

气体传感器鉴别花椒产地研究

庞涛¹, 杨霄¹, 陈晓燕^{2,3*}, 陶怀亮¹, 李蒙良¹

(1. 四川农业大学机电学院, 雅安 625000; 2. 四川农业大学信息工程学院, 雅安 625000;
3. 四川农业大学农业信息工程四川省重点实验室, 雅安 625000)

摘要: 目前花椒产地鉴别基本以感官评定为主, 缺乏客观性, 在实施应用时难以做到量化和标准化, 难以做出判断。因此设计研发一种快速鉴别花椒的智能装置。该装置以气体传感器阵列为核心, 能够独立对花椒气味信息进行检测和鉴别, 区分不同产地的同类花椒。利用主成分分析和 Wilks Λ 统计分析对检测数据进行处理。提取主成分 5 个, 累积贡献率为 94.41%, 其对应 Fisher 判别模型训练集平均准确率达到 88.6%, 验证集 90%, Wilks Λ 统计分析最终选取 8 个变量, 其对应判别 Fisher 模型训练集平均准确率 91.82%, 验证集 95%。对 Wilks Λ 统计所选取变量建立细分类交叉验证的 Fisher 判别模型, 平均正确率达到 97.27%, 将模型移植到采集装置, 完成智能花椒品种鉴别装置。该方法是一种简便高效的花椒品种鉴别方法, 可为今后进一步研究花椒产地、分级提供检测仪器和理论依据。

关键词: 传感器; 农作物; 气体监测器; 花椒产地鉴别; Fisher 判别

doi: 10.11975/j.issn.1002-6819.2019.18.032

中图分类号: S-3

文献标志码: A

文章编号: 1002-6819(2019)-18-0267-06

庞涛, 杨霄, 陈晓燕, 陶怀亮, 李蒙良. 气体传感器鉴别花椒产地研究[J]. 农业工程学报, 2019, 35(18): 267—272. doi: 10.11975/j.issn.1002-6819.2019.18.032 http://www.tcsae.org

Pang Tao, Yang Xiao, Chen Xiaoyan, Tao Huailiang, Li Mengliang. Identification of *Zanthoxylum bungeanum* origin based on gas sensor[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2019, 35(18): 267—272. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.2019.18.032 http://www.tcsae.org

0 引言

花椒是中国的重要经济作物, 中国西部地区如四川、陕西、甘肃、青海等地花椒种植的面积很大, 成为重要的经济作物甚至是地方经济的支柱产业^[1]。市面上花椒品种众多, 外形、色泽、风味各不相同, 这也造成了不同品种的花椒价格差异很大, 中国约有 39 种花椒, 14 个变种^[2], 大部分花椒品种仍处于野生状态, 人工栽培的花椒主要分为青花椒和红花椒, 不同产地的花椒质量不同。但随着花椒深度开发和利用得到越来越多的重视, 市场需求快速增长, 随之而来的是市场上以劣充优、以陈冒新、掺假掺杂等现象层出不穷。一方面, 消费者开始更多地关心他们所消费的农产品的来源与真实性, 另一方面, 企业与花椒原产地渴望寻求保护自身品牌的有效方法, 因此研究开发出一种简单、快速、无损的花椒产地鉴别检测方法, 具有重要的现实意义。

现有花椒品种鉴别的常用方法有感官分析技术、气相色谱法、与气质联用技术 (GC-MS) 等。感官辨识对于专业人员的要求较高, 需要长期的经验, 而这些专业人员很难做到将不同地区的花椒完全辨认。而如吴习宇等采用近红外光谱技术鉴别花椒产地^[3], 吴莉莉等采用机

器视觉对花椒品种进行鉴别^[4]。此外, 也有众多研究发现不同种或不同品系的花椒间化学成分也有较大差异^[5]。这一类的非感官的鉴别手段, 测试周期较长, 运行成本高, 在操作上不利于推广。

气体传感器是一种模仿生物嗅觉的气体检测系统^[6]。而气体传感器的核心部件为由多种气敏传感器组成的气敏传感器阵列, 其原理是将多个传感器感知到的时间或空间上互补或冗余的信息, 并在某种准则下进行综合与分析, 以获得单个或单类传感器无法获得的有价值的综合信息, 从而形成对观测对象客观的描述^[7-9]。鉴于气体传感器对待测气体样品的信息综合分析能力, 它已经在农业生产^[10-12]、生物医学^[13-15]、环境监测^[16-18]、食品检测^[19-21]等领域得到了广泛应用。

本文自行研发设计出一种基于气体传感器阵列的花椒信息采集装置, 用于 4 种花椒气味信息检测, 依据 Wilks Λ 统计量对气体传感器中的传感器阵列数据进行优化, 剔除冗余信息。将优化数据利用 Fisher 判别建立分类器, 结合花椒信息采集装置, 建立一种快速、简单、准确且便于推广的花椒产地鉴别方法, 为保证名品花椒、产地优选提出新的方法。

1 材料与方法

1.1 材料选取

试验花椒分为 4 个品种, 分别是四川汉源红花椒、陕西韩城红花椒 (大红袍)、四川汉源青花椒、云南鲁甸青花椒 (后文简称云南青花椒)。4 种花椒均取自原产

收稿日期: 2019-03-18 修订日期: 2019-04-26

基金项目: 四川省教育厅一般项目 (自然科学) 立项编号: 17ZB0333 基于高光谱图像技术的苹果品质无损检测方法研究

作者简介: 庞涛, 讲师, 主要从事农业信息检测。Email: 349380993@qq.com

*通信作者: 陈晓燕, 教授。主要从事农产品无损检测。

Email: chenxy@sicau.edu.cn

地种植园区，均为干燥花椒果皮。其中汉源青、红花椒和云南青花椒各取 60 份，陕西红花椒取 40 份，每份均 15 g，此 220 份花椒用作训练样本。各类花椒取样本各 20 份作为验证集。

1.2 试验设备

智能花椒品种鉴别装置。该装置系统结构图如图 1a 所示，作为整个设备的设计指导。图 1b 为实际设计时的气室结构设计剖图，从图中可以看出，气室，样品槽，传感器阵列，加热片，离心风机的组装结构，风机从 upper 端吸收被加热片加热的花椒气体，由下侧向四周的传感器阵列散发，使得各个传感器均匀接触到花椒气体，气

体再由外侧返回气室，达成气体循环。图 1c 为装置设计的外部结构，从外部只能看到气室和样品槽，无法看到气室内部的具体结构。图 1d 为拆下外壳的实物俯视图，俯视结构中可以看到本装置在设计时包含了气体流通通道，可以保证气体实现循环。图 1e 是在 1d 的基础上，拆除传感器阵列的保护外壳，传感器设置为倒置摆放，阵列通道内为线路连接，只将传感器探头露出，保证传感器传递线路不受到花椒气味的颗粒影响，确保实验安全进行。图 1f 为实物的侧放图，将侧面挡板取下，内部为搭建好的控制电路及外围电路等，挡板可拆卸保证了随时可以对装置程序，电路进行修改。

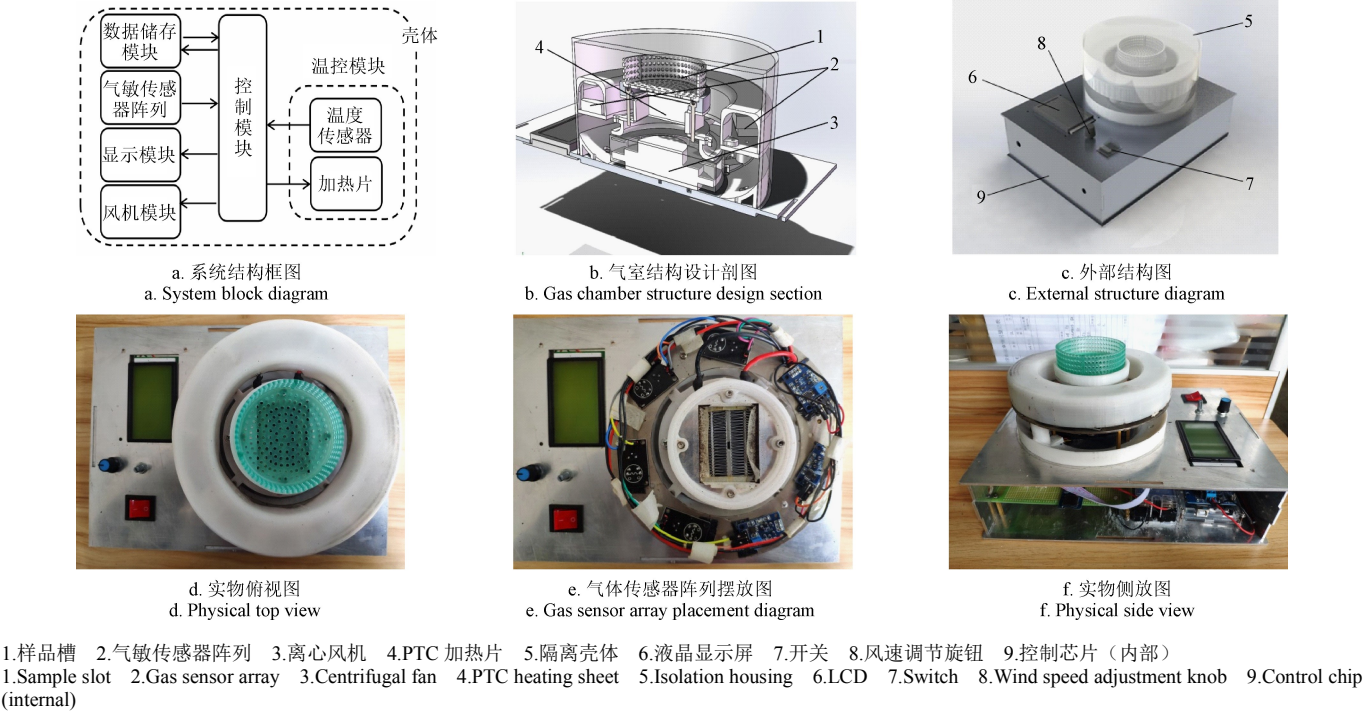


图 1 智能花椒品种鉴别装置图
Fig.1 Intelligent *Zanthoxylum bungeanum* variety identification device

本装置所采用的传感器选型时参照花椒气体组成成分进行选择，并对传感器本身的灵敏性，恢复性，稳定性进行筛选，最终所选的 7 个传感器分别为 MQ135、MQ5、MQ2、TGS2611、TGS2600、TGS2610 和 TGS2602。图 2 为本文所用气体传感器阵列实物图，表 1 为传感器所对应敏感响应特性。该装置有采集检测和鉴别种类 2 种工作模式，检测到的数据和鉴别结果会在显示屏上显示。

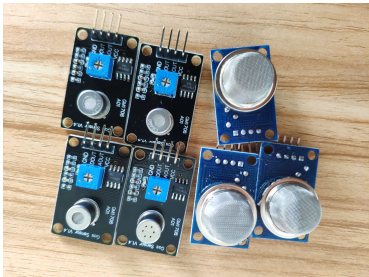


图 2 试验所用传感器
Fig.2 Sensor used in experiment

表 1 气体传感器响应特性		
Table 1 Gas sensor response characteristics		
传感器序号 Sensor serial number	传感器名称 Sensor name	主要敏感性 Primary sensitivity
1	MQ135	苯系、硫化物
2	MQ2	醇类
3	MQ5	丁烷、丙烷、甲烷
4	TGS2611	烷烃，甲烷
5	TGS2600	香草类气体
6	TGS2610	烷烃，丙烷，丁烷
7	TGS2602	醇、苯、醛

1.3 信息采集和选择

首先在采样前，将颗粒状活性炭置入样品槽中，进行清洗，直至传感器阵列的响应信号稳定，将此时传感器阵列的响应值作为基准值，清洗时间为 10 min。将花椒样本送入样品槽，开启气味采集工作模式，先对气味采集装置进行约 60 s 的预热，使传感器阵列处于正常工作状态，气室内的温度稳定在 26℃（±1℃）。进行数据采集，采样

时间约为 10 min，每个传感器分别采集 50 次数据，每次采样间隔 10 s，取 50 次采样数值的平均值（aver），最大值（max），最小值（min）作为一个样本的记录值，每组花椒样本包含 21 个数据参数，构成完整的数据带。

1.4 数据分析

1.4.1 主成分分析

主成分分析（principal component analysis, PCA），是一种通过正交变换将一组可能存在相关性的变量转换为一组线性不相关的变量，从而代表所有变量的统计方法，转换后的这组变量叫主成分^[22]。对测定花椒的 21 个参数进行主成分分析，选取方差累积贡献率达到 90% 以上的稳定水平的主成分，并确定其数量。

1.4.2 Wilks Λ 统计分析

Wilks Λ 统计量是常用的数据筛选方式，用于检验多个母体的判别效果和各个变量的判别能力^[23]。其实质是，样本组内离差平方和与样本总体离差平方和之比。其比值越小说明样本贡献越大。

对 2 种分析方法选定的成分分别建立 Fisher 判别分析，选择最优方式并将最优判别分析移植到鉴别装置中。

2 结果与分析

2.1 主成分提取与 Wilks Λ 统计筛选

对 220 个样本所测试花椒 21 个特征参数进行主成分分析，方差累积结果如图 3，主成分分析在第 5 个以后，累积方差逐增长渐趋于平稳，故提取主成分 5 个，方差累积贡献率达到 94.41%。

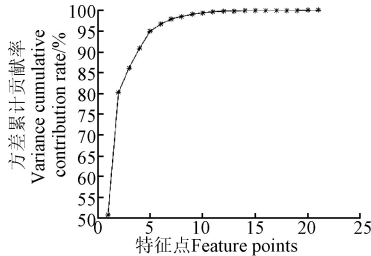


图 3 主成分方差累积贡献率

Fig.3 Cumulative contribution rate of principal component variance

Wilks Λ 统计分析筛选出 8 个变量，涉及试验传感器 6 个，如表 2 所示。

表 2 Wilks Λ 统计量选择
Table 2 Wilks Λ statistical selection

序号 No.	输入 Input	Wilks Λ
1	TGS2602_max	0.091
2	MQ5_max	0.062
3	MQ2_aver	0.052
4	MQ135_aver	0.045
5	TGS2610_aver	0.040
6	TGS2602_min	0.037
7	TGS2602_aver	0.030
8	TGS2600_aver	0.029

2.2 Fisher 判别分析

根据筛选的变量 PCA 和 Wilkss Λ 统计筛选的变量，

分别建立 PCA-Fisher 判别分析模型与 Wilks-Fisher 判别分析模型，2 种模型均包含 3 个判别函数，PCA-Fisher 判别函数如下：

$$PF_1=1.655\times PC_1+0.101\times PC_2-0.247\times PC_3-1.362\times PC_4+0.739\times PC_5 \tag{1}$$

$$PF_2=0.224\times PC_1-0.376\times PC_2+0.073\times PC_3+0.619\times PC_4-0.531\times PC_5 \tag{2}$$

$$PF_3=0.072\times PC_1+0.498\times PC_2+0.308\times PC_3+0.519\times PC_4-0.534\times PC_5 \tag{3}$$

式中 PF₁, PF₂, PF₃ 表示分类坐标值，PC₁~PC₅ 表示 5 个主成分。

Wilks-Fisher 判别函数如下：

$$WF_1=-0.001\times X_1-0.031\times X_2+0.028\times X_3-0.045\times X_4+0.015\times X_5+0.026\times X_6+0.005\times X_7-0.004\times X_8-5.56 \tag{4}$$

$$WF_2=-0.013\times X_1-0.049\times X_2+0.002\times X_3+0.093\times X_4-0.048\times X_5+0.066\times X_6-0.001\times X_7+0.036\times X_8-2.417 \tag{5}$$

$$WF_3=-0.002\times X_1-0.008\times X_2-0.024\times X_3-0.074\times X_4-0.138\times X_5-0.005\times X_6+0.034\times X_7+0.116\times X_8+5.101 \tag{6}$$

式中 WF₁, WF₂, WF₃ 表示分类坐标值，X₁~X₈ 表示表 2 中序号 1~8 的 8 个筛选变量。将输入导入 2 种模型，分别与其对应质心相比较，求得距离最短即为分析结果。分析结果如表 3 所示。

表 3 Fisher 判别分类结果
Table 3 Fisher discriminant classification results

函数 Functions	数据集 Data set	判别样本分类 Discriminant sample classification				平均正确率/ Average correct rate/%
		汉源青 Hanyuan green	云南青 Yunnan green	汉源红 Hanyuan red	陕西红 Shaanxi red	
PF	训练集	54	49	52	40	88.6
	验证集	19	16	17	20	90
WF	训练集	58	51	53	40	91.82
	验证集	20	17	17	20	95

由表 3 知，Wilks-Fisher 判别分析的整体正确率相对较高。其中在 PCA-Fisher 判别模型中，云南青花椒与汉源红花椒的识别率明显较低，这是因为，不同花椒的挥发性成分相对含量不同。文献[24]中指出，汉源青花椒和汉源红花椒的挥发性气体成分含量差异极大，汉源青花椒醇类化合物含量较高，而汉源红花椒酯类化合物较高，这是导致传感器在响应 2 种花椒时数值差异的主要原因，以表 2 中出现最多参数的传感器 TGS2602 为例，图 4a 是传感器 TGS2602 每组所采集样本的平均值（aver）折线图，从中可以明显得看出 4 种花椒的响应差异值较为明显；图 4b 是表 2 中未筛选传感器 TGS2611 每组采集样本的平均值折线图，可以发现，该传感器对 4 种花椒的响应值差异较低，印证了该传感器所提供信息存在较多冗余，而 PCA 会提取这一部分的冗余信息，这是导致 PCA-Fisher 模型识别精度较低的原因，也证明了 Wilks 统计分析对剔除冗余数据的有效性。图 4 c 是 4 类花椒在 Fisher 判别分析坐标下的坐标点，从图中可以看出，陕西

红花椒相对其他 3 种花椒区分度明显, 云南青花椒和汉源青花椒、汉源红花椒在一定程度上具有重叠部分, 这可能是因为汉源鲁甸 2 地相隔较近, 在地理位置, 土壤等因素较为接近引起的, 这也导致了云南青花椒和汉源红花椒的判别率相对较低, 故仅仅使用 Fisher 判别分析, 并不能完全将产地鉴别, 还需要对函数或数据进行优化分类。

2.3 细分类交叉验证下的 Wilks—Fisher 判别

对 Wilks—Fisher 判别模型进行改进, 将样本集进行

区分, 分别训练出红花椒的判别函数和青花椒的判别函数。为降低交互数据的干扰性, 对分析案例进行交叉验证。交叉验证是在机器学习建立模型和验证模型参数时常用的办法。交叉验证, 就是重复的使用数据, 把得到的样本数据进行切分, 组合为不同的训练集和测试集, 用训练集来训练模型, 用测试集来评估模型预测的好坏。在此基础上可以得到多组不同的训练集和测试集, 某次训练集中的某样本在下次可能成为测试集中的样本, 即所谓“交叉”。

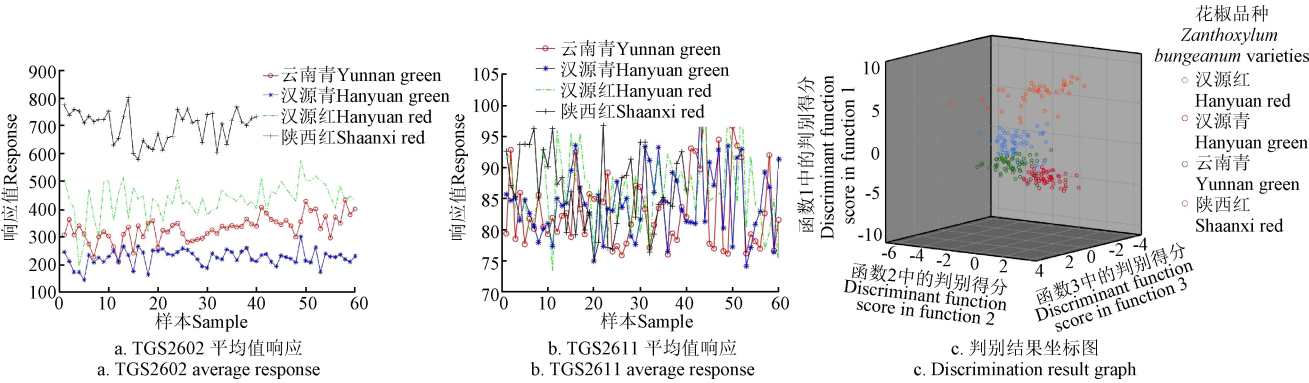


图 4 传感器部分响应及判别坐标
Fig.4 Sensor partial response and discrimination coordinates

采用 S 折交叉验证, 根据样本数量 S, 将对应训练集样本分成 S-1 份, 剩下 1 份作为测试机, 循环 S 次, 直到每个样本均作为测试集对象进行测试。在判别时, 将数据分为青红花椒进行分别验证, 结果如表 4 所示。

表 4 青红花椒分类交叉验证结果

Table 4 Cross-validation results of classification of Sichuan <i>Zanthoxylum bungeanum</i>				
花椒品种 <i>Zanthoxylum bungeanum</i> varieties	判别正确 Correct discriminati on	判别错误 Error discriminati on	合计 Total	正确率 Correct rate/%
汉源青 Hanyuan green	59	1	60	98.3
云南青 Yunnan green	56	4	60	93.3
汉源红 Hanyuan red	59	1	60	98.3
陕西红 Shaanxi red	40	0	40	100

由表 4 可知, 在分类判别时, 汉源青花椒的正确率为 98.3%, 误判 1 组, 云南青花椒的正确率为 93.3%, 误判 4 组, 陕西红花椒的正确率为 100%, 汉源红花椒的正确率为 98.3%, 误判 1 组, 平均正确率达到 97.27%, 所有验证结果较原模型检测均有所提高。该验证结果表明分类交叉验证的 Wilks—Fisher 判别模型可以较好地鉴别花椒的产地, 将判别函数写入采集装置并调试程序, 即构成完整的智能花椒品种鉴别系统。

3 结 论

本文采用自制的智能花椒品种鉴别装置, 利用气体传感器阵列对花椒数据进行采集, 通过主成分分析、Wilks A 统计分析提取有效变量, 建立对应的 Fisher 判别模型。

试验表明, Wilks—Fisher 具有更好的结果, 训练集的平均判别正确率达到 91.82%, 验证集平均判别正确率达到 95%。在细分类交叉验证的下的该模型判别率达到 97.27%。该模型能够正确地识别 4 个产地的花椒, 其检测结果较为理想。而将模型函数写入装置后, 构成智能花椒品种鉴别系统装置, 实现花椒产地的无损智能鉴别。该装置无需样品预处理, 不需要使用检测理化值所需的精密仪器, 检测成本较高光谱等精细设备极具优势, 为花椒产地的识别及追溯提供了技术支持, 可以有效控制以次充优、假冒伪劣等掺假手段, 为名优名品花椒, 具有地理标志保护花椒的鉴别提供了简便快捷的新思路, 具有广阔的应用前景。

[参 考 文 献]

[1] 毕君, 赵京献, 王春荣, 等. 国内外花椒研究概况[J]. 经济林研究, 2002, 20(1): 46—49.

[2] Huang Chengjiu, Rutaceae, Dicotyledoneae, Angiospermae. Flora Reipublicae Popularis Sinicae[M]. Beijing: Science Press, 1997.13.

[3] 吴习宇, 祝诗平, 黄华, 等. 近红外光谱技术鉴别花椒产地[J]. 光谱学与光谱分析, 2018(1): 68—72.

Wu Xiyu, Zhu Shiping, Huang Hua, et al. Near infrared spectroscopy for determination of the geographical origin of Huajiao[J]. Spectroscopy and Spectral Analysis, 2018(1): 68—72. (in Chinese with English abstract)

[4] 吴莉莉, 邢玉清, 郑宝周, 等. 机器视觉在花椒品种鉴别中的初步研究[J]. 传感技术学报, 2016, 29(1): 136—140.

Wi Lili, Xing Yuqing, Zheng Baozhou, et al. A preliminary

- study of zanthoxylum bungeanum maxim varieties discriminating by computer vision[J]. Chinese Journal of Sensors and Actuators, 2016, 29(1): 136—140. (in Chinese with English abstract)
- [5] Ito C, Katagiri H, Sato A, et al. Pharmacognostical Studies on the Sino-Japanese crude drugs “Huajiao” and “Sansho”: Part 5. On essential oils and pungent principles of pericarps of subgenus *Zanthoxylum* plants and commercial samples[J]. Natural Medicines, 1997, 51(3): 249.
- [6] Chen F Q. Means for rapid detection of toxic gases summary[J]. Autom Instrum, 2014(10): 79.
- [7] 韩崇昭, 朱洪艳, 段战胜, 等. 多源信息融合[M]. 北京: 清华大学出版社, 2006: 15—16.
- [8] 王欣. 多传感器数据融合问题的研究[D]. 吉林: 吉林大学, 2006: 13—15.
- [9] Hall D L, Llinas J. Handbook of Multisensor Data Fusion[M]. New York: CRC Press, 2001: 10—13.
- [10] 周亦斌, 王俊. 基于电子鼻的番茄成熟度及贮藏时间评价的研究[J]. 农业工程学报, 2005, 21(4): 113—117.
- Zhou Yibin, Wang Jun. Evaluation of maturity and shelf life of tomato using an electronic nose[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2005, 21(4): 113—117. (in Chinese with English abstract)
- [11] 殷勇, 郝银凤, 于慧春. 基于多特征融合的电子鼻鉴别玉米霉变程度[J]. 农业工程学报, 2016, 32(12): 254—260.
- Yin Yong, Hao Yinfeng, Yu Huichun. Identification method for different moldy degrees of maize using electronic nose coupled with multi-features fusion[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2016, 32(12): 254—260. (in Chinese with English abstract)
- [12] 张婷婷, 孙群, 杨磊, 等. 基于电子鼻传感器阵列优化的甜玉米种子活力检测[J]. 农业工程学报, 2017, 33(21): 275—281.
- Zhang Tingting, Sun Qun, Yang Lei, et al. Vigor detection of sweet corn seeds by optimal sensor array based on electronic nose[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2017, 33(21): 275—281. (in Chinese with English abstract)
- [13] 郝燕, 王永清, 王悦, 等. 检测细胞代谢气体成分诊断肺癌的方法[J]. 浙江大学学报: 工学版, 2008, 42(2): 294.
- [14] Brekelmans M P, Fens N, Brinkman P, et al. Smelling the diagnosis: The electronic nose as diagnostic tool in inflammatory arthritis[J]. PLOS One, 2016, 11(3): e0151715.
- [15] 张斌, 吴哲, 何宏建, 等. 结合磁共振成像和脑机接口的新型在体生物电子鼻的研究[J]. 中国生物医学工程学报, 2018, 37(1): 57—63.
- Zhang Bin, Wu Zhe, He Hongjian, et al. Characterization of in vivo bioelectronic nose with combined manganese-enhanced MRI and brain-computer interface[J]. Chinese Journal of Biomedical Engineering. 2018, 37(1): 57—63.
- [16] 方向生, 施汉昌, 何苗, 等. 气体传感器在环境监测中的应用与进展[J]. 环境科学与技术, 2011, 34(10): 112.
- [17] Tian F C, Zhang J, Yang S X, et al. Suppression of strong background interference on E-nose sensors in an open country environment[J]. Sensors, 2016, 16(2): 233.
- [18] 俞守华, 董绍娴, 区晶莹. 猪舍有害气体 NH_3 、 H_2S 的电子鼻定量识别[J]. 农业工程学报, 2009, 25(7): 153—157.
- Yu Shouhua, Dong Shaoxian, Ou Jingying. Quantitative identification of pernicious gases NH_3 and H_2S in piggery using electronic nose[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2009, 25(7): 153—157. (in Chinese with English abstract)
- [19] 殷勇, 赵玉珍, 于慧春. 基于多种变量分析方法鉴别食醋种类电子鼻信号特征筛选[J]. 农业工程学报, 2018, 34(15): 290—297.
- Yin Yong, Zhao Yuzhen, Yu Huichun. Feature selection of electronic nose signal for vinegar discrimination based on multivariable analysis [J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2018, 34(15): 290—297. (in Chinese with English abstract)
- [20] 沈飞, 刘鹏, 蒋雪松, 等. 基于电子鼻的花生有害霉菌种类识别及侵染程度定量检测[J]. 农业工程学报, 2016, 32(24): 297—302.
- Shen Fei, Liu Peng, Jiang Xuesong, et al. Recognition of harmful fungal species and quantitative detection of fungal contamination in peanuts based on electronic nose technology[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2016, 32(24): 297—302. (in Chinese with English abstract)
- [21] 裴高璞, 史波林, 赵镭, 等. 典型掺假蜂蜜的电子鼻信息变化特征及判别能力[J]. 农业工程学报, 2015, 31(增刊 1): 325—331.
- Pei Gaopu, Shi Bolin, Zhao Lei, et al. Information variation feature and discriminant capabilities of electronic nose for typical adulteration honey identification[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2015, 31(Suppl.1): 325—331. (in Chinese with English abstract)
- [22] 金慧芳, 史东梅, 陈正发, 等. 基于聚类及 PCA 分析的红壤坡耕地耕层土壤质量评价指标[J]. 农业工程学报, 2018, 34(7): 155—164.
- Jin Huifang, Shi Dongmei, Chen Zhengfa, et al. Evaluation indicators of cultivated layer soil quality for red soil slope farmland based on cluster and PCA analysis[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2018, 34(7): 155—164. (in Chinese with English abstract)
- [23] 高惠璇. 应用多元统计分析[M]. 北京: 北京大学出版社, 2005: 199—205.

- [24] 蒲凤琳, 孙伟峰, 车振明, 等. 水蒸气蒸馏结合 GC-MS 法分析比较四川汉源青、红花椒挥发性香气成分[J]. 中国调味品, 2017, 42(1): 23—27.
- Pan Fenglin, Sun Weifeng, Che Zhenming, et al. Analysis of

volatile aroma components in hanyuan green and red *Zanthoxylum bungeanums* by steam distillation and GC-MS[J]. China Condiment, 2017, 42(1): 23—27. (in Chinese with English abstract)

Identification of *Zanthoxylum bungeanum* origin based on gas sensor

Pang Tao¹, Yang Xiao¹, Chen Xiaoyan^{2,3*}, Tao Huailiang¹, Li Mengliang¹

(1. College of Mechanical and Electrical Engineering, Sichuan Agricultural University, Yaan 625000, China;

2. College of Information and Engineering, Sichuan Agricultural University, Yaan 625000, China;

3. The Lab of Agricultural Information Engineering, Sichuan Key Laboratory, Sichuan Agricultural University, Yaan 625000, China)

Abstract: At present, the identification of the origin of *Zanthoxylum bungeanum* is basically based on sensory evaluation, lack of objectivity, and it is difficult to quantify standardize when applying, and is difficult for non-professionals to make judgments. Therefore, in this paper, a smart device to quickly identify *Zanthoxylum bungeanum* was designed and developed. The device was based on the gas sensor array, including a control module, a temperature module, a data storage module, a fan module, and a display module, it could not only independently detect and identify the odor information of the *Zanthoxylum bungeanum*, but also distinguish the same kind of *Zanthoxylum bungeanum* from different places. The sensor array contained seven gas sensors, which could respond to irritating gases emitted by *Zanthoxylum bungeanums* such as benzene, alkanes, alcohols, and aldehydes. When the temperature was stable at about 26 degrees Celsius, it could effectively collect information on the odor emitted by *Zanthoxylum bungeanum*. Each group of *Zanthoxylum bungeanum* samples was collected 50 times, and the average value, the maximum value, and the minimum value were taken as sample recording parameters. In this paper, four kinds of *Zanthoxylum bungeanums* were selected as experimental subjects. Two kinds of green *Zanthoxylum bungeanums* were from Ludian in Yunnan and Hanyuan in Sichuan. At the same time, the two kinds of red *Zanthoxylum bungeanums* were from Hancheng in Shaanxi and Hanyuan in Sichuan. A total of 220 samples were collected as training sets, including 40 red *Zanthoxylum bungeanums* in Shaanxi and 60 samples in the remaining three samples. Another 80 samples were taken as the verification set, the number of samples for each *Zanthoxylum bungeanum* was 20 in the verification set as well. The detection data were processed using principal component analysis (PCA) and Wilks statistical analysis. Five principal components were extracted, and the cumulative contribution rate was 94.41%. The average accuracy rate of the training model corresponding to the Fisher discriminant model was only 88.6%, and the verification set was 90%. As a comparison, the Wilks statistical analysis finally eliminated 13 variables as well as selected 8 variables, and only TGS2611 sensor acquisition was not used. The average accuracy of the Fisher model training set was 91.82%, and the validation set was 95%. The results of the comparison of the two models indicate that the variables screened by Wilks are more effective in discriminating the *Zanthoxylum bungeanum* field. Among the four kinds of *Zanthoxylum bungeanums*, the recognition rate of Yunnan green *Zanthoxylum bungeanum* and Hanyuan red *Zanthoxylum bungeanum* was relatively lower than the others, and there was a phenomenon that the boundary data overlaps in the discrimination result graph. Then, to solve the problem, a Fisher discriminant model with cross-validation was established for the variables selected by Wilks statistic. In addition, the average accuracy rate reached 97.27%. Finally, the model was transplanted to the collection device to complete the identification device of intelligent *Zanthoxylum bungeanum* variety. It was a simple and efficient method for identifying *Zanthoxylum bungeanum* varieties and could provide a testing instrument and theoretical basis for further research on the origin and classification of *Zanthoxylum bungeanum*.

Keywords: sensors; crops; gas detectors; identification of *Zanthoxylum bungeanum* field; fisher discriminant