

基于手持式电子鼻的白酒识别

李金金, 孙哲华, 孟庆浩*

(天津大学 电气自动化与信息工程学院, 天津, 300072)

摘 要 现有电子鼻存在体积偏大、价格昂贵等问题, 阻碍了其普及应用。为解决这一问题, 该文提出了一种用于白酒检测的手持式电子鼻设计方案, 传感器阵列由 6 种贴片式气体传感器和 1 个温湿度传感器组成, 并置于一个微型仿生气室中。基于所设计的手持式电子鼻对 6 种白酒的识别问题进行了研究。首先, 从电子鼻传感器响应中提取特征, 并采用主成分分析法 (principal component analysis, PCA)、核主成分分析法 (kernel principal component analysis, KPCA) 和核熵成分分析法 (kernel entropy component analysis, KECA) 进行特征降维; 然后, 通过支持向量机 (support vector machine, SVM)、K 近邻 (k-nearest neighbor, KNN) 和反向传播人工神经网络 (back-propagation artificial neural network, BP-ANN) 对白酒进行识别。结果表明, 所设计的手持式电子鼻可以对不同种类的白酒进行在线检测与准确识别。

关键词 白酒; 手持式电子鼻; 气室; 特征提取; 特征降维; 分类

白酒作为世界六大蒸馏酒之一, 是国人的重要日常消费品。用于白酒检测的方法主要包括感官评价法及分析仪器法^[1-4]。感官评价法存在重复性低、主观性强及个体差异大等问题, 而分析仪器等检测设备体积大、价格高, 较难实现实时评测。因此, 有必要研发用于白酒快速检测和识别的新技术。

电子鼻是一种混合气味检测与识别设备, 具有快速、无损等优点, 在食品检测领域中发挥了重要作用^[5-7], 在白酒检测方面应用^[8-16]。

便携性或小型化对提高电子鼻的实用性非常重要。在便携式电子鼻的研究中, 美国研制的商业化电子鼻 Cyranose320 是其典型代表, 它采用 32 个聚合物传感器组成阵列, 尺寸为 (10 × 22 × 5) cm^[17]。徐后坤^[18]、ZHANG^[19]设计了用于 VOC 气体检测的手持式电子鼻。张延军^[20]和 LI^[21]开发了用于猪肉新鲜度检测的手持式电子鼻系统。李宇晓^[22]开发了基于 51 单片机的无线式手持式电子鼻, 用于 7 种中草药的识别。目前便携式电子鼻方面的研究^[14, 23-24]大多集中在数据处理。

本文以 6 种白酒为检测和识别对象, 开发了一种基于仿生气室的手持式电子鼻系统, 并根据采集响应信号特点, 分析比较了主成分分析 (principal component analysis, PCA)、核主成分分析 (kernel principal

component analysis, KPCA) 和核熵成分分析 (kernel entropy component analysis, KECA) 3 种降维方法以及支持向量机 (support vector machine, SVM)、K 近邻 (K-nearest neighbor, KNN) 和反向传播人工神经网络 (back-propagation artificial neural network, BP-ANN) 算法对白酒识别率的影响, 选出了最优算法, 最后将其应用到白酒识别实验中, 验证了手持式电子鼻的适用性。

1 材料与方法

1.1 手持式电子鼻系统

根据白酒香味物质特点 (乙酸乙酯、醇、醛等) 对气体传感器阵列进行了优化, 最终选定 6 种传感器用于手持式电子鼻, 分别是 CCS801 和 CCS803 (英国 CCS), TGS8100 (日本 Figaro), MiCS-5914、MiCS-5524 和 MiCS-4514 (英国 E2V), 并将最终选定的气体传感器和温湿度传感器制成传感器阵列电路。

手持式电子鼻系统采用动态顶空采样方法, 并设计了泵吹式气路系统。为了实现电子鼻小型化, 对各个硬件模块的选型进行了严格的控制。

自行设计的白酒检测用手持式电子鼻系统实物如图 1 所示。该电子鼻系统总重 283 g, 长 11 cm, 宽 9 cm, 高 3 cm, 平均功率在 5 W 以内。图 1 左侧图给出了系统的主要组成部分, 图 1 右侧展示了所设计电子鼻系统与一个 6.4 英寸手机的对比。

图 2 给出了仿生气室结构示意图。其中, 序号 1 和 6 分别是气室的入口和出口, 序号 2 代表气室的流

第一作者: 硕士研究生 (孟庆浩教授为通讯作者, E-mail: qh_meng@tju.edu.cn)。

基金项目: 国家自然科学基金 (61573253)

收稿日期: 2019-08-01, 改回日期: 2019-09-16

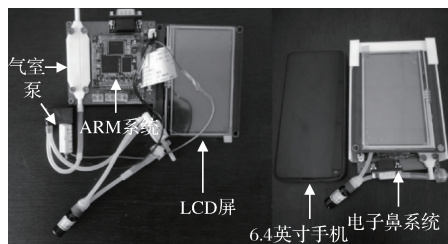


图1 自制手持式白酒检测电子鼻系统实物图

Fig.1 Photos of the self-designed handheld e-nose system for liquor detection

线型外壳,序号3是竖直隔板,序号4是水平挡板,序号5表示小圆柱结构,4个小圆柱结构通过螺丝将气体传感器阵列电路板紧固。

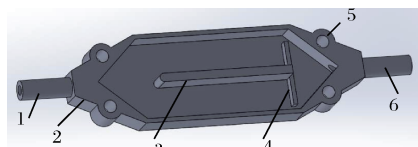


图2 仿生气室结构示意图

Fig.2 Schematic diagram of the bionic chamber

为比较不同气室方案对电子鼻系统的影响,设定总仿真时间为70 s,其中采样时间1 s,清洗时间69 s;取仿真时间为1 s时传感器阵列处的平均气压(相对于标准大气压)和气室出口处平均气流速度进行对比,结果如表1所示。

表1 各气室方案对应的气流速度和气压值

Table 1 Airflow velocity and pressure of each chamber scheme

气室方案	结构	传感器区 平均气压值/Pa	出口气流 平均速度/(m·s ⁻¹)
方案1	方形	153.90	4.412
方案2	流线外形	191.549	4.324 3
方案3	加中间隔板	192.237	4.295 6
方案4	挡板高度1 mm	202.953	4.247 4
方案5	挡板高度2 mm	210.043	4.184 0
方案6	挡板高度3 mm	217.967	4.155 9
方案7	挡板高度4 mm	223.086	4.110 5
方案8	挡板高度5 mm	234.436	3.993 9

综合考虑气压、气流速度和气体残留情况,选择挡板高度为3 mm的气室方案6作为最优仿生气室,并将其应用于手持式电子鼻的设计。

1.2 白酒数据处理方法

1.2.1 采样数据预处理

为了消除来自传感器阵列原始数据的噪声干扰,使用了中位值平均滤波和 Savitzky-Golay (S-G) 滤波。

中位值滤波可以有效滤除硬件电路中孤立的脉冲噪声,S-G滤波对气体传感器响应值的上下波动具有较好的平滑效果。

中位值滤波是将某一点的真实值用该点邻域窗口内所有点的中值代替,这种方法既能解决偶然因素引起的波动现象,也能使曲线更加平滑,如公式(1)所示:

$$y(i) = \text{Median}\{x(i-N), \dots, x(i), \dots, x(i+N)\} \quad (1)$$

式中: $x(i)$ 表示传感器第*i*个采样数值,即为位于窗口中心的传感器响应值, $y(i)$ 是更新后的窗口中心值, N 为正整数, $L = 2N + 1$ 。

S-G滤波是一种在时域内基于局域多项式最小二乘法拟合方法,具体方法如下:

首先建立一个宽度为 $n = 2m + 1$ 的滤波窗口, m 为窗口半径,取为51,设各个测量点为 $x = (-m, -m+1, \dots, m)$,采用 $k-1$ 次多项式对窗口内的数据点进行拟合,为保证每个参数有解,取 $k < n$,如公式(2)所示:

$$y = a_0 + a_1x + a_2x^2 + \dots + a_{k-1}x^{k-1} \quad (2)$$

式中:将这 n 个测量点带入方程,然后用最小二乘法将系数 a_i 拟合出来,最后将中间点 x_0 代入原多项式中,得到对应的 y_0 ,用该值作为最佳拟合值代替原来对应的值。

本文采用的数据预处理方法是上述中位值滤波和S-G滤波的叠加,即先进行 $L = 51$ 的中位值滤波,然后再进行滑动窗口为51,多项式阶数为3的S-G滤波。图3给出了传感器CCS801的一条原始响应曲线,使用上述两种滤波方法后的响应曲线如图4所示。由图4可知,滤波后传感器响应曲线更平滑,可以较好地去除传感器响应中存在的噪声。

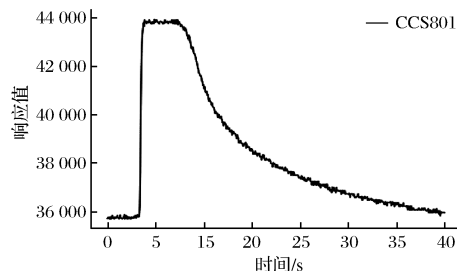


图3 CCS801 传感器响应曲线

Fig.3 Response curve of CCS801 sensor

1.2.2 特征提取

依据手持式电子鼻的采样方式和数据曲线的形状特点,最终选取了5个特征参数如下:一阶微分最大值 Max_{der} 、一阶微分最小值 Min_{der} 、二阶微分最大值

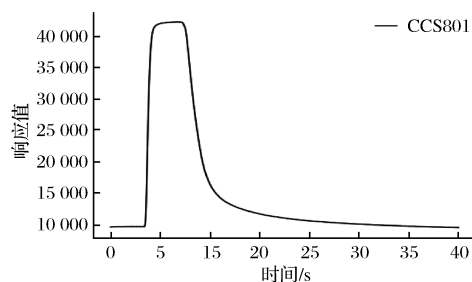


图4 滤波后的 CCS801 传感器响应曲线

Fig.4 Response curve of CCS801 sensor after filtering

Max_{der} 、一阶微分最大值处对应的响应值 R_e 、响应相对差值 R_r , 提取特征的公式如下:

$$\text{Max}_{\text{der}} = \max\left(\frac{dS_t}{dt}\right) \quad (3)$$

$$\text{Min}_{\text{der}} = \min\left(\frac{dS_t}{dt}\right) \quad (4)$$

$$\text{Max}_{\text{der}} = \max\left(\frac{d^2 S_t}{dt^2}\right) \quad (5)$$

$$R_e = \text{Max}_{\text{der}} \times dt \quad (6)$$

$$R_r = \max(S_t) - S_{i_0} \quad (7)$$

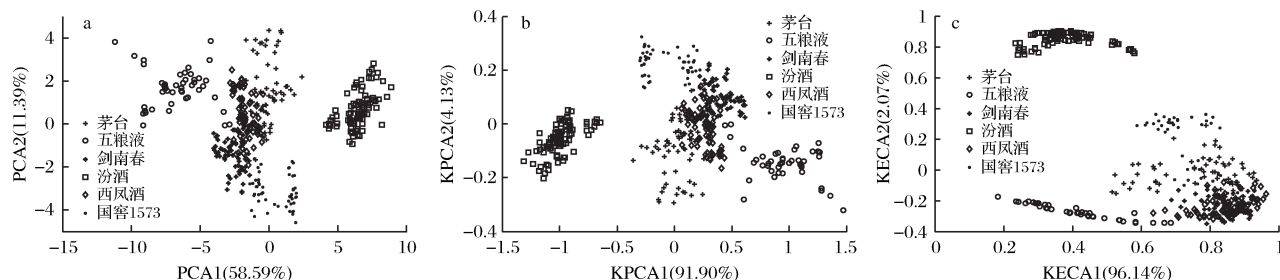
式中: S_t 为传感器输出电压信号对应的数字量, S_{i_0} 为响应起始点对应的传感器输出电压信号所对应的数字量。

1.2.3 特征降维和分类方法

经过特征提取后,传感器响应数据仍为高维阵列矩阵,特征之间存在一定的冗余信息。为了消除这些冗余信息的影响,需要进一步降维才可用于分类器。本文分别使用 PCA、KPCA 和 KECA 3 种方法进行降维,并对其进行了对比,选出最佳的降维方法。分类算法选择了 3 种比较常用的模式识别算法来分别进行分类识别,分别是 SVM、KNN 和 BP-ANN。

2 结果与分析

2.1 实验材料以及采样过程



a - PCA 降维后的主成分分布; b - KPCA 降维后的主成分分布; c - KECA 降维后的主成分分布

图5 基于3种不同的降维方法白酒样本主成分分布

Fig.5 Distribution of principal components in liquor samples based on three different dimension reduction methods

本文选取了 6 种不同类型的白酒,其具体信息见表 2。采样时,室内温度 28 ℃,相对湿度 30%。实验采用的采样参数设置如下^[25]:采样频率为 400 Hz;采样时间 40 s,其中吸气时间为 1 s,吸气等待时间为 3 s,呼气时间为 35 s,呼气等待时间为 1 s;气室清洗时间为 30 s,然后再针对采集到的数据进行上述所提到的预处理以及特征提取。

表2 实验白酒样本信息

Table 2 Information of Chinese liquors used in the experiment

品牌	酒精度/(% vol)	香型	产地
汾酒	42	清香型	山西/汾阳
国窖 1573	52	浓香型	天津/红桥
剑南春	52	浓香型	天津/红桥
飞天茅台	53	酱香型	四川/泸州
五粮液	52	浓香型	安徽/蚌埠
西凤	52	凤香型	陕西/宝鸡

2.2 三种降维方法对比

本文分别采用 PCA、KPCA、KECA 3 种方法对提取出的特征进行降维,对降维后的数据分别取前 2 个主成分绘图以探究其降维后对于分类的效果,如图 5 所示,其中横纵坐标分别是第一主成分和第二主成分,括号中的值为其贡献值。从前 2 个主成分的贡献值可以看出,KECA 和 KPCA 的前 2 个主成分元的累积贡献率分别达到了 98.21% 和 96.03%,而 PCA 前 2 个主成分元的贡献率为 69.98%,可以看出 KECA 和 KPCA 两种降维方法对于数据的主要信息的提取程度要比 PCA 要强,并且 KECA 效果最好。从主成分分析图中可以看出,3 种降维方法的分类效果类似,都有 2 到 3 种样本的分类效果较好,其他种类的样本则存在一定程度的混乱,体现在图中即为不同类型的点相互混杂。所以还需要进一步分析,才能确定针对白酒的最优降维方法。

为了定量地描述不同降维方法的分类性能,本文引入了样本特征类间离散度 S_b 和类内离散度 S_w , 其中 S_b 表示不同类别样本之间的分离程度, S_w 代表同一类别样本的分离程度,其定义如公式(8)、(9)所示:

$$S_b = \sum_{i=1}^c (m_i - m)(m_i - m)^T \quad (8)$$

$$S_w = \sum_{i=1}^c \sum_{j=1}^{n_i} (x_{ij} - m_i)(x_{ij} - m_i)^T \quad (9)$$

式中: c 为样本类别数, n_i 是第 i 类样本的样本数, m_i 是第 i 类样本的均值, m 是所有样本的均值, x_{ij} 是第 i 类的第 j 个样本。

从类内离散度和类间离散度的计算过程来看,其中必要的步骤之一就是求得样本之间的欧氏距离,而欧式距离存在取值范围大的特征的影响会高于取值范围小的特征的问题,所以在计算类内离散度和类间离散度之前首先需要进行归一化以解决上述问题,归一化的定义如公式(10):

$$a'_i = \frac{a_i - \min(a_i)}{\max(a_i) - \min(a_i)} \quad (10)$$

式中: a_i 代表第 i 个特征中的一个特征值, $\max(a_i)$ 、 $\min(a_i)$ 分别代表所有样本的第 i 个特征中的最大值和最小值。

从分类角度来看,类间离散度应该越大越好,而类内离散度应该越小越好,这样有利于后续的分类。针对不同降维方法的类间离散度、类内离散度以及二者比值如表3所示,KECA降维方法的比值明显大于其他两种降维方法。综上,使用KECA进行降维有利于后续的分类识别,所以在后续分类识别算法中将采用KECA来进行降维。为了尽可能保证信息的完全,通过主元成分累计贡献率达到99.8%来确定最终降维后的维数,最终确定的KECA降维后的维数为10维。

表3 三种降维方法对应的类间离散度和类内离散度

Table 3 Inter-class dispersion and intra-class dispersion corresponding to three dimension reduction methods

特征	原始特征集	PCA	KPCA	KECA
类间离散度 S_b	5 250.635	42.829	47.813	160.245
类内离散度 S_w	1 4916.577	56.449	55.893	54.457
比值	0.352	0.759	0.855	2.943

2.3 三种分类算法对比

基于仿生气室的白酒样本共有426个,随机取其中的50个作为测试样本,其余作为训练样本进行分类识别。进行多次训练集和测试集的选取和训练,以保证结果的客观性。

将仿生气室下的手持式电子鼻采样数据作为分类样本,采用KECA进行降维,6类白酒样本共有396个,其中每类白酒有52~90个样本,取出总数的10%作为训练样本,并且保证训练样本中包含所有类别白酒,经过多次随机挑选训练样本后取分类正确率的平均值以分析分类效果,3种分类方法的识别结果如表4所示。

表4 三种分类算法的正确率

Table 4 Accuracies of three classification algorithms

识别正确率	SVM/%	KNN/%	BP-ANN/%
茅台酒	97.72	97.76	99.32
五粮液	98.51	98.58	99.58
剑南春	94.89	96.10	97.39
汾酒	98.40	99.00	100
西凤酒	95.34	96.10	98.48
国窖1573	95.78	97.72	99.40
总正确率	96.50	97.70	99.10

由表4可知,经过KECA降维后,BP-ANN对白酒的识别正确率最高(99.10%),其次是KNN和SVM。但是从运行时间来讲,在相同训练、测试样本数量以及迭代次数(即重新选择训练集和样本集)的情况下,SVM的分类识别时间为2.99 s,KNN的分类识别时间为1.10 s,而BP-ANN神经网络需要的时间为85.18 s。从手持式电子鼻对白酒识别的在线和快速方面考虑,综合识别正确率和运行时间,建议选择KNN作为分类算法。

3 结论

通过对6种白酒的采样、数据预处理、特征提取、降维和分类处理等操作,验证了本文自主研发的手持式电子鼻系统可用于白酒的在线检测与识别;分析比较了PCA、KPCA和KECA三种方法对气体传感器信号特征的降维效果,基于前2个主成分贡献值以及类间离散度和类内离散度比值,发现KECA效果更好;综合比较了SVM、KNN和BP-ANN三种分类算法的识别率和运行时间,3种方法识别率均可达到94%以上,但KNN的识别时间最短。

参 考 文 献

- [1] 沈怡方. 中国白酒感官品质及品评技术历史与发展[J]. 酿酒, 2006, 33(4): 3-4.
- [2] 张健,赵镭,欧阳一非,等. 现代仪器分析技术在白酒感官评价研究中的应用[J]. 食品科学, 2007, 28(10): 561-565.
- [3] 胡国栋. 气相色谱在白酒分析中的应用现状与回顾

- [J]. 食品与发酵工业, 2003, 29(10): 65-69.
- [4] 季克良, 郭坤亮, 朱书奎, 等. 全二维气相色谱/飞行时间质谱用于白酒微量成分的分析[J]. 酿酒科技, 2007(3): 100-102.
- [5] 王俊, 胡桂仙, 于勇, 等. 电子鼻与电子舌在食品检测中的应用研究进展[J]. 农业工程学报, 2004, 20(2): 292-295.
- [6] 吴守一, 邹小波. 电子鼻在食品行业中的应用研究进展[J]. 江苏理工大学学报(自然科学版), 2000, 21(6): 13-17.
- [7] 杜锋, 雷鸣. 电子鼻及其在食品工业中的应用[J]. 食品科学, 2003, 24(5): 161-163.
- [8] 马泽亮, 国婷婷, 殷廷家, 等. 基于电子鼻系统的白酒掺假检测方法[J]. 食品与发酵工业, 2019, 45(2): 190-195.
- [9] 周宇, 孟庆浩, 亓培锋, 等. 基于单类支持向量机的电子鼻在白酒真假识别中的应用[J]. 传感技术学报, 2015, 28(12): 1741-1746.
- [10] 张君生, 李臻峰, 宋飞虎, 等. 基于 zNoseTM 电子鼻的新产原酒品质检测[J]. 食品与发酵工业, 2018, 44(12): 216-220.
- [11] 徐晚秀, 李臻峰, 张振, 等. 基于电子鼻的中国白酒酒龄检测[J]. 食品与发酵工业, 2016, 42(2): 144-149.
- [12] 王辉, 李臻峰, 邓霞, 等. 基于电子鼻对不同香型白酒的快速识别和分类[J]. 食品工业科技, 2017, 38(6): 62-65; 83.
- [13] 柯永斌, 周红标, 李珊, 等. 基于电子鼻的不同香型白酒快速识别[J]. 酿酒科技, 2013(11): 1-3; 8.
- [14] 张松, 张覃轶, 张顺平. 基于便携式电子鼻的同香型白酒识别[J]. 酿酒科技, 2018(6): 25-29.
- [15] 殷勇, 申晓鹏, 于慧春. 基于 KECA + FDA 的白酒电子鼻多特征鉴别方法[J]. 农业机械学报, 2018, 49(4): 374-380.
- [16] 亓培锋, 孟庆浩, 井雅琪, 等. 用于白酒识别的电子鼻数据分析与参数优化[J]. 天津大学学报(自然科学与工程技术版), 2015, 48(7): 643-651.
- [17] DUUTA R, HINES E L, GARDNER J W, et al. Bacteria classification using Cyranose 320 electronic nose [J]. Biomedical Engineering Online, 2002, 1(1): 4.
- [18] 徐后坤. 基于 μ PSD3234A 单片机的手持式电子鼻的研制[D]. 武汉: 华中科技大学, 2007.
- [19] ZHANG G L, ZHANG X M, JIN J J, et al. Development of portable electronic nose for VOC detection[J]. Applied Mechanics and Materials, 2014, 568: 420-423.
- [20] 张延军. 基于嵌入式系统的电子鼻研制[D]. 杭州: 杭州电子科技大学, 2011.
- [21] LI H, CHEN Q, ZHAO J, et al. Non-destructive evaluation of pork freshness using a portable electronic nose (E-nose) based on a colorimetric sensor array[J]. Analytical Methods, 2014, 6(16): 6271-6277.
- [22] 李宇骁. 基于 51 单片机的无线式手持电子鼻系统的开发与应用研究[D]. 武汉: 华中科技大学, 2012.
- [23] 陆艺莹, 张覃轶, 龙海仙, 等. 基于便携式电子鼻的白酒快速识别[J]. 酿酒科技, 2016(1): 83-86.
- [24] 李志华. 用于白酒识别的手持式电子鼻系统设计与应用[D]. 天津: 天津大学, 2017, 11.
- [25] QI P F, MENG Q H, JING Y Q, et al. A bio-inspired breathing sampling electronic nose for rapid detection of Chinese liquors [J]. IEEE Sensors Journal, 2017, 17(15): 4689-4698.

Chinese liquors recognition based on a handheld electronic nose

LI Jinjin, SUN Zhehua, MENG Qinghao*

(School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China)

ABSTRACT The existing electronic noses (e-noses) have problems such as large volume and high price, which hinders its popularization and application. To solve these problems, a design scheme of hand-held electronic nose for liquor detection is proposed. The sensor array consists of six patch gas sensors and one temperature and humidity sensor, and is placed in a miniature bionic chamber. Based on the designed hand-held electronic nose, the recognition of six kinds of liquor was studied. Firstly, the features are extracted from the e-nose sensors' responses, and feature reduction is performed by principal component analysis (PCA), kernel principal component analysis (KPCA) and kernel entropy component analysis (KECA). Then, the liquors are classified through support vector machine (SVM), K-nearest neighbor (KNN) and back-propagation artificial neural network (BP-ANN). The results show that the designed handheld electronic nose can detect and accurately identify different kinds of Chinese liquors online.

Key words liquor; handheld electronic nose; chamber; feature extraction; feature dimensionality reduction; classification