

文章编号 1004-924X(2015)02-0349-07

## 近红外高光谱成像技术快速鉴别国产咖啡豆品种

鲍一丹, 陈 纳, 何 勇\*, 刘 飞, 张 初, 孔汶汶  
(浙江大学 生物系统工程与食品科学学院, 浙江 杭州 310058)

**摘要:** 结合近红外高光谱成像技术和不同的判别分析模型对 4 种国产咖啡豆品种进行了快速无损判别。通过高光谱成像仪提取 874~1 734 nm 波段内的光谱数据, 去除首尾噪声波段后, 分别基于 925~1 680 nm 波段的全谱波段和通过连续投影算法(SPA)选择的特征波长, 建立了偏最小二乘判别分析(PLS-DA)、随机森林(RF)、K 最邻近算法(KNN)、支持向量机(SVM)模型和极限学习机(ELM)5 种判别分析模型。基于上述判别模型对咖啡豆品种进行鉴别; 然后通过准确率、命中率和否定率 3 个参数对鉴别结果进行了评价。实验显示, 基于全谱和特征波段建立的模型均取得了较好的判别效果, 其中 ELM 模型效果均为最优, 每个品种建模集和预测集的准确率、命中率和否定率均在 93.5% 以上。研究结果表明, 基于近红外高光谱成像技术结合模型判别分析方法可以实现对国产咖啡豆品种的识别, 特征波长的选择减少了变量数, 但判别效果与全谱相当。

**关键词:** 近红外高光谱成像; 咖啡豆; 无损判别; 判别分析模型; 极限学习机

**中图分类号:** O433.4 **文献标识码:** A **doi:** 10.3788/OPE.20152302.0349

### Rapid identification of coffee bean variety by near infrared hyperspectral imaging technology

BAO Yi-dan, CHEN Na, HE Yong\*, LIU Fei, ZHANG Chu, KONG Wen-wen

(College of Biosystem Engineering and Food Science, Zhejiang University, Hangzhou 310058, China)

\* Corresponding author, E-mail: yhe@zju.edu.cn

**Abstract:** Four different Chinese domestic coffee beans were identified rapidly by combining near infrared hyperspectral imaging technique and five kinds of discriminant models. A near-infrared hyperspectral imaging system covering the spectral range of 874–1 734 nm was set up to capture hyperspectral images of coffee bean samples. The head and end of the spectra with obvious noises were removed, and the spectral data in the range of 925–1 680 nm were extracted to establish discriminant models in the experiment. The sensitive wavelengths were selected from the full spectra by Successive Projections Algorithm (SPA). Five discriminant methods, including Partial Least Square-discriminant Analysis (PLS-DA), Random Forest (RF), K-nearest Neighbor algorithm (KNN), Support Vector Machine (SVM) and Extreme Learning Machine (ELM) were applied to the establishment of discriminant models based on the full spectra and the selected sensitive wavelength variables. The properties of the models were compared and valued by three parameters, sensitivity, precision and specificity.

收稿日期: 2014-07-10; 修订日期: 2014-08-21.

基金项目: 国家 863 高技术研究发展计划资助项目(No. 2012AA101903); 中央高校基本科研业务费专项资金资助项目(No. 2014FZA6005)

Among all discriminant models, the ELM models based on the full spectra and the selected sensitive wavelength variables show the best identification results, respectively. For each coffee bean cultivar, the sensitivity, precision and specificity of ELM models based on full spectra and the sensitive wavelengths are all over 93.5% in both the calibration set and the prediction set. It concludes that Chinese domestic coffee beans could be identified by near-infrared hyperspectral imaging combined with discriminant models rapidly. Selecting the sensitive wavelengths reduces variables, but the identification effect is the same as that of the full spectra.

**Key words:** near-infrared hyperspectral imaging; coffee bean; non-destructive identification; discriminant analysis model; extreme learning machine

## 1 引言

咖啡豆是制作咖啡的原材料,由咖啡生豆烘焙而成。不同品种的咖啡豆,其外观色泽、气味以及内部化学成分存在很大差异<sup>[1]</sup>。咖啡豆种类按颗粒大小主要分为小粒咖啡(Arabica coffee)、中粒咖啡(Robusta coffee)和大粒咖啡(Liberica coffee)。我国主要种植小粒咖啡和中粒咖啡,分布于云南省和海南省。咖啡豆品种对咖啡品质的影响很大。一般地,人们通过观察豆的外观、色泽等,或利用化学方法来检测咖啡豆品质的变化。曾凡逵等以云南 Arabica 咖啡、兴隆 Robusta 咖啡和越南 Robusta 咖啡为原料,通过气相色谱-质谱联用仪(Gas Chromatograph-Mass Spectrometer-computer, GC/MS)分析了生咖啡及不同烘焙程度咖啡的脂肪酸含量,根据总脂肪和脂肪酸含量的差异来区分咖啡<sup>[2]</sup>。Semmelroch 等运用稳定同位素稀释法分析了 14 种由 Arabica 咖啡豆和 Robusta 咖啡豆制成的咖啡中的芳香提取物,以鉴别咖啡豆品种<sup>[3]</sup>。这些方法存在评价标准不统一,检测时间长且成本高等缺点<sup>[4]</sup>。因此,需要研究一种能够快速鉴别咖啡豆品种的无损检测方法。

近红外光谱分析技术是一种速度快、操作简便、成本较低的无损检测方法,广泛用于分类判别和品质评价研究<sup>[5-6]</sup>,在咖啡分类判别中也有应用。王艳艳等利用可见-近红外光谱技术对市场上雀巢、麦斯威尔和帕拉 3 种品牌的咖啡品种进行了鉴别<sup>[7]</sup>。Ivanira 等结合反相高效液相色谱法和近红外光谱法,对传统的咖啡豆品种波旁和其他 3 种转基因咖啡豆品种进行了区分<sup>[8]</sup>。Esteban-Diez 等采用近红外光谱技术实现了对 Ara-

bica、Robusta 咖啡以及二者混合品的鉴别<sup>[9-10]</sup>。Downey 等采用近红外光谱与中红外光谱实现了对 Arabica 和 Robusta 咖啡的区分<sup>[11]</sup>。Santos 等实现了对巴西咖啡的近红外光谱分析<sup>[12]</sup>。Myles 等应用近红外光谱技术实现了对 Arabica 和 Robusta 咖啡的区分<sup>[13]</sup>。

近年来,高光谱成像技术发展迅速,该技术将光谱检测和图像检测的优势合为一体,再结合化学计量方法有效而全面地分析光谱信息,不仅降低了数据计算量,而且提高了模型判别的准确率,目前已广泛用于农业<sup>[14]</sup>、食品<sup>[15]</sup>等领域。不过国内外尚未有人采用近红外高光谱成像技术对咖啡豆进行分类,因此本文研究了采用近红外高光谱成像技术对国产咖啡豆品种鉴别的可行性。

## 2 材料与方 法

### 2.1 材料

实验所用的咖啡豆主要采用我国种植的中粒咖啡和小粒咖啡。中粒咖啡豆采用海南的福山咖啡和兴隆咖啡,小粒咖啡豆采用云南的卡蒂莫小粒咖啡豆和铁毕卡小粒咖啡豆,4 种咖啡均采用中度烘焙。从所收集的 4 种样本原料中分别挑选出 600 粒颗粒完整的咖啡豆作为试验样本。将每粒咖啡豆擦拭干净,依次排列在平板上,并用高光谱成像仪进行扫描。

### 2.2 高光谱图像采集

本次实验在黑暗密闭的暗箱中进行,采用芬兰 Spectral Imaging 公司生产的 N17E-QE 光谱仪和型号为 OLES22 的 C-mount 成像镜头采集咖啡豆的高光谱图像信息,图像分辨率为 320 pixel×256 pixel,采用美国 Dolan Jenner Industries 公司生产的 Fiber-LiteDC950 为线光源。在

仪器校正后,将试验样本放在由台湾五铃光学股份有限公司生产的电控位移平台上,由配套计算机控制整个实验过程,并保存实验数据。

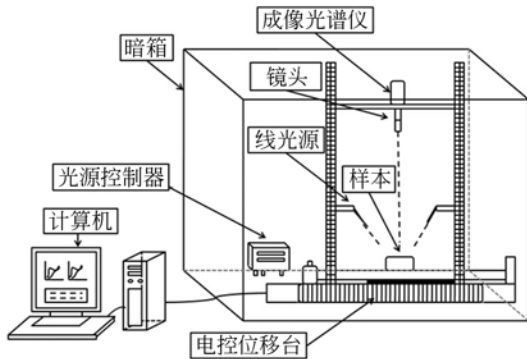


图 1 高光谱成像系统

Fig. 1 Hyper-spectral imaging system

对咖啡豆的光谱信息采集前,必须调节高光谱成像仪的相机曝光时间、移动控制平台的移动速度和物镜高度,以避免采集的图像变形失真或模糊不清。经过反复尝试,物镜高度设置为 17.9 cm,曝光时间设置为 3 500  $\mu$ s,平台移动速度设置为 13.8 mm/s。

通过高光谱成像系统的采集软件采集咖啡豆的光谱图像信息,然后对采集的光谱图像信息进行图像校正。图像校正公式为:

$$I = \frac{R - D}{W - D}, \quad (1)$$

其中: $I$ 是校正过的图像, $R$ 是原始图像, $W$ 是白板校正图像, $D$ 是黑板校正图像。图像校正后,运用 ENVI4.6 软件提取高光谱图像中样本的光谱信息。选取整颗咖啡豆为感兴趣区域(Region of Interest, ROI),将 ROI 内所有光谱信息的平均值作为对应样本的光谱值。

### 2.3 化学计量学方法

光谱数据分析是在全谱波长和通过连续投影算法(Successive Projections Algorithm, SPA)选择特征波长后的波段范围内,分别建立偏最小二乘判别分析(Partial Least Square-Discriminant Analysis, PLS-DA)<sup>[16-17]</sup>、RF 模型(Random Forest, RF)<sup>[18]</sup>、K 最邻近算法(K-Nearest Neighbor Algorithm, KNN)<sup>[19]</sup>、支持向量机(Support Vector Machine, SVM)<sup>[20-22]</sup>和极限学习机(Extreme Learning Machine, ELM)<sup>[23]</sup>等 5 种判别分析模型。其中,PLS-DA 算法在 Unscrambler<sup>®</sup> 10.1 上

实现,而其它分类算法在 Matlab R2009b 软件上进行。

判别分析模型以准确率、命中率、否定率为评价指标<sup>[24]</sup>。其中,准确率  $S_n$  表示通过判别模型将  $a$  类样本正确归为  $a$  类的概率;命中率  $S_p$  表示判别模型只将  $a$  类样本归为  $a$  类的能力;否定率  $S_p$  表示判别模型将不属于  $a$  类的样本归属为不属于  $a$  类的能力。

## 3 实验结果与分析

### 3.1 咖啡豆的光谱

本次实验中,每个品种分别选取 600 颗咖啡豆,共提取 2 400 颗咖啡豆的 256 个光谱波段信息,去除首尾明显噪声,选择第 16~240 波段进行分析,获取 925~1 680 nm 波段的光谱信息。将样本按照 2:1 的比例用 Kennard-Stone(K-S)算法划分为建模集和预测集,不同品种咖啡豆的赋值以及建模集和预测集样本的划分如表 1 所示。图 2 所示为建模集中每个品种 400 颗咖啡豆的平均光谱曲线,4 个品种的咖啡豆光谱曲线的趋势一致,差异性不大。

表 1 咖啡豆样本的建模集和预测集区分以及品种赋值  
Tab. 1 Division of coffee bean samples and sample coding

样本种类	类别 赋值	样本 总数	建模集	预测集
云南铁皮卡小粒咖啡豆	1	600	400	200
云南卡蒂莫小粒咖啡豆	2	600	400	200
海南福山中粒咖啡豆	3	600	400	200
海南兴隆中粒咖啡豆	4	600	400	200

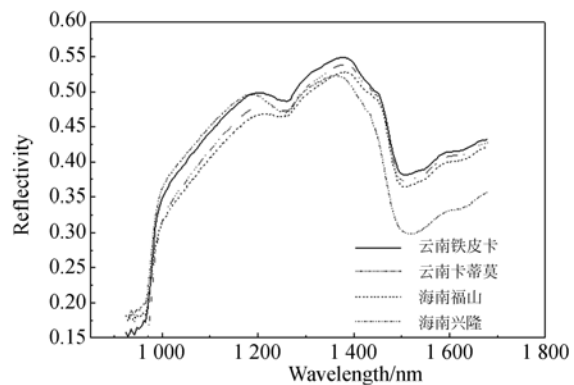


图 2 咖啡豆建模集样本光谱图

Fig. 2 Reflectance spectra of coffee bean modeling set samples

### 3.2 基于全谱数据的咖啡豆品种识别模型

为建立不同品种的咖啡豆光谱曲线与咖啡豆品种之间的关系,分别建立了基于全谱数据的 PLS-DA、KNN、RF、SVM 和 ELM 模型,并运用准确率  $S_n$ 、命中率  $P_r$  和否定率  $S_p$  3 个参数分析判别结果,具体数据如表 2 所示。

表 2 基于全谱数据建立识别模型的判别结果

Tab. 2 Identified results of different classification models using raw full spectra preprocessed (%)

	建模集			预测集			
	$S_n$	$P_r$	$S_p$	$S_n$	$P_r$	$S_p$	
PLS-DA	1	88.5	97.79	96.83	92	98.92	98.33
	2	86.5	71.93	86.08	84	71.49	87.50
	3	76.25	79.02	90.50	74	77.89	91.67
	4	83.75	99.11	99.42	90	99.45	99.83
KNN	1	94	97.79	84.42	95.5	60.98	87.83
	2	56	71.93	94.17	61.5	82	95.5
	3	68.25	79.02	94	83	89.25	96.67
	4	97.25	99.11	99.42	99.5	99.5	99.83
RF	1	100	100	100	97	89.40	96.17
	2	100	100	100	75	91.46	97.67
	3	100	100	100	93	85.32	94.5
	4	100	100	100	100	100	100
SVM	1	99.25	99.25	99.75	100	89.40	100
	2	94.5	95.70	98.58	92	97.35	99.17
	3	96.25	95.06	98.33	97.5	92.42	97.33
	4	100	100	100	100	100	100
ELM	1	94	100	100	99	100	100
	2	99.75	99.25	99.75	97	99.49	99.83
	3	99.25	99.75	99.92	99.5	97.07	99.33
	4	100	100	100	100	100	100

PLS-DA 模型的隐含变量数为 9, SVM 采用网格搜索法确定最佳模型参数 ( $c = 256, g = 0.0359$ ), ELM 采用逐步选择法确定的模型隐含层节点数为 396。5 种判别模型中, ELM 判别分析模型的识别效果最好,  $S_n, P_r$  和  $S_p$  普遍高于 97%, 而品种 4 的  $S_n, P_r$  和  $S_p$  均达到了 100%, 表示 ELM 模型能准确分类咖啡豆, 且准确率可达 100%。判别效果较好的是 SVM 模型, 所有品种的  $S_n, P_r$  和  $S_p$  均在 89.4% 以上。RF 模型、PLS-DA 模型和 KNN 模型的预测效果依次下降, 它们对品种 1, 4 的预测效果优于对品种 2, 3 的预测。KNN 模型下, 品种 2 的预测集  $S_n$  最低, 只有 61.5%, 基本不能判别品种 2。从整体上看,

基于全谱数据的各种识别模型对品种 2, 3 的识别准确率低于品种 1, 4 的识别准确率, 这可能与品种自身的内外部品质与其它品种相近有关。

### 3.3 基于特征波段选择建立咖啡豆品种识别模型

为简化模型, 减少计算量, 选择 SPA 对原始光谱数据进行特征波长选择。SPA 是一种常用的变量选择算法, 用于在样本数量大的前提下筛选光谱的特征波长<sup>[25]</sup>。将 SPA 选择特征波长个数设置为 5~30, 得出最优的波长个数为 14 个 (如图 3 所示, 分别是 941, 952, 972, 1 032, 1 099, 1 106, 1 214, 1 352, 1 386, 1 413, 1 511, 1 632, 1 656 和 1 662 nm)。1 214 nm 与 C-H 二级倍频有关, 1 352 nm 与 C-H 组合键的一级倍频有关, 1 413 nm 与 O-H 的一级倍频有关, 1 662 nm 与 C-H 的一级倍频有关<sup>[26]</sup>。与全波段光谱分析 225 个波长数相比, 由 SPA 选出的波长数缩减为原来的 6.22%。由图 3 可知, 利用 SPA 算法选择的特征波段主要集中在 940~1 670 nm, 光谱变量的减少降低了校正模型的复杂度。

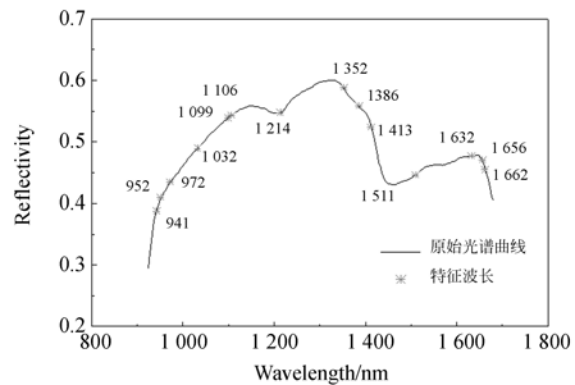


图 3 咖啡豆经过 SPA 特征波长选择后的样本光谱图  
Fig. 3 Reflectance spectra of coffee beans after SPA characteristics of wavelength selection

基于特征波长建立的判别分析模型的判别效果如表 3 所示。PLS-DA 模型的隐含变量数为 7, SVM 参数为  $c = 256, g = 1$ , ELM 模型隐含层节点数为 421。5 种判别模型中, ELM 模型和 SVM 模型的识别效果最好, 4 个品种的准确率  $S_n, P_r$  和  $S_p$  均在 90% 以上。RF 模型的预测效果稍差, 体现为对品种 2 的预测效果差, 其余品种的  $S_n, P_r$  和  $S_p$  均超过 87%。PLS-DA 模型和 KNN 模型的预测效果较差。

### 3.4 基于全谱与特征波段的判别分析模型比较

基于全谱与基于特征波段的 ELM 判别模型

对品种1和品种4的识别效果最好,优于对品种2和品种3的识别。在5种模型中,基于全谱建立的ELM模型的识别率较高。而基于特征波段的ELM模型和基于全谱波段的识别效果相似。其次,判别效果较好的是SVM模型。基于全谱与基于特征波段的SVM模型对品种4的识别效果最明显,而基于特征波段的SVM模型效果对品种1的识别更明显。基于全谱波段和基于特征波段的RF模型对4种咖啡豆的判别效果相近,对品种1,4的预测效果较好,对品种2,3的预测效果稍逊,但仍能识别品种2和品种3。基于特征波段的KNN识别模型对品种4的识别效果最明显,但对品种1,品种2和品种3的识别效果与基于全谱波段的识别效果相近,均能识别。针对品种3的识别,基于特征波段的PLS-DA模型判别效果优于基于全谱的PLS-DA模型。对品种1、品种2和品种4的识别,基于全谱的PLS-DA模型的识别效果优于基于特征波段的PLS-DA模型。通过比较,ELM模型取得了最优的识别效果,基于全谱和特征波段,ELM模型对咖啡豆品种的 $S_n$ ,  $P_r$ 和 $S_p$ 均在93.5%以上,表明利用高光谱图像技术结合判别分析模型能识别咖啡豆种类。对比可知,本文方法与采用近红外光谱技术的鉴别结果相近<sup>[8-13]</sup>。

基于全谱与基于特征波段的判别分析模型的判别效果相似,表明特征波段的选择能有效地代表原始样本的信息,减少了冗余变量,降低了模型的复杂程度和计算量。ELM模型与SVM模型的判别效果较优,这是由于二者对线性与非线性数据都具有较好的处理效果和较强的泛化能力。

基于全谱和基于特征波段建立的模型,对品种2和品种3的识别效果比对品种1和品种4的差,可能与实验样品特性有关。由于品种1和品种4的产地与品种都不同,品种1为云南小粒种,品种4为海南中粒种,其内外品质均有差异,因此它们的识别效果较好。品种1和品种2均属于小粒种,两者相似度大,分辨难度大。品种2咖啡是品种1咖啡基因的改良品种,与品种1咖啡具有相似基因,因此品种1,2间难以分辨。品种3海南福山咖啡和品种4海南兴隆咖啡都属于中粒咖啡,种植环境相似,其内外部品质有一定的相似性,因此利用光谱信息进行区分有一定难度。

表3 基于SPA特征波段选择建立识别模型的判别结果

Tab. 3 Results of different classification models using SPA spectra preprocessed (%)

	建模集			预测集		
	$S_n$	$P_r$	$S_p$	$S_n$	$P_r$	$S_p$
PLS-DA1	86.75	94.29	98.25	86	97.18	99.17
	2 84.75	70.48	88.17	80	71.43	89.33
	3 76.5	84.53	95.33	80	80.81	93.67
	4 90.5	98.91	99.67	86	97.97	99.33
KNN	1 71.29	71.29	87.42	95	74.8	89.33
	2 82.53	82.53	95.17	67	79.76	94.33
	3 85.17	85.17	95.75	79	88.76	96.67
	4 99.5	99.5	99.83	100	100	100
RF	1 100	100	100	97	87.39	95.33
	2 100	100	100	76.5	87.43	96.33
	3 100	100	100	88.5	87.19	95.67
	4 100	100	100	100	100	100
SVM	1 98.75	98.01	99.33	100	99	99.67
	2 92.5	95.61	98.58	94	96.91	99
	3 96.75	94.39	98.08	97	90.65	96.67
	4 100	100	100	100	100	100
ELM	1 100	99.75	99.92	99	99.5	99.83
	2 98	98.99	99.67	93.5	96.89	99
	3 99	98.26	99.42	96.5	93.69	97.83
	4 100	100	100	100	100	100

## 4 结 论

本文利用高光谱成像技术结合建立不同的判别分析模型,实现了对中国产咖啡品种的鉴别。通过高光谱成像仪提取874~1734 nm波段的256个波段光谱信息,并选取925~1680 nm波段的光谱数据进行分析。基于原始光谱采用SPA算法选择了14个特征波长,分别基于全谱和特征波长建立了PLS-DA、KNN、RF、SVM和ELM5种判别分析模型,取得了比较理想的识别效果。通过比较分析,基于全谱与特征波段的ELM模型的判别效果最好,建模集和预测集的准确率( $S_n$ ),命中率( $P_r$ )和否定率( $S_p$ )3个参数值均在93.5%以上。SVM、RF、KNN以及PLS-DA模型的效果依次变差。尽管判别分析的效果较好,但由于咖啡豆的品种来源、产地等因素会对判别分析造成影响,可能造成错判。研究结果表明,基于近红外高光谱成像技术结合判别分析的方法可以实现对国产咖啡豆品种的识

别,特征波长的选择大大减少了变量数,且判别效果与全谱相当。

#### 参考文献:

- [1] 胡荣锁,初众,谷凤林,等. 海南主要地域生咖啡豆挥发性化学成分对比研究[J]. 光谱学与光谱分析, 2013,33(2):548-553.  
HU R S, CHU ZH, GU F L, *et al.*. Comparison of green coffee beans volatiles chemical composition of Hainan main area[J]. *Spectroscopy and Spectral Analysis*, 2013, 33(2): 548-553. (in Chinese)
- [2] 曾凡远,刘爱琴,谭乐和,等. 3个不同产地咖啡脂肪酸组成比较[J]. 热带作物学报, 2011, 32(8): 1460-1463.  
ZENG F K, LIU A Q, TAN L H, *et al.*. Fatty acid composition of coffee oil from three different origins [J]. *Chinese Journal of Tropical Crops*, 2011, 32(8):1460-1463. (in Chinese)
- [3] SEMMELROCH P, LASKAWY G, BLANK I, *et al.*. Determination of potent odourants in roasted coffee by stable isotope dilution assays [J]. *Flavour and Fragrance Journal*, 1995, 10: 1-7.
- [4] 詹家芬,陆舍铭,曲国福,等. 老挝咖啡的挥发性和半挥发性成分提取分析[J]. 食品研究与开发, 2008,29(2):125-129.  
ZHAN J F, LU SH M, QU G F, *et al.*. Analysis on volatile and semi-volatile components of Laos' coffee [J]. *Food Research and Development*, 2008, 29(2):125-129. (in Chinese)
- [5] 刘燕德,周延睿,彭彦颖. 基于近红外漫反射光谱检测鸡蛋品质[J]. 光学精密工程, 2013,21(1):40-45.  
LIU Y D, ZHOU Y R, PENG Y Y. Detection of egg quality by near infrared diffuse reflectance spectroscopy[J]. *Opt. Precision Eng.*, 2013, 21(1): 40-45. (in Chinese)
- [6] 杜敏,巩颖,林兆洲,等. 样品表面近红外光谱结合多类支持向量机快速鉴别枸杞子产地[J]. 光谱学与光谱分析, 2013,32(5):1211-1214.  
DU M, GONG Y, LIN ZH ZH, *et al.*. Rapid identification of wolfberry fruit of different geographic regions with sample surface near infrared spectra combined with multi-class SVM [J]. *Spectroscopy and Spectral Analysis*, 2013, 32(5): 1211-1214. (in Chinese)
- [7] 王艳艳,何勇,邵永妮,等. 基于可见-近红外光谱的咖啡品牌鉴别研究[J]. 光谱学与光谱分析, 2007, 27(4):702-706.  
WANG Y Y, HE Y, SH Y N, *et al.*. Discrimination among different brands of coffee by using vis-near infrared spectra[J]. *Spectroscopy and Spectral Analysis*, 2007, 27(4):702-706. (in Chinese)
- [8] IVANIRA M, IEDA S S. Chemometric discrimination of genetically modified *coffea arabica* cultivars using spectroscopic and chromatographic fingerprints [J]. *Talanta*, 2013, 107: 416-422.
- [9] ESTEBAN-DIEZ I, GONZALEZ-SAIJ J M, PIZARRO C. An evaluation of orthogonal signal correction methods for the characterisation of arabica and robusta coffee varieties by NIRS [J]. *Analytica Chimica Acta*, 2004, 514(1): 57-67.
- [10] ESTEBAN-DIEZ I, GONZALEZ-SAIJ J M, SAENZ-GONZALEZ C, *et al.*. Coffee varietal differentiation based on near infrared spectroscopy [J]. *Talanta*, 2007, 71(1):221-229.
- [11] DOWNEY G, BRIANDET R, WILSON R H, *et al.*. Near-and mid-infrared spectroscopies in food authentication; coffee varietal identification [J]. *Journal of Agricultural and Food Chemistry*, 1997, 45(11): 4357-4361.
- [12] SANTOS K M, MOURA M F V, AZEVEDO F G, *et al.*. Classification of Brazilian coffee using near-infrared spectroscopy and multivariate calibration [J]. *Analytical Letters*, 2012, 45(7): 774-781.
- [13] MYLES A J, ZIMMERMAN T A, BROWN S D. Transfer of multivariate classification models between laboratory and process near-infrared spectrometers for the discrimination of green Arabica and Robusta coffee beans [J]. *Applied Spectroscopy*, 2006, 60(10): 1198-1203.
- [14] KONG W, LIU F, ZHANG C, *et al.*. Fast detection of peroxidase (POD) activity in tomato leaves which infected with *Botrytis cinerea* using hyperspectral imaging [J]. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2014, 118: 498-502.
- [15] XIONG Z J, SUM D W, ZENG X A, *et al.*. Recent developments of hyperspectral imaging systems and their applications in detecting quality attributes of red meats; a review [J]. *Journal of Food Engineering*, 2014, 132: 1-13.
- [16] 周扬,戴曙光,葛丁飞. 近红外光谱稀疏分量分析检测柴油品质参数[J]. 光学精密工程, 2014, 22

- (2):296-303.
- ZHOU Y, DAI SH G, GE D F. Detection of diesel quality parameters by near infrared spectroscopy with sparse component analysis[J]. *Optics and Precision Engineering*, 2014, 22(2): 296-303. (in Chinese)
- [17] ZHU F L, ZHANG H L, SHAO Y N, *et al.*. Mapping of fat and moisture distribution in atlantic salmon using near-infrared hyperspectral imaging [J]. *Food and Bioprocess Technology*, 2014, 7: 1208-1214.
- [18] LEE S, CHOI H, CHA K, *et al.*. Random forest as a potential multivariate method for near-infrared (NIR) spectroscopic analysis of complex mixture samples gasoline and naphtha [J]. *Microchemical Journal*, 2013, 110: 739-748.
- [19] ZHANG L G, ZHANG X, NI L J, *et al.*. Rapid identification of adulterated cow milk by non-linear pattern recognition methods based on near infrared spectroscopy [J]. *Food Chemistry*, 2014, 145:342-348.
- [20] ALVES J C L, POPPI R J. Biodiesel content determination in diesel fuel blends using near infrared (NIR) spectroscopy and support vector machines (SVM) [J]. *Talanta*, 2013, 104: 155-161.
- [21] LIU F, HE Y. Discrimination of producing areas of auricularia auricula using visible/near infrared spectroscopy [J]. *Food and Bioprocess Technology*, 2011, 4(3):387-394.
- [22] 刘燕德,周延睿,潘圆媛. 基于最小二乘支持向量的辣椒可溶性固形物和维生素 C 含量近红外光谱检测[J]. *光学精密工程*, 2014, 22(2):281-288.
- LIU Y D, ZHOU Y R, PAN Y Y. Determination of soluble solid contents and vitamin C of fresh peppers based on NIR spectrometry and least square support vector machines [J]. *Optics and Precision Engineering*, 2014, 22(2): 281-288. (in Chinese)
- [23] KONG W W, ZHANG C, LIU F, *et al.*. Irradiation dose detection of irradiated milk powder using visible and near-infrared spectroscopy and chemometrics [J]. *Journal of Dairy Science*, 2013, 96(8): 4921-4927.
- [24] 褚小立. 化学计量学方法与分子光谱分析技术 [M]. 北京:化学工业出版社,2011.
- ZHU X L. *Molecular Spectroscopy Analytical Technology Combined with Chemometrics and Its Applications* [M]. Beijing: Chemical Industry Press, 2011. (in Chinese)
- [25] OUYANG A, LIU J. Classification and determination of alcohol in gasoline using NIR spectroscopy and the successive projections algorithm for variable selection [J]. *Measurement Science and Technology*, 2013, 24(2): 025502-025507.
- [26] RIBEIRO J S, FERREIRA M M C, SALVA T J G. Chemometric models for the quantitative descriptive sensory analysis of Arabica coffee beverages using near infrared spectroscopy[J]. *Talanta*, 2011, 83: 1352-1358.

## 作者简介:



鲍一丹(1963—),女,浙江杭州人,博士,副教授,硕士生导师,主要研究方向为精细农业领域内现代检测技术和自动控制、虚拟仪器技术等研究。E-mail: ydbao@zju.edu.cn



陈 纳(1989—),女,浙江慈溪人,硕士研究生,主要从事近红外光谱无损检测方面的研究。E-mail: chenna123ok@126.com

## 通讯作者:



何 勇(1963—),男,浙江慈溪人,博士,教授,博士生导师,浙江大学求是特聘教授,主要从事数字农业与精细农业、农业物联网技术、智能农业装备、光谱与高光谱成像技术、管理信息系统与专家系统等方面的科研和教学工作。E-mail: yhe@zju.edu.cn