

# 基于 Tamura 纹理特征的煤岩壳质组显微组分分类

王培珍<sup>a,c</sup>, 任 将<sup>a</sup>, 杜存铃<sup>a</sup>, 张代林<sup>b</sup>

(安徽工业大学 a. 电气与信息工程学院; b. 煤的洁净转化与综合利用安徽省重点实验室; c. 冶金减排与资源综合利用教育部重点实验室, 安徽 马鞍山 243032)

**摘要:** 根据煤岩壳质组各显微组分的纹理特点及其差异, 采用基于人类视觉感知的 Tamura 纹理描述方法, 提取其粗糙度、对比度、方向度、线像度、规则度等特征量, 对壳质组显微组分中角质体、藻类体和树皮体等3类组分典型样本进行描述, 并对特征量的可区分性进行分析。鉴于壳质组显微组分样本数的局限, 构建适合于小样本分类问题的支持向量机分类器, 分别采用不同特征量及特征量组合对3种典型组分进行分类。实验结果表明: 单个特征难以实现对显微组分的有效分类; 采用组合特征可明显提高分类效果, 其中线像度和方向度的组合效果最佳, 分类平均准确率可达98.9%。该结果可为煤岩其它显微组分的分类与识别提供参考。

**关键词:** 煤岩; 壳质组; 显微组分; Tamura 纹理; 支持向量机; 分类

**中图分类号:** TP 391.4 ; TQ 533.6      **文献标志码:** A      **doi:** 10.3969/j.issn.1671-7872.2018.02.007

## Classification of Macerals in Exinite of Coal Based on Tamura Features

WANG Peizhen<sup>a,c</sup>, RENG Jiang<sup>a</sup>, DU Cunling<sup>a</sup>, ZHANG Dailin<sup>b</sup>

(a. School of Electrical & Information Engineering; b. Anhui Key Laboratory of Clean Conversion and Utilization; c. Key Laboratory of Metallurgical Emission Reduction & Resources Recycling, Ministry of Education, Anhui University of Technology, Ma anshan 243032, China)

**Abstract:** According to the textural characteristics of macerals in exinite of coal and the difference between them, features of coarseness, contrast, directionality, linelikeness and regularity of samples were extracted with Tamura texture description method, which is based on human visual apperception, to describe the typical three macerals of cutinite, alginite and barkinite, and the distinguishability of features were analyzed. In view of the limitations of the sample number of exinite macerals, a support vector machine based classifier, which is fit for small sample size problem was built, features (single or joint) were employed for the classification of the three typical macerals. Experimental results show that it is difficult to achieve an effective classification with single feature, but after joint two of them, the effect of classification is improved obviously, and the combination of linelikeness and directionality performs its best, achieves an accuracy of 98.9%, which can offer a reference for further classification or recognition of other macerals of coal.

**Key words:** coal; exinite; maceral; Tamura texture; support vector machine; classification

煤岩不同显微组分的构成与其性质直接相关, 因此实现煤岩显微组分的自动分类与识别对煤岩工艺性质的界定具有重要意义。煤岩的显微组分来源多样, 在漫长的形成过程中受多种因素干扰, 其形态结构各异且较为复杂<sup>[1]</sup>。目前, 已有学者对煤显微图像分类问题进行研究, 并取得一定进展。如阮晓东等<sup>[2]</sup>通过提取煤岩显微图像中区域的轮廓, 根据轮廓的交点判断有重叠的圆形颗粒, 分析煤岩显微组分的形貌; 尹文花<sup>[3]</sup>通

收稿日期 2018-05-04

基金项目: 国家自然科学基金项目(51574004); 安徽省高校学科拔尖人才学术资助重点项目(2016041)

作者简介: 王培珍(1966—), 女, 安徽泾县人, 博士, 教授, 主要研究方向为图像处理与模式识别。

过显微组分灰度值的分布测算煤岩显微图像中相应组分含量,由于灰度值仅为煤岩显微组分的特征之一,且存在不同组分灰度值相近甚至相同的情况,致使测定的准确性不高;王素婷等<sup>[4]</sup>提取煤岩显微图像的局部二值模式(LBP)和灰度共生矩阵(GLCM)特征,采用支持向量机对煤岩显微组分大类别进行分类,但分类正确率有待提高。王培珍等<sup>[5-7]</sup>对于煤岩显微图像分析与识别进行了较为系统的研究,如采用改进的圆形LBP均匀模式对煤显微图像的纹理特征进行分析;采用小波变换将煤岩惰质组显微图像进行分解,提取不同尺度下的纹理特征并对其进行有效分类;采用主成分分析法对由灰度分布特征量和基于灰度共生矩阵的纹理特征量构成的煤岩镜质组初始特征量集进行抽取,并构建支持向量机对其分类,取得了一定的效果。壳质组是煤岩三大显微组分之一,主要来源于高等植物的角质层、木栓层等稳定成分或藻类等微生物的降解物<sup>[1]</sup>,形貌复杂多样,本课题组<sup>[8]</sup>通过改进传统的SUSAN(Univalued Segment Assimilating Nucleus)算法,对其显微组分之一析出沥青体的锐角角点进行检测,该方法可将析出沥青体从其它组分中分离出来。壳质组中其它组分的有效分类还鲜见报道。壳质组显微组分存在明显的纹理特征(随机纹理),且不同组分间纹理存在一定差异。鉴于Tamura纹理与人类视觉感知相吻合<sup>[9-10]</sup>在表征随机纹理时有其独特的特点,本文采用Tamura纹理描述方法对壳质组显微图像中典型组分特征量进行提取,并构建支持向量机分类器对其进行分类,以探索对煤岩壳质组显微组分进行有效分类的新方法。

## 1 壳质组显微图像的特点

壳质组显微组分类别较多,包括孢子体、角质体、木栓质体、树皮体、树脂体、藻类体、沥青质体、渗出沥青质体、荧光体、碎屑体等<sup>[11]</sup>。但某些组分较少出现,因此文中选择较为典型的角质体、藻类体、树皮体3类显微组分为对象,其显微图像如图1。

形成角质体植物的叶、枝、茎等表皮组织会产生分泌物(角质),这些分泌物经煤化作用后形成较为稳定的组分<sup>[12]</sup>,其显微图像在垂直层理的切面中呈长条状,且外缘平滑内缘呈锯齿状,如图1(a)藻类体来源于藻类,其形状亦类似于藻类,呈锯齿状、海绵状或其它不规则形状,如图1(b)中颜色较亮的不规则区域;树皮体由植物的茎和根等外部栓质化组织形成,大多呈扁平的长方形或砖形,排列比较规则,如图1(c)。综合图1可看出,煤岩壳质组显微组分的自然形貌具明显的纹理结构,且不同组分间纹理具一定的差异。

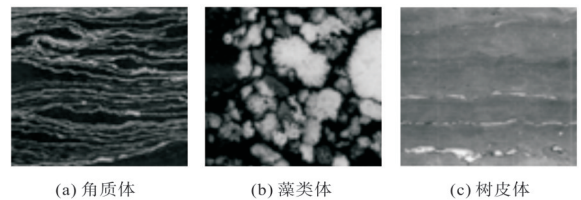


图1 3类典型壳质组显微图像

Fig. 1 Microscopic images of three kinds of typical macerals of exinite

## 2 特征提取

基于人类对纹理视觉感知的心理学研究,Tamura等<sup>[9]</sup>提出一种纹理特征的描述方法,其中定义了6个用于描述纹理的特征量,分别为粗糙度(coarseness)、对比度(contrast)、方向度(directionality)、线像度(linelikeness)、规则度(regularity)和粗略度(roughness)。粗略度是粗糙度和对比度两个属性的合成。故文中采用粗糙度、对比度、方向度、线像度和规则度等5个特征量对壳质组显微组分进行描述与分析。

### 2.1 特征量的构建

#### 2.1.1 粗糙度

粗糙度反映图像灰度变化的剧烈程度,纹理基元尺寸越大则感觉越粗糙,其定义如下。

计算图像中以坐标 $(x,y)$ 为中心、大小为 $2^k \times 2^k$ 的活动窗口(如 $1 \times 1, 2 \times 2, \dots, 32 \times 32$ ,窗口以中每个像元的灰度均值 $A_k(x,y)$ ,

$$A_k(x,y) = \frac{1}{2^{2k}} \sum_{i=x-2^{k-1}}^{x+2^{k-1}-1} \sum_{j=y-2^{k-1}}^{y+2^{k-1}-1} g(i,j) \quad (1)$$

其中 $(i,j)$ 表示邻域中像元的位置; $g(i,j)$ 表示位于 $(i,j)$ 处像元的灰度值; $k$ 确定参与均值计算的像素范围。

对于每个像元,分别计算其在水平及垂直方向上互不重叠的活动窗口之间的平均灰度差 $E_{k,v}, E_{k,h}$ ,

$$E_{k,v} = |A_k(x, y + 2^{k-1}) - A_k(x, y - 2^{k-1})| \quad (2)$$

$$E_{k,h} = |A_k(x + 2^{k-1}, y) - A_k(x - 2^{k-1}, y)| \quad (3)$$

对于每一个像元,寻找最优的尺寸  $S_{k_{\text{best}}} = 2^k$ ,使其平均灰度值  $E$  值达到最大  $E_{k_{\text{best}}}$ 。

$$E_{k_{\text{best}}} = \arg \max_k (E_{k,v}, E_{k,h}) \quad (4)$$

粗糙度  $F_{\text{crs}}$  定义为整幅图像中各像元最优尺寸  $S_{k_{\text{best}}}$  的平均值,

$$F_{\text{crs}} = \frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N S_{k_{\text{best}}}(i,j) \quad (5)$$

其中  $M$  和  $N$  分别表示图像宽度和高度。

### 2.1.2 对比度

对比度反映灰度图像中最暗和最亮灰度的层级,其差异范围的大小决定对比度的大小,此外灰度级的动态范围、边缘的尖锐程度、模式的重复周期等也会影响对比度。对比度  $F_{\text{con}}$  定义为

$$F_{\text{con}} = \frac{\sigma}{\alpha_4^{1/4}} \quad (6)$$

其中:  $\sigma$  表示图像灰度值的均方差;  $\alpha_4$  表示图像灰度统计量的峰态,  $\alpha_4 = \mu_4 / \sigma^4$ ;  $\mu_4$  表示四阶矩均值。

### 2.1.3 方向度

方向度描述纹理沿某方向集中的强度。方向度的计算基于每个像素点的梯度向量,该向量的模和方向分别定义为

$$|\Delta G| = (|\Delta_h| + |\Delta_v|) / 2 \quad (7)$$

$$\theta = \tan^{-1}(\Delta_v / \Delta_h) + \frac{\pi}{2} \quad (8)$$

其中  $\Delta_h$  和  $\Delta_v$  分别为  $(x, y)$  处水平与垂直方向灰度差分,分别如下所示的 2 个  $3 \times 3$  的模板计算。

$$\begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

量化各像素点的方向角  $\theta$  值,统计其相应方向角上边缘像素数量  $N_\theta(i)$ ,挑选出边缘强度(模)大于预定阈值的边缘像素点,设其数量为  $N_\theta(\phi)$ ,则相应方向上的直方图  $H_D(\phi)$  为

$$H_D(\phi) = N_\theta(\phi) / \sum_{i=0}^{n-1} N_\theta(i) \quad (9)$$

其中  $\phi = 0, 1, 2, \dots, n-1$ 。则方向度定义为

$$F_{\text{dir}} = 1 - r n_p \sum_p \sum_{\phi \in w_p} (\phi - \phi_p)^2 H_D(\phi) \quad (10)$$

其中:  $n_p$  为峰的数量;  $w_p$  为围绕该角度峰值与谷值之间的范围;  $r$  为与角度  $\theta$  量化水平相关的归一化因子;  $\phi_p$  为第  $p$  个峰值的位置。

### 2.1.4 线像度

线像度  $F_{\text{lin}}$  表示图像纹理是否具有线性结构,定义如下

$$F_{\text{lin}} = \frac{\sum_i \sum_j P_d(i,j) \cos\left[(i-j) \frac{2\pi}{m}\right]}{\sum_i \sum_j P_d(i,j)} \quad (11)$$

其中:  $P_d(i, j)$  表示相距为  $d$  的一组元素、方向编码分别为  $i, j$  时出现的次数;  $m$  表示划分角度的数目,文中取 16。

### 2.1.5 规则度

纹理具有重复排列的自然景象,其规则度往往难以用数学公式表示。若图像中存在与整幅图像纹理特征不一致的局部区域纹理特征,则图像的规则度减弱。所以可通过对原图像分区后每个子图像属性方差来测度纹理的规则度,这里考虑子图像的 4 个特征来衡量纹理的规则度  $F_{\text{reg}}$ 。

$$F_{\text{reg}} = 1 - \eta(\sigma_{\text{crs}} + \sigma_{\text{con}} + \sigma_{\text{dir}} + \sigma_{\text{lin}}) \quad (12)$$

其中  $\eta$  为标准化因子;  $\sigma_{crs}$ ,  $\sigma_{con}$ ,  $\sigma_{dir}$ ,  $\sigma_{lin}$  分别为  $F_{crs}$ ,  $F_{con}$ ,  $F_{dir}$ ,  $F_{lin}$  的标准方差。

2.2 样本特征数据分析

壳质组显微图像样本由光度计在油浸反光下获取。分别选取角质体、藻类体、树皮体 80 幅显微图像,按照 2.1 中关于 Tamura 纹理的定义分别计算上述 3 类壳质组典型显微组分的 5 个特征量。由于篇幅所限,表 1 仅给出其中 5 组典型样本数据。

为直观分析样本数据的分布特点及特征量的可区分性,对于每类组分选取 30 个样本绘制其特征量的分布,结果如图 2。由图 2 可看出:粗糙度作为特征量具有一定的可区分性,从小到大排序依次是角质体、树皮体、藻类体,主要原因在于角质体线性纹理比较细密,粗糙度较小,树皮体存在块状的纹理区域,相比于角质体粗糙度要大一些,而藻类体图像中有大量的藻类状区域,粗糙度最大;藻类体与角质体的对比度特征数据间有较多的交错;角质体由于纹理具明显的方向性,其方向度值较大,与其它两类组分具有明显的可区分性;线像度对于 3 个类别组分具有较好的区分性;对于 3 类组分来说,规则度分布相互交错,可区分性较差。综上所述,采用 Tamura 方法对煤岩壳质组纹理特征进行描述时,粗糙度、线像度对 3 类组分具有较好的可区分性;对比度、方向度次之;而规则度的可区分性较差,本文的后续工作中舍弃该特征量。

表 1 壳质组典型显微组分特征量

Tab.1 Texture features of typical macerals of exinite

显微组分	粗糙度	对比度	方向度	线像度	规则度
角质体	4.49	28.22	80.62	0.34	0.94
	5.08	36.08	80.38	0.32	0.93
	4.96	39.27	93.34	0.40	0.93
	5.12	31.43	80.25	0.19	0.91
	5.36	50.38	94.13	0.38	0.92
藻类体	6.14	47.81	19.69	0.08	0.94
	6.09	29.98	39.40	0.03	0.92
	6.32	66.23	27.09	0.06	0.90
	6.43	59.57	27.09	0.08	0.91
	6.16	59.98	52.26	0.04	0.90
树皮体	5.85	19.50	54.35	0.12	0.93
	5.45	19.65	70.89	0.12	0.92
	5.54	30.16	83.76	0.13	0.93
	6.09	36.86	75.90	0.14	0.93
	5.99	23.25	71.90	0.13	0.93

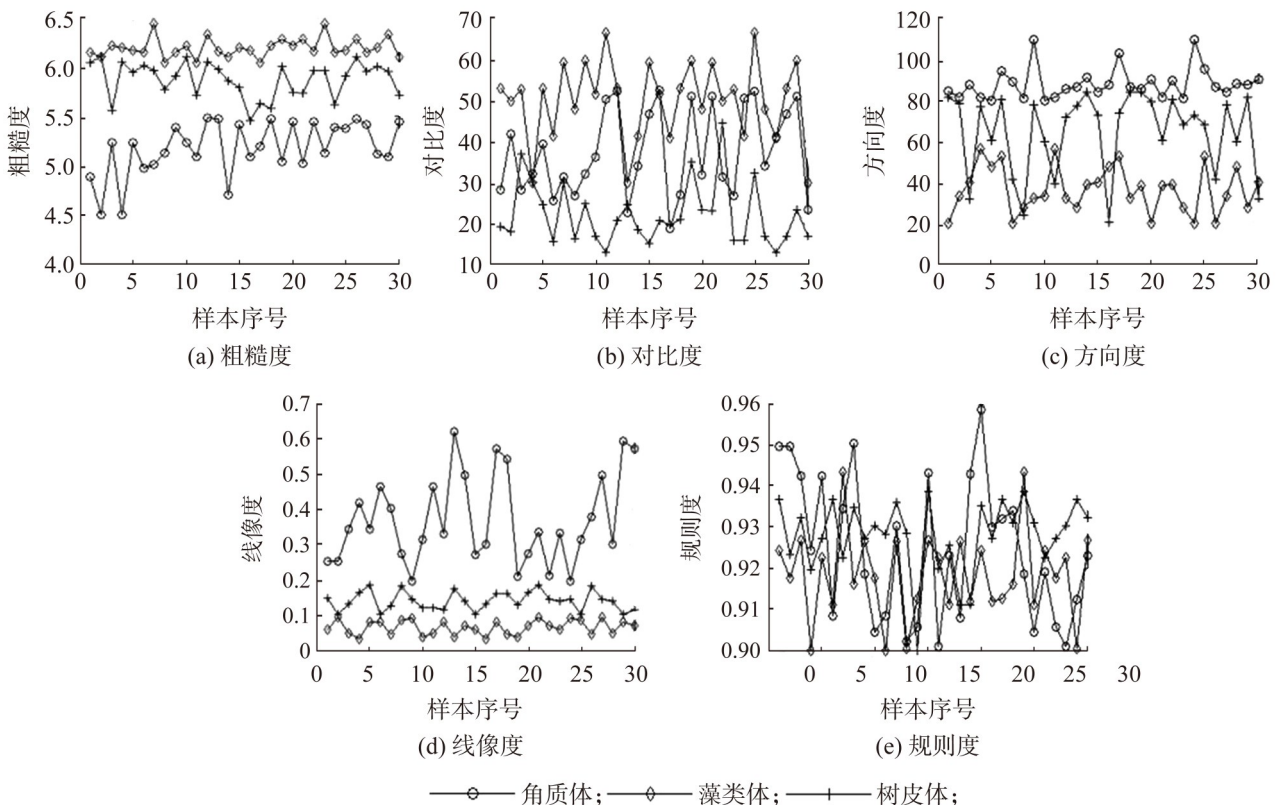


图 2 壳质组显微组分纹理特征分布

Fig. 2 Distribution of texture features of macerals of exinite

### 3 分类器的构建

支持向量机(Support Vector Machine, SVM)分类器以训练误差作为优化问题的条件,是目前解决小样本问题的最优分类方法之一,已在多领域的分类与预测问题中得到应用<sup>[13-14]</sup>。针对文中研究对象,构建支持向量机分类器对3类典型壳质组显微组分进行分类。

设有训练样本集  $\{(x_i, y_i), i = 1, 2, \dots, l\}$ ,  $x_i \in R^K$  表示样本数据,  $l$  表示集合中包含的样本个数,  $K$  为  $x_i$  的维数,  $y_i$  为  $x_i$  的类别编号。最优超平面为  $\omega^T x_i + b = 0 (i = 1, \dots, l)$ ,  $\omega$  和  $b$  分别为权重向量和偏置量。对于线性不可分问题,需满足

$$y_i(\omega^T x_i + b) \geq 1 - \zeta_i \quad (13)$$

其中  $\zeta_i$  为松弛变量,表示线性可分问题的偏离程度。为寻找可使训练样本错误分类误差最小的超平面,得以下优化问题

$$\varphi(\omega, \zeta) = \frac{1}{2} \omega^T \omega + C \sum_{i=1}^l \zeta_i \quad (14)$$

即寻找合适的权值向量  $\omega$ , 使  $\varphi(\omega, \zeta)$  最小,其中  $C$  为惩罚系数,表示对错分样本的惩罚程度。求解中,为实现非线性问题向线性问题的转化,引入核函数  $k(x_i, x)$ , 则上述二次型优化问题最终可表示为

$$f(x) = \text{sgn}[\sum_{i=0}^l \lambda_i k(x_i, x) + b] \quad (15)$$

其中  $\lambda_i$  为拉格朗日乘子。鉴于径向基函数(RBF)所需参数较少及对于缺少先验知识时所具备的良好性能,文中选择RBF核函数,形式为

$$k(x_i, x) = \exp(-\gamma \|x - x_i\|^2) \quad (16)$$

其中  $\gamma$  为核参数,控制函数的径向作用范围。

### 4 实验结果与分析

实验过程中,选取壳质组3类典型显微组分样本各80幅,其中训练样本各50幅,测试样本各30幅,样本图像大小为  $100 \times 100$ 。

采用单个特征量进行分类,结果如表2。由表2可看出:粗糙度对于3类组分的分类准确率均达83%以上,其中角质体的分类准确率96.7%。这是因为角质体的粗糙度较小,其特征量与树皮体只有少量重叠,而树皮体粗糙度特征量分布介于角质体与藻类体之间,且与两者都有交错,故分类准确率较低;采用对比度特征量进行分类时,分类效果均不理想,主要原因在于3类组分对比度特征分布相互交织,可区分性差;采用方向度进行分类时,角质体与藻类体可有效区分出来,树皮体的特征值分布与其它两类都有重叠,分类效果较差;线像度对于3类组分的分类准确率均达82%以上,尤其是藻类体和树皮体均达96.7%,具有良好的区分度;采用规则度作为分类特征时,几乎不能进行有效区分,这与3类组分中规则度的分布相互交错有关。上述分类结果与2.2关于特征量分析结果相映衬。

分析表2知,仅用单特征量对3类组分进行分类,整体分类效果不理想。为此,采用特征量的组合对其进行分类实验。根据前述分析及单特征分类结果,线像度对3类组分具有较好的可区分性,故将线像度作为第一特征量与其它特征组合,组合特征量分类结果如表3。由表3可看出,采用多特征量进行分类时,分类的准确率明显提高。其中线像度、方向

表2 基于单个特征量分类结果

Tab. 2 Results of classification with single feature

特征向量	组分	分类准确率/%	
		训练样本	测试样本
粗糙度	角质体	100	96.7
	藻类体	91.7	86.7
	树皮体	85.2	83.3
对比度	角质体	76.9	70.0
	藻类体	75.0	70.0
	树皮体	70.4	73.3
方向度	角质体	96.1	96.7
	藻类体	100	96.7
	树皮体	55.6	50.0
线像度	角质体	84.6	82.1
	藻类体	100	96.7
	树皮体	100	96.7

度两个特征量组合后分类效果最好,平均分类准确率达98.9%。这是因为3类典型组分中,线像度具有较好的可区分性,且方向度对于角质体、藻类体两类具有较强的可区分性,两个特征量组合后相得益彰,分类准确率大为提高。

## 5 结 论

在分析壳质组中角质体、树皮体、藻类体等组分显微图像特点的基础上,采用Tamura纹理分析方法提取上述3类显微组分的特征量,分析各特征量的可区分性,并构建RBF支持向量机进行分类实验,得如下主要结论:

1) 基于Tamura纹理的5个特征量中,线像度具有较好的可区分性,方向度、粗糙度次之,规则度可区分性最差;

2) 采用单特征量进行分类时,线像度整体效果较好,方向度和粗糙度次之,但方向度对树皮体的分类效果较差;采用单特征量进行分类的整体效果不理想;

3) 将特征量进行组合后,分类的准确率大为提高,其中线像度与方向度组合所得分类效果最好。

本文仅对壳质组中角质体、树皮体、藻类体3类组分进行了较为有效的分类,当添加其它组分至分类器时,其特征量间可区分性会减弱,分类的准确率会下降,但本文的分类方法具有一定的通用性,可推广至壳质组其它组分和煤岩其它组别显微组分的分类与识别中。

### 参考文献:

- [1] 中国煤田地质总局. 中国煤岩学图鉴[M]. 徐州:中国矿业大学出版社,1996:5-14.
- [2] 阮晓东,赵文峰. 煤粉显微图像中重叠颗粒识别的方法[J]. 煤炭学报,2005,30(6):769-771.
- [3] 尹文花. 基于空间域数字图像处理方法的煤岩分析[D]. 太原:太原理工大学,2010:19-27.
- [4] 王素婷,朱宪坤,吕青. 基于RILBP-GLCM算法的煤岩显微组分识别[J]. 煤炭技术,2017,36(3):142-144.
- [5] 王培珍,董双,张代林,等. 基于圆形LBP均匀模式的煤镜质体显微组分纹理分析[J]. 安徽工业大学学报(自然科学版),2014,31(2):147-151.
- [6] 王培珍,殷子皖,丁海涛,等. 基于小波变换的煤岩惰质组显微图像分类[J]. 安徽工业大学学报(自然科学版),2016,33(3):278-283.
- [7] 王培珍,殷子皖,王高,等. 一种基于PCA与RBF-SVM的煤岩显微组分镜质组分类方法[J]. 煤炭学报,2017,42(4):977-984.
- [8] 王培珍,杜存铃,孙瑞,等. 煤岩显微组分中渗出沥青体角点特征检测方法[J]. 安徽工业大学学报(自然科学版),2015,32(4):376-380.
- [9] TAMURA H, MORI S, YAMAWAKI T. Textural features corresponding to visual perception[J]. IEEE Transactions on Systems, Man and Cybernetics, 1978, 8(6):460-473.
- [10] 梅军,张森林,樊臻. 基于Tamura纹理特征的织物组织识别算法[J]. 轻工机械,2017,35(4):52-55.
- [11] 全国煤岩标准化技术委员会. 烟煤显微组分分类:GB/T 15588-2013[S]. 北京:中国标准出版社,2014:1-6.
- [12] 韩德馨. 中国煤岩学[M]. 徐州:中国矿业大学出版社,1996:349-451.
- [13] 郭明玮,赵宇宙,项俊平,等. 基于支持向量机的目标检测算法综述[J]. 控制与决策,2014,29(2):193-200.
- [14] OLGUN M, ONARCAN A O, ÖZKAN K, et al. Wheat grain classification by using dense SIFT features with SVM classifier[J]. Computers & Electronics in Agriculture, 2016, 122(2016):185-190.

责任编辑:何莉

表3 基于组合特征量分类结果

Tab. 3 Results of classification with joint features

特征组合	组分	识别率/%	
		训练样本	测试样本
线像度+ 粗糙度	角质体	100	100
	藻类体	89.1	90.3
	树皮体	100	96.7
线像度+ 方向度	角质体	98.0	100
	藻类体	100	100
	树皮体	100	96.7
线像度+ 对比度	角质体	71.0	80.2
	藻类体	70.1	76.0
	树皮体	100	96.7