

宋镇, 姬长英, 张波. 基于光谱与图像信息的杏鲍菇多糖含量检测 [J]. 华南农业大学学报, 2019, 40(3): 104-110.
SONG Zhen, JI Changying, ZHANG Bo. Detection of polysaccharide content in *Pleurotus eryngii* based on spectral and image information[J]. Journal of South China Agricultural University, 2019, 40(3): 104-110.

基于光谱与图像信息的杏鲍菇多糖含量检测

宋 镇, 姬长英, 张 波
(南京农业大学 工学院, 江苏 南京 210031)

摘要:【目的】利用高光谱成像技术实现杏鲍菇 *Pleurotus eryngii* 多糖含量的快速无损检测。【方法】利用高光谱图像采集系统获取 350~1 021 nm 波长范围内的杏鲍菇高光谱图像, 同时利用苯酚-硫酸法测定对应样本的多糖含量。通过波段运算和阈值分割构建掩膜图像, 使样本与背景相分离。采用主成分分析 (PCA) 处理原始高光谱图像, 获得代表原始图像 99% 信息的 2 个主成分图像 (PC1、PC2), 然后利用连续投影算法 (SPA) 选出 554.4、772.8、811.4、819.1、855.6、986.3 和 1 019.5 nm 7 个特征波长及对应的光谱特征, 分别提取 7 个特征波长图像和 2 个主成分图像的纹理与颜色特征, 最后利用偏最小二乘回归 (PLSR) 建立杏鲍菇样本基于不同图像特征与多糖含量之间的关系模型。【结果】从校正集决定系数 (R_c^2) 来看, 基于特征光谱+特征波长图像特征+主成分图像特征的模型效果最好, $R_c^2=0.954$, $RMSE_c=0.341$; 从预测集决定系数 R_p^2 来看, 基于特征光谱+特征波长图像特征的模型效果最好, $R_p^2=0.868$, $RMSE_p=0.539$ 。【结论】该研究结果可为杏鲍菇多糖含量的快速、无损检测提供一定的参考。

关键词: 高光谱图像; 杏鲍菇; 多糖; 主成分分析; 特征波长; 纹理特征
中图分类号: S646.9 **文献标志码:** A **文章编号:** 1001-411X(2019)03-0104-07

Detection of polysaccharide content in *Pleurotus eryngii* based on spectral and image information

SONG Zhen, JI Changying, ZHANG Bo
(College of Engineering, Nanjing Agricultural University, Nanjing 210031, China)

Abstract: 【Objective】 To quickly and non-destructively detect polysaccharide content in *Pleurotus eryngii* using hyperspectral imaging technology. 【Method】 Hyperspectral images of *P. eryngii* in the visible and near infrared (390-1050 nm) regions were acquired using the hyperspectral imaging system. Polysaccharide contents in corresponding *P. eryngii* samples were measured by phenol sulfuric acid method. The binary mask image was constructed by the method of band operation and threshold segmentation to separate the sample area from the background area. Principal component analysis (PCA) was used to process the original hyperspectral images, and two principal component images (PC1, PC2) representing 99% information of the original image were obtained. Seven characteristic wavelengths of 554.4, 772.8, 811.4, 819.1, 855.6, 986.3, and 1 019.5 nm were selected using the successive projection algorithm (SPA). Texture and color data were extracted from two principal component images and seven characteristic wavelength images, and spectral data were also extracted from seven characteristic wavelength images. Using partial least squares regression (PLSR) models were established based on the correlations of different image features and polysaccharide contents in *P. eryngii* samples. 【Result】 According to the determination coefficient of the calibration set (R_c^2), the best model is the one based

on characteristic spectra, characteristic wavelength images and principal component images with $R_c^2=0.954$ and $RMSE_c=0.341$. According to the determination coefficient of the prediction set (R_p^2), the best model is the one based on characteristic spectra and characteristic wavelength images with $R_p^2=0.868$ and $RMSE_p=0.539$.

【Conclusion】 This study provides references for fast and non-destructive detection of polysaccharide content in *Pleurotus eryngii*.

Key words: hyperspectral image; *Pleurotus eryngii*; polysaccharide; principal component analysis; characteristic wavelength; texture feature

杏鲍菇 *Pleurotus eryngii* 又名刺芹侧耳, 是近年来培育成功的集食用、药用、食疗于一身的珍稀食用菌新品种^[1]。杏鲍菇含有丰富的蛋白质和人体必需的多种氨基酸, 更重要的是含有丰富的多糖类物质, 多糖可以降血脂、降血糖、增强人体免疫能力, 还具有一定的抗癌作用^[2-4], 因此, 多糖含量成为衡量杏鲍菇营养价值的最主要指标^[5]。但是杏鲍菇和大多数菌类一样, 不宜储藏, 因此, 杏鲍菇采摘后的深加工是杏鲍菇储藏的一个重要环节。目前杏鲍菇的储藏以干制及多糖类物质的提取为主要加工方法。干制可以延长杏鲍菇的储藏时间, 但干制过程中杏鲍菇的营养物质会受到一定的损耗, 因此对干燥工艺提出了较高的要求, 不仅要延长杏鲍菇的储藏时间, 也要保证杏鲍菇干制品的营养价值。多糖作为最主要的营养指标, 实现杏鲍菇多糖的快速无损检测对指导和改进干燥工艺有着重要意义。传统的多糖检测采用苯酚-硫酸法, 测定前需要制作葡萄糖标准曲线, 使用该方法测定多糖费时费力且具有破坏性, 不能满足快速、无损检测的要求。

高光谱成像技术是近年发展起来的快速、无损检测方法, 它结合了图像技术和光谱技术, 既能获得图像纹理信息又能获得多波段的光谱信息, 因此在农产品外观品质评价和内部成分含量的检测方面得到越来越广泛的应用^[6-9]。梁琨等^[10] 基于高光谱和 CARS-IRIV 算法实现了库尔勒香梨 *Pyrus sinkiangensis* 可溶性固形物含量检测, 结果表明 CARS-IRIV 算法能有效减少 CARS 算法提取的变量个数且预测精度更高。李晓丽等^[11] 比较了不同的光谱预处理方法和建模方法, 实现了茶叶表没食子儿茶素没食子酸酯 (Epigallocatechin gallate, EGCG) 浓度的检测和 EGCG 的分布可视化。此外, 有学者研究发现, 由于农产品存在内外品质各向异性的特点, 单独的光谱检测采用点源式, 难以反应空间的差异性, 综合利用高光谱图像的纹理信息和光谱信息, 可以提高模型的可靠性^[12]。如于慧春等^[13] 融合光谱信息和多个图像纹理信息建立枸杞 *Lycium* 多糖和总糖预测模型, 结果表明基于融合特征的模

型预测精度高于单一特征模型。近年来, 虽然利用高光谱技术取得了很多科研成果, 但高光谱设备普遍较贵, 制约了其在实际生产中的推广应用。但随着越来越多的技术人员投入到高光谱的研发中, 未来的高光谱设备成本将会大大降低。另外, 政府和科研单位合作, 成果共享, 建立高光谱图像数据库, 也会推动高光谱技术的发展。从果蔬农药残留检测到化妆品重金属检测, 水体、土壤等环境污染监测, 再到牙齿、皮肤等医学检测, 高光谱技术在检测方面展现了其得天独厚的优势, 高光谱技术的应有前景可期。

本试验拟对杏鲍菇进行干燥处理, 获取不同干燥时间、不同干燥条件下的杏鲍菇样本, 通过高光谱成像系统获得杏鲍菇样本的高光谱图像纹理信息和光谱信息, 分别提取针对杏鲍菇多糖的有效特征集, 建立基于图像特征、光谱特征以及两者融合特征的预测模型, 以期实现杏鲍菇多糖含量的准确预测。

1 材料与方法

1.1 试验样本制备

试验所用杏鲍菇均购于南京市浦口区果蔬市场, 挑选新鲜、表皮完整无损伤的杏鲍菇若干。对杏鲍菇进行清洗, 去除表面污垢, 将杏鲍菇沿菇长方向作切片处理, 切片是为了方便杏鲍菇的干燥以及后续高光谱图像的采集。为了减小试验结果的偶然性, 将杏鲍菇样本进行分组, 每组采用不同的干燥温度、湿度、风速等因素组合。干燥过程中随机取样, 获取其高光谱图像后, 立即进行多糖的提取和测定, 试验共计获得 220 个杏鲍菇片样本的高光谱图像以及对应的多糖含量数据。

1.2 高光谱图像的采集

试验中所用的高光谱系统包括: Imspector 型光谱仪 (芬兰 Specim 公司)、CCD 相机 (美国 Imperx 公司)、镜头、21 V/150 W 线性卤素灯 (美国 Illumination 公司)、暗箱、电控移动平台以及计算机等部件。成像光谱仪波长范围: 350~1 021 nm; 光谱仪的分辨

率为 2.8 nm; 图像分辨率: 1 632 像素×971 像素。参数设置如下: 曝光时间 52 ms, 样本与镜头的距离为 330 mm, 传送带移动速度为 1.08 mm/s, 将杏鲍菇片放在黑色背景布上进行高光谱图像采集。

为了减少光照和暗电流对试验的影响, 增强高光谱图像的可靠性, 需要对获取的高光谱图像进行黑白校正, 校正公式为^[14]:

$$R = \frac{R_0 - R_b}{R_w - R_b}, \tag{1}$$

式中: R 为校正后的光谱图像, R_0 为校正前的漫反射光谱图像, R_b 为关电源并合上镜盖后的暗图像, R_w 为标准白板的漫反射光谱图像。

1.3 高光谱图像预处理

1.3.1 主成分分析 主成分分析 (Principal component analysis, PCA,) 是一种应用广泛的降维算法^[15]。主成分分析通过正交变换将原变量转换为一组新的、互相无关的变量, 尽可能多地解释原变量的所有信息。利用 ENVI4.8 软件进行主成分分析, 可以获得由原始高光谱图像数据中所有波段下的图像经过线性组合而形成的主成分图像。利用 PCA 降维后的图像用式 (2) 表示。

$$PC_m = \sum_{i=1}^m \beta_{i,m} \lambda_i, \tag{2}$$

式中: PC_m 为第 m 个主成分图像; λ_i 为第 i 个波段所对应的图像; $\beta_{i,m}$ 为第 m 个主成分在 λ_i 下的权重系数, $\beta_{i,m}$ 的值越大, 表明 λ_i 对主成分图像 PC_m 的贡献度越大。

由于前 2 个主成分图像已代表了原始高光谱图像数据 99% 的信息, 所以选取前 2 个主成分进行分析 (图 1)。

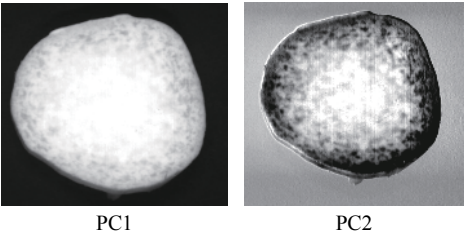


图 1 前 2 个主成分图像

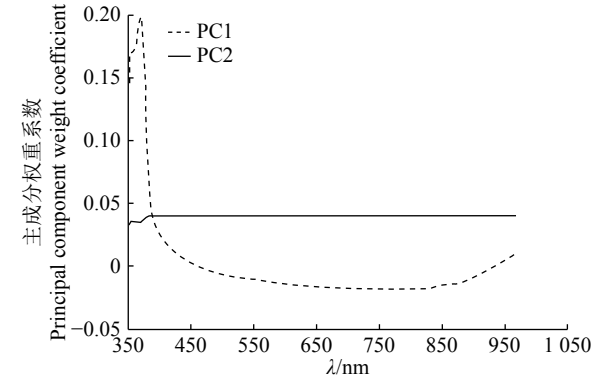


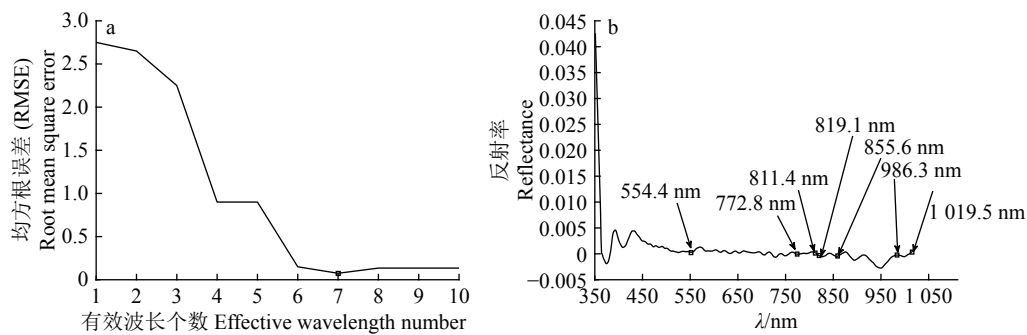
图 2 前 2 个主成分图像下各波段的平均权重系数

Fig. 2 The average weight coefficient of each band in images of the first two principal components

连续投影算法 (Successive projection algorithm, SPA)^[16] 是一种使矢量空间共线性最小化的前向变量选择算法, 对每一次循环所提取的变量, 分别建立多元线性回归模型, 校正集交互验证均方根误差 (Root mean square error, RMSE) 的最小值所对应的特征集, 即为最优特征波长集。在进行特征波长提取前, 采用 SG (Savitzky-Golay) 平滑、一阶微分 (First derivative, FD) 算法对光谱数据进行预处理, SG 平滑可以消除噪音, 提高信噪比, 一阶微分可以消除基线干扰, 提高分辨率和灵敏度^[17-18]。图 3 为 SPA 算法筛选特征波长的过程。由图 3a 可知, 随着有效波长数目的增加, RMSE 不断减小, 当筛选波长数为 7 时, RMSE 最小, 因此选取 7 个有效特征波长: 554.4、772.8、811.4、819.1、855.6、986.3 和 1 019.5 nm (图 3b)。提取 220 个样本在 7 个波长下的光谱数据作为样本的光谱特征, 将光谱数据存储为 220×7 的矩阵格式。

1.3.3 图像特征提取 图像特征主要包含颜色特征、纹理特征、形状特征和空间关系特征。形状特征和空间关系特征与多糖含量无明显关系, 干燥过程中, 糖类物质会发生焦糖化作用, 这不仅会导致样本的光谱特性发生变化, 同时也会导致样本的颜色特征和纹理特征发生一定的变化, 焦糖化作用越明显, 杏鲍菇片样本颜色越深, 像素点的灰度值会明显不同。颜色矩是一种有效的颜色特征提取方法, 将图像中的颜色分布用其矩表示。利用颜色一阶矩 (平均值 Average)、颜色二阶矩 (方差 Variance) 和颜色三阶矩 (偏斜度 Skewness) 描述颜色分布。本试验计算灰度图像单通道的一至三阶矩。灰度共生矩阵 (Gray-level co-occurrence matrix, GLCM) 是一种有效的纹理特征提取方法, 利用 Graycomatrix 函数 (Matlab), 距离参数值选取 1, 方向依次选取 0、45°、90°、135°, 分别对图像中的杏鲍菇样本进行对

1.3.2 特征波长的选择 首先基于主成分图像权重系数选择特征波长, 图 2 为 220 个杏鲍菇片样本的前 2 幅主成分图像下各波长的平均权重系数曲线, 从图 2 中可以看出, 波峰、波谷不明显, 且集中在噪声较大的两端, 不利于提取特征波长。因此本文直接对提取出的光谱数据进行特征波长选择。



a: 均方根误差随变量个数的变化; b: SPA 优选特征波长分布图
a: Variation of root mean square error with the number of variables; b: Distribution map of SPA characteristic wavelength

图 3 SPA 优选特征波长过程

Fig. 3 Process of selecting characteristic wavelength based on SPA

比度、能量、同质性和相关性纹理特征提取。1 个样本提取 2 幅主成分图像、7 幅特征波长图像, 共计 144 个纹理特征和 27 个颜色特征。

1.4 杏鲍菇片样本多糖含量测定

根据前人研究结果^[19-20], 采用水浴加热法提取杏鲍菇样品内的多糖, 采用苯酚-硫酸法测定多糖含量。

1.5 偏最小二乘回归

偏最小二乘回归 (Partial least squares regression, PLSR)^[21] 集多元线性回归分析+典型相关分析+主成分分析于一身, 能够提取对系统解释能力较强的综合变量, 从而排除无解释能力的信息, 还可以有效地解决变量多重共线性的问题。利用 PLSR 算法建立回归模型。

以决定系数 (R^2) 和均方根误差 (RMSE) 作为模型的评价指标。 R^2 反应模型的稳定性, 越接近于 1, 说明模型的稳定性越好, 拟合程度越高。RMSE 反映模型的预测精度, RMSE 越小, 模型预测能力越好。利用 Matlab R2016a 软件进行建模。

2 结果与分析

2.1 高光谱图像掩膜

为了分割样本和背景, 采用 ENVI 软件进行掩膜处理。样本区域与背景区域在不同波长下的反射

率曲线如图 4 所示。由于事先将杏鲍菇样本放置在黑色背景布上进行高光谱图像采集, 因此背景区域的反射率接近于 0, 与样本区域的反射率差异非常明显, 有利于二者的分割, 由于高光谱设备硬件原因, 高光谱图像在开始波段和结束波段受噪声影响较大, 背景区域反射率曲线在一开始的波动也说明了这点。因此将波段选取在中间位置进行分割, 经过多次分割结果对比, 750 nm 处的分割效果较好。具体的分割处理过程如图 5 所示。图 5a 为杏鲍菇样本在 R 分量为 663.2 nm, G 分量在 555.4 nm, B 分量在 450.2 nm 下的 RGB 图像。图 5b 为 750 nm 波长下的高光谱图像, 对该波长下的图像设定合适

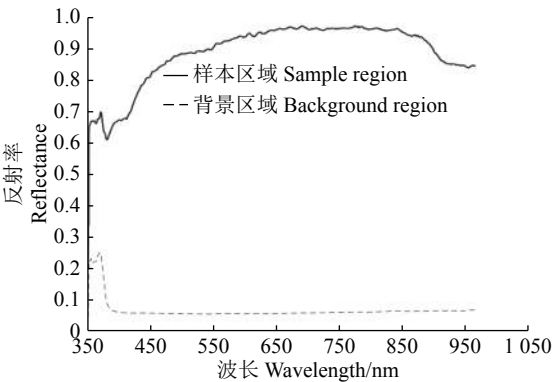
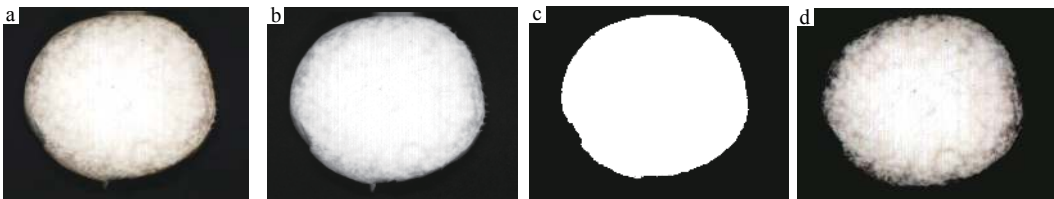


图 4 样本区域与背景区域的光谱曲线
Fig. 4 Spectral curves of the sample region and the background region



a: 原始高光谱图像; b: 750 nm 处图像; c: 二值掩膜图像; d: 掩膜处理后的高光谱图像
a: Original hyperspectral image; b: Image of 750 nm; c: Image with binary mask; d: Hyperspectral image after masking

图 5 图像分割流程

Fig. 5 Process of image segmentation

的阈值即可分离背景和样本,分割阈值应大于背景的最大反射率,小于杏鲍菇样本的最小反射率,由图 4 可知阈值选取范围较广,经过多次设定阈值分割,阈值取 0.5 时,分割出的杏鲍菇二值掩膜图像样本最为完整(图 5c);然后将原始图像与杏鲍菇掩膜图像作逻辑与运算,得到去除背景后的杏鲍菇样本高光谱图像(图 5d)。220 个杏鲍菇样本均作上述处理提取样本区域作为感兴趣区域(ROI)。

表 1 校正集与预测集样本多糖含量的统计结果

样本集 Sample set	n	w(多糖)/(g·kg ⁻¹) Polysaccharide content		
		最大值 Maximum	最小值 Minimum	平均值 Average
校正集 Calibration set	132	64.2	23.6	43.2
预测集 Prediction set	88	58.1	31.5	40.2

2.3 特征光谱数据建模

220 个杏鲍菇片样本的光谱数据经 SG-FD 方法预处理后,分别提取其 7 个特征波长对应的光谱平均值,利用 PLSR 模型建立杏鲍菇片样本光谱数据与多糖含量之间的关系模型,对不同样本的多糖含量进行预测。由表 2 可知,基于特征光谱的 PLSR 模型的校正集决定系数 R^2 偏低,模型的稳定性不高。这是由于 SPA 算法在寻找最低限度冗余信息的变量时会降低信号的信噪比^[23],SPA 算法共筛选出 7 个特征变量,与 616 个变量相比数目明显减少,虽然简化了模型,却降低了精度。因此需要进一步增加有效的特征信息,来提高模型的预测能力。

2.4 图像数据建模

图像数据的提取包括 2 幅主成分图像和 7 幅特征波长图像,基于颜色矩计算灰度图像单通道的一至三阶矩,提取颜色特征数据。基于灰度共生矩阵(Gray level co-occurrence matrix, GLCM)从 4 个角度(0、45°、90°、135°)对图像进行对比度、能量、同质性和相关性的数据提取。单个样本共计提取 144 个纹理数据和 27 个颜色数据。纹理数据较多,影响模型运算速度,因此有必要进一步对纹理数据进行筛选。本文将提取出的原始纹理数据与对应的多糖含量进行相关性分析,结果如表 3 所示。由表 3 可以看出,主成分图像 PC2 的 4 个纹理特征与多糖含量相关性较低,因此剔除主成分图像 PC2 的纹理特征。当纹理特征为对比度时,其值均与多糖含量呈负相关,相关系数达到了-0.7 左右,相关性较好。当纹理特征为能量、同质性和相关性时,其值均与多糖含量呈正相关,其中能量的相关系数达到了

2.2 样本划分结果

利用 Kennard-Stone(KS) 算法^[22]进行样本划分以提高回归模型稳定性。设定校正集与预测集之比为 3:2,选择 132 个样本作为校正集,88 个样本作为预测集,样本集多糖含量统计结果如表 1 所示。由表 1 可见,校正集的多糖含量范围能覆盖预测集的多糖含量范围,并且 2 个样本集的多糖含量平均值较为接近。因此,样本划分合理。

0.6 左右,相关性较好。出现上面结果的原因可能是:干燥会导致多糖的损耗,多糖含量低说明杏鲍菇片干燥到了一定阶段,此时杏鲍菇片由于失水表

表 2 基于不同特征的模型对多糖含量的预测性能比较

模型 ¹⁾ Model	校正集 Calibration set		预测集 Prediction set	
	R_c^2	RMSE _c	R_p^2	RMSE _p
1	0.876	0.511	0.833	0.632
2	0.778	0.721	0.738	0.778
3	0.732	0.782	0.628	0.933
4	0.821	0.650	0.772	0.714
5	0.912	0.451	0.868	0.539
6	0.855	0.546	0.757	0.767
7	0.954	0.341	0.815	0.661

1) 1: 基于特征光谱模型;2: 基于特征波长图像特征模型;3: 基于主成分图像特征模型;4: 基于特征波长图像特征+主成分图像特征模型;5: 基于特征光谱+特征波长图像特征模型;6: 基于特征光谱+主成分图像特征模型;7: 基于特征光谱+特征波长图像特征+主成分图像特征模型

1) 1: Model based on characteristic spectra; 2: Model based on feature of characteristic wavelength images; 3: Model based on feature of principal component images; 4: Model based on features of characteristic wavelength images and principal component images; 5: Model based on characteristic spectra and feature of characteristic wavelength images; 6: Model based on characteristic spectra and feature of principal component images; 7: Model based on characteristic spectra and features of characteristic wavelength images and principal component images

表 3 不同角度下特征波长图像和主成分图像的纹理特征与多糖含量的相关性

Table 3 Correlation between polysaccharide content and texture parameters of characteristic wavelength image and principal component image at different angles

项目 Item	λ/nm 或主成分 or PC	对比度				能量				同质性				相关性			
		Contrast				Energy				Homogeneity				Correlation			
		0	45°	90°	135°	0	45°	90°	135°	0	45°	90°	135°	0	45°	90°	135°
特征波长图像 Characteristic wavelength image	554.4	-0.7	-0.7	-0.8	-0.8	0.6	0.6	0.6	0.6	0.2	0.2	0.2	0.3	0.4	0.3	0.3	0.3
	772.8	-0.6	-0.8	-0.8	-0.8	0.7	0.7	0.7	0.7	0.2	0.2	0.2	0.2	0.5	0.5	0.5	0.4
	811.4	-0.7	-0.8	-0.7	-0.7	0.7	0.7	0.7	0.7	0.2	0.2	0.2	0.2	0.4	0.4	0.4	0.4
	819.1	-0.7	-0.7	-0.7	-0.7	0.6	0.7	0.6	0.7	0.3	0.3	0.2	0.2	0.4	0.4	0.5	0.4
	855.6	-0.8	-0.7	-0.7	-0.7	0.7	0.7	0.6	0.6	0.3	0.2	0.3	0.2	0.3	0.4	0.4	0.4
	986.3	-0.7	-0.7	-0.7	-0.8	0.6	0.6	0.6	0.6	0.2	0.2	0.2	0.2	0.5	0.5	0.4	0.4
	1 019.5	-0.7	-0.6	-0.6	-0.6	0.7	0.7	0.6	0.6	0.3	0.3	0.3	0.3	0.4	0.5	0.4	0.5
主成分图像 Principal component image	PC1	-0.7	-0.8	-0.8	-0.7	0.6	0.6	0.6	0.6	0.3	0.4	0.4	0.4	0.5	0.4	0.5	0.5
	PC2	-0.4	-0.5	-0.5	-0.4	0.5	0.4	0.4	0.4	0.2	0.2	0.2	0.2	0.3	0.3	0.4	0.3

面皱缩, 沟纹加深, 而对比度与能量分别反映图像纹理的深浅和粗细度, 因此与多糖含量相关性较好。相关性与同质性则是反应图像局部的灰度相关性, 局部的灰度变化与多糖含量相关程度不大。综上, 选取对比度和能量作为最终的纹理特征 (数据矩阵由 220×144 降到 220×64) 分别建立基于主成分图像、特征波长图像和两者纹理特征融合的 PLSR 模型。建模结果如表 2 所示。由表 2 可知, 基于两类图像特征融合所建立的模型预测结果优于各自建立的模型, 这是由于融合特征包含的信息全面, 但三者的预测集决定系数都较低, 都低于基于光谱特征建立的模型, 说明相比于图像特征, 光谱特征能更好地反映样本多糖含量的变化。

2.5 特征光谱与图像数据融合建模

将图像特征光谱数据与图像数据在特征层进行融合, 建立 3 个基于融合特征的预测模型。由表 2 可知, 在 3 个融合模型中, 通过校正集决定系数 R_c^2 来看, 基于特征光谱+特征波长图像特征+主成分图像特征的模型校正集结果最好, 这是因为融合三者的信息最全面, 因此构建出的模型具有较好的自测能力, 但对不同类的特征进行叠加, 既增加了模型计算量, 又引入了一部分冗余信息, 从而影响了模型预测集的预测性能, 导致预测集决定系数 R_p^2 偏低。通过预测集决定系数 R_p^2 来看, 基于特征光谱+特征波长图像特征的模型效果最好, 且优于单独利用光谱数据建立的模型, 这是因为 SPA 算法筛选的特征波长变量较少, 而特征光谱信息与特征波长图像的纹理信息具有较强的相关性, 光谱与纹理信息的融合增强了模型的预测能力。

3 结论

基于高光谱成像技术获取杏鲍菇片在 350~1 021 nm 的高光谱图像, 通过 ENVI 4.8 软件的主成分分析提取前 2 幅主成分图像, 光谱数据经 (SG) 平滑、一阶微分 (FD) 预处理后, 基于连续投影算法 (SPA) 筛选出 7 个特征波长, 并获取 7 个特征波长下的图像, 然后进行光谱数据和图像数据的提取, 最后基于特征光谱数据、图像数据和融合数据共建立 7 个偏最小二乘回归 (PLSR) 预测模型。比较 7 个模型性能发现: 1) 基于光谱数据建立的模型性能优于图像数据建立的模型, 说明光谱数据和多糖含量有更好的相关性; 2) 从模型性能指标校正集决定系数 R_c^2 来看, 基于特征光谱+特征波长图像特征+主成分图像特征的模型效果最好, $R_c^2 = 0.954$, $RMSE_c = 0.341$; 3) 从预测集决定系数 R_p^2 来看, 基于特征光谱+特征波长图像的模型效果最好, $R_p^2 = 0.868$, $RMSE_p = 0.539$ 。

参考文献:

[1] 李月梅, 采俊香, 牛瑞青. 不同基质配方工厂化栽培杏鲍菇研究 [J]. 北方园艺, 2012(7): 177-179.

[2] MARIGA A M, PEI F, YANG W J, et al. Immunopotential of *Pleurotus eryngii* (DC. ex Fr.) Quel.[J]. *J Ethnopharmacol*, 2014, 153(3): 604-614.

[3] ZHANG A, LI X, XING C, et al. Antioxidant activity of polysaccharide extracted from *Pleurotus eryngii* using response surface methodology.[J]. *Int J Biol Macromol*, 2014, 65(5): 28-32.

[4] 郑素玲, 郭立英, 范永山. 杏鲍菇多糖对老龄小鼠抗疲劳能力的影响 [J]. 食品科学, 2010, 31(7): 269-271.

[5] 姚秋萍, 杨琼, 卫亚丽, 等. 杏鲍菇多糖研究进展 [J]. 食

- 品工业科技, 2017, 38(21): 347-351.
- [6] M. ELMASRY G, NAKAUCHI S. Image analysis operations applied to hyperspectral images for non-invasive sensing of food quality: A comprehensive review[J]. *Biosyst Eng*, 2016, 142: 53-82.
- [7] 李江波, 饶秀勤, 应义斌. 农产品外部品质无损检测中高光谱成像技术的应用研究进展 [J]. *光谱学与光谱分析*, 2011, 31(8): 2021-2026.
- [8] 高俊峰, 张初, 谢传奇, 等. 应用近红外高光谱成像技术预测甘蔗可溶性固形物含量 [J]. *光谱学与光谱分析*, 2015, 35(8): 2154-2158.
- [9] ZHANG B H, HUANG W Q, LI J B, et al. Principles, developments and applications of computer vision for external quality inspection of fruits and vegetables: A review[J]. *Food Res Int*, 2014, 62(62): 326-343.
- [10] 梁琨, 刘全祥, 潘磊庆, 等. 基于高光谱和 CARS-IRIV 算法的‘库尔勒香梨’可溶性固形物含量检测 [J]. *南京农业大学学报*, 2018, 41(4): 760-766.
- [11] 李晓丽, 魏玉震, 徐劼, 等. 基于高光谱成像的茶叶中 EGCG 分布可视化 [J]. *农业工程学报*, 2018, 34(7): 180-186.
- [12] 孙俊, 金夏明, 毛罕平, 等. 基于高光谱图像光谱与纹理信息的生菜氮素含量检测 [J]. *农业工程学报*, 2014, 30(10): 167-173.
- [13] 于慧春, 王润博, 殷勇, 等. 基于高光谱图像光谱与纹理信息的枸杞多糖及总糖含量检测 [J]. *核农学报*, 2018, 32(3): 523-531.
- [14] 何勇, 彭继宇, 刘飞, 等. 基于光谱和成像技术的作物养分生理信息快速检测研究进展 [J]. *农业工程学报*, 2015, 31(3): 174-189.
- [15] KAMRUZZAMAN M, SUN D W, EIMASRY G, et al. Fast detection and visualization of minced lamb meat adulteration using NIR hyperspectral imaging and multivariate image analysis[J]. *Talanta*, 2013(103): 130-136.
- [16] YE S F, WANG D, MIN S G. Successive projections algorithm combined with uninformative variable elimination for spectral variable selection[J]. *Chemometrics Intell Lab Syst*, 2008, 91(2): 194-199.
- [17] 尼珍, 胡昌勤, 冯芳. 近红外光谱分析中光谱预处理方法的作用及其发展 [J]. *药物分析杂志*, 2008, 28(5): 824-829.
- [18] 程长阔, 杨鹏程, 李燕, 等. 光谱预处理方法对硝酸盐测量模型影响实验研究及软件设计 [J]. *海洋技术学报*, 2015, 34(2): 69-73.
- [19] 张丽, 彭小列, 张建锋, 等. 杏鲍菇多糖的提取及其抑菌作用 [J]. *贵州农业科学*, 2010, 38(9): 90-92.
- [20] 王瑞君. 枸杞多糖对德氏乳酸杆菌在体外生长与保存活力的影响 [J]. *河北科技大学学报*, 2012, 33(2): 115-118.
- [21] 于雷, 洪永胜, 耿雷, 等. 基于偏最小二乘回归的土壤有机质含量高光谱估算 [J]. *农业工程学报*, 2015, 31(14): 103-109.
- [22] 潘国锋. 基于 K-S 算法的水质硝酸盐含量光谱检测方法研究 [J]. *光谱实验室*, 2011, 28(5): 2700-2704.
- [23] 何嘉琳, 乔春燕, 李冬冬, 等. 可见-近红外高光谱成像技术对灵武长枣 V_C 含量的无损检测方法 [J]. *食品科学*, 2018, 39(6): 194-199.

【责任编辑 李晓卉, 庄 延】