43	86%	50	46	92%	50	45	90%	50	43	86%
8	12 × 12	50	41	82%	50	46	92%	50	46	92%	50	45	90%
12	8 × 8	50	42	84%	50	47	94%	50	46	92%	50	42	84%
12	10 × 10	50	43	86%	50	47	94%	50	47	94%	50	42	84%
12	12 × 12	50	41	82%	50	45	90%	50	47	94%	50	45	90%
16	8 × 8	50	42	84%	50	46	92%	50	45	90%	50	43	86%
16	10 × 10	50	42	84%	50	46	92%	50	46	92%	50	45	90%
16	12 × 12	50	42	84%	50	46	92%	50	46	92%	50	45	90%

3.3 结果分析

根据表 3 的分类结果可知,基于 VSM 分类的平均准确性大于 85% (接近 90%),比单纯基于底层特征的 BP 神经网络分类准确性高(仅 80% 左右),而且该方法相对于 BP 神经网络分类更加简便。同时可以看到,图像的分块数目和泡沫状态词汇表的规模对分类结果也有影响。状态词汇越多分类效果越好,但图像的分块并不是越多越好。若分块数目过多(即每个分块过小),由于图像存在噪声污染(亮点区域,暗点区域等)得到的纹理特征和颜色特征并不能表征其实际值,导致泡沫状态词汇匹配错误。泡沫状态词汇表越大,计算量也越大。因此,需要综合考虑计算量和准确性,选取适宜长度的泡沫状态词汇表,并进行合理的图像分块。

4 结 论

向量空间模型(VSM)在文本分类中得到了广泛应用,本文结合泡沫图像实际情况,提出了基于 VSM 的泡沫图像分类方法。该分类方法首先用聚类方法得到泡沫状态词汇表,然后用词袋的方法对泡沫图像进行描述,得到泡沫图像的词袋向量表示,从而保证描述的图像更便于人们理解;最后用向量内积分类方法进行图像的有监督分类。基于工业现场获取的浮选泡沫图像进行了实验验证,比较了所提出分类方法和基于图像底层特征的 BP 神经网络分类方法的分类效果,验证了所提方法的有效性。实验结果表明,所提的基于 VSM 的分类方法其分类结果稳定可靠,平均精确率接近 90%。

从实验结果可以看到,图像分块大小以及词汇

表的规模对最终的分类识别有影响,文中对图像的分块只是进行均匀划分且块数相同,词汇表规模的选取也是选定在一定范围内。但根据直观理解,若能够针对采集到的不同图像进行不定块数、不定大小的智能划分,则能够更好地区分各类图像,提高分类识别的准确率。因此泡沫图像的智能分块方法以及词汇表规模的选取方法等是下一步研究的方向。

参考文献

[1] Aldrich C, Marais C, Shean B J, et al. Online monitoring and control of froth flotation systems with machine vision: A review. *International Journal of Mineral processing*, 2010, 96(4): 1-13

[2] Bouchard J, Anderd D, Delvillar R, et al. Column flotation simulation and control: An overview. *Minerals Engineering*, 2009, 22(6): 519-529

[3] Moolman D W, Aldrich C, Van Deventer J S J, et al. The interpretation of flotation froth surface by using digital analysis and neural networks. *Chemical Engineering Science*, 1995, 50(22): 3501-3513

[4] 刘文礼,路迈西,王凡等. 煤泥浮选泡沫图像纹理特征的提取及泡沫状态的识别. *化工学报*, 2003, 54(6): 830-835

[5] 郝元宏,韩静,齐春. 一种新的浮选泡沫图像识别方法. *西安交通大学学报*, 2011, 45(4): 104-108

[6] 王红平,齐春,李金标等. 基于主成分分析的矿物浮选泡沫图像分类与识别. *矿冶*, 2005, 14(9): 79-82

[7] Salton G, Wang A, Yang C S. A vector space model for automatic indexing. *Communication of ACM*, 1975, 18(11): 613-620

[8] Salton G. Developments in automatic text retrieval. *Science*, 1991, 253(5023): 974-979

- [9] Rajan K, Ramalingam V, Ganesan M, et al. Automatic classification of Tamil documents using vector space model and artificial neural network. *Expert Systems with Applications*, 2009, 36(8): 10914-10918
- [10] 张妹, 赵铁军, 赵华等. 基于内容相似度的文摘自动评测方法及其有效性分析. *高技术通讯*, 2006, 16(3): 241-245
- [11] Haralick R M, Shanmugam K, Dinstein I. Texture features for image classification. *IEEE Trans on Sys, Man and Cyb*, 1973, 3(6): 610-621

Flotation froth image classification based on vector space model

Wang Yalin, Zhang Runqin, Xie Yongfang, Gui Weihua

(School of Information Science and Engineering, Central South University, Changsha 410083)

Abstract

A flotation froth image classification method based on vector space model (VSM) is presented to solve the problem that existing flotation froth image classification methods based on froth bottom characteristics have the shortcomings of low accuracy and semantic gap. Referring to the method of text classification, the new method divides froth images into blocks, extracts these blocks' bottom characteristics and clusters them to build the table of the froth status words. Based on the table, the similarity between the words and the word frequency are calculated, and then the froth images is described with a bag-of-word vector. Finally, the classification and recognition of real time froth image is realized based on the VSM method. The results of the experiment performed by using the real plant data show that the proposed method can increase the classification accuracy to nearly 90%, which is higher than the accuracy of classification method based on bottom characteristics. The proposed method can solve the problem of semantic gap to a certain extent, showing its important value in industrial application.

Key Words: flotation image, classification and recognition, vector space model (VSM), froth state words, texture