

龙井茶叶品质的电子鼻检测方法^{*}

于慧春 王 俊 张红梅 于 勇

【摘要】 针对茶叶品质感官审评的不足,采用电子鼻检测手段,对4个不同等级的龙井茶作等级判别。对传感器信号进行多因素方差分析得出:不同容器容积和不同采样时刻对传感器的响应信号有着显著的影响。通过主成分(PCA)、线性判别(LDA)和BP神经网络方法对各茶叶样品进行了分类判别。PCA对于等级差别较近的茶叶区分结果不太理想;而LDA相对于PCA有较好的区分效果;设计BP神经网络拓扑结构为30-12-4,通过对网络进行适当训练,总的测试回判率可达到90%。

关键词: 龙井茶 电子鼻 BP神经网络

中图分类号: TS272.7

文献标识码: A

Measurement of the LongJing Tea Quality by Using an Electronic Nose

Yu Huichun Wang Jun Zhang Hongmei Yu Yong
(Zhejiang University)

Abstract

An investigation was made to determine the four tea samples with different quality grade by using an electronic nose (e-nose). The response signals of e-nose were analyzed under different sampling conditions by variance analysis and multivariate analysis. Analytical results showed that the different volume of vials and the different collection times have significant effect on the response signals of the e-nose. Then the data were processed using principal components analysis (PCA), linear discrimination analysis (LDA) and artificial neural network (ANN). The results analyzed by LDA were superior to that by PCA, which could distinguish all the tea samples completely. However, PCA method could not estimate sample of A280 correctly. Further 90% correct classification was achieved for all the tea samples using the BP neural network.

Key words LongJing tea, Electronic nose, BP neural network

引言

目前,茶叶品质大多数是通过人的感官评定的。人感觉器官的灵敏度易受外界因素的干扰而改变,从而影响评定的准确性。电子鼻技术的兴起让人们看到综合评价气味整体信息的巨大潜力^[1]。近年来,电子鼻得到了广泛的研究和应用,尤其在食品和饮料工业方面^[2~7]。其中,最主要的一个方面就是根据各样品的不同挥发物进行识别分类。电子鼻把具有

不同特性的气体传感器组合成气体传感器阵列,使其检测范围更宽;同时,在模式识别数据处理方面也在不断发展完善各种识别方法,如线性识别函数法、人工神经网络法等。人工神经网络的最大优点就是可实现复杂的非线性映射,并且具有良好的容错性,有助于提高气体检测的精度^[8]。本文运用电子鼻对不同等级的西湖龙井茶叶香气进行检测,然后进行适当的特征提取,结合统计模式识别方法和神经网络分析方法对输入的茶叶香气信号进行模式识别,

收稿日期: 2006-02-24

^{*} 国家自然科学基金资助项目(项目编号:30571076)和国家教育部新世纪优秀人才支持计划资助项目(项目编号:NCET-04-0544)

于慧春 浙江大学生物系统工程与食品科学学院 博士生, 310029 杭州市

王 俊 浙江大学生物系统工程与食品科学学院 教授 博士生导师 通讯作者

张红梅 浙江大学生物系统工程与食品科学学院 博士生

于 勇 浙江大学生物系统工程与食品科学学院 博士后

实现对龙井绿茶等级的评定。

1 试验材料、仪器、方法与信号处理

1.1 试验材料和仪器

茶叶来自于浙江大学茶学系的西湖龙井, 试验前锡纸袋密封贮藏。不同等级的茶叶样品分别编号为 A 120、A 280、A 380、A 600。

试验采用德国 A R S E N S E 公司的 P E N 2 型便携式电子鼻 (portable electronic nose), 包含 10 个金属氧化物传感器。该电子鼻具有自动调整、自动校准及系统自动富集的功能。响应信号为传感器接触到样品挥发物后的电导率 G 与传感器在经过标准活性炭过滤气体后的电导率 G_0 的比值。仪器主要包含: 传感器通道、采样通道、计算机。

1.2 试验方法

对龙井茶 A 280 进行取样, 每个样品 5 g。分别放置在 50 mL、150 mL、250 mL、500 mL 的烧杯中, 瓶口用双层薄膜密封静置, 静置时间分别为 45 m i n 和 120 m i n。各样品重复测量 15 次, 每次采样时间为 60 s, 每秒采样一次, 电子鼻软件每秒自动记录一次数据。通过对上述试验数值的分析, 确定出在本试验中 (针对 5 g 茶叶样品) 应采用的最佳烧杯容积和样品静置时间的搭配, 以获得较佳响应信号。然后再对 4 个不同等级的龙井茶进行测量, 各个等级的茶叶分别重复测量 45 次。试验在室温 25℃ 下进行。

1.3 数据处理方法

采用单因素及多因素方差分析来确定较优的试验条件和特征提取点; 采用主成分分析、线性判别分析以及 B P 神经网络对不同茶叶样品进行定性判别。

2 结果与讨论

2.1 方差分析

2.1.1 多因素方差分析

A 280 级茶叶在不同静置时间和不同烧杯容积的条件下重复 15 次, 测得的数值进行多因素方差分析, 考察顶空静置时间 A (45 m i n、120 m i n)、烧杯容积 B (50 mL、150 mL、250 mL、500 mL) 和采样时刻 C (1~ 60 s) 3 个因素的变化对响应信号影响的显著性。使用 S A S 数据处理软件进行多元方差分析, 分析结果如表 1 所示。

从表 1 可以知: $F(B) = 31.505 \gg F_{0.01}(3, \infty) = 3.78$, 对电子鼻响应信号有极显著的影响; $F(C) = 2.091 > F_{0.01}(59, \infty) = 1.47$, $F(A \times B) = 4.225 > F_{0.01}(3, \infty) = 3.78$, 对电子鼻响应信号影响高度显著; 而 $F(A) = 0.478 < F_{0.05}(1, \infty) = 3.84$, $F(A \times$

$C) = 0.006 < F_{0.05}(59, \infty) = 1.32$, $F(B \times C) = 0.110 < F_{0.05}(\infty, \infty) = 1$, $F(A \times B \times C) = 0.019 < F_{0.05}(\infty, \infty) = 1$, 没有显著影响。

表 1 A 280 级茶叶的 3 因素(A、B、C) 方差分析
Tab 1 The VAOAN for tea A280

变异来源	平方和	自由度	均方	F 值	显著水平
A	0.189	1	0.189	0.478	0.489
B	37.426	3	12.475	31.505	0.000
C	48.843	59	0.828	2.091	0.000
A × B	5.019	3	1.673	4.225	0.005
A × C	0.131	59	0.002	0.006	1.000
B × C	7.692	177	0.043	0.110	1.000
A × B × C	1.324	177	0.007	0.019	1.000
误差	1 710.621	4 320	0.396		
总和	1 811.245	4 799			

从 F 值可知, 测量时烧杯容积对信号响应的影响最大; 其次是顶空静置时间 A 和烧杯容积 B 的交互作用 $A \times B$ 的影响; 然后是采样时刻 C 的影响。并且, 只有顶空静置时间 A 和烧杯容积 B 之间交互作用高度显著, 采样时刻与顶空静置时间和烧杯容积之间的交互作用不显著。由于顶空静置时间 A 对电子鼻的响应信号影响不显著, 所以后面的试验采用 45 m i n 的静置时间。

2.1.2 容积的方差分析

对 4 个等级的茶叶在不同容积烧杯中顶空静置 45 m i n 样品的测试数据分别进行单因素方差分析, 考察烧杯容积对电子鼻响应信号的影响, 如表 2 所示 (重复 15 次)。可以看出, 采用 B_3 (250 mL) 的烧杯时, 对茶叶的影响最大 ($F = 102.2254$)。所以, 采用 250 mL 的烧杯时对不同等级茶叶区分效果比较好, 其次是采用 B_4 (500 mL) 的烧杯。

表 2 不同烧杯容积对不同等级茶叶的方差分析
Tab 2 Multivariate analysis for 4 tea groups with different vials (15 duplicators for each tea group)

容积	茶叶样品	Wilks' Lambda	F	显著水平
B_1	A 120、A 280、A 380、A 600	0.237	68.429	0.000
B_2	A 120、A 280、A 380、A 600	0.290	83.883	0.000
B_3	A 120、A 280、A 380、A 600	0.354	102.225	0.000
B_4	A 120、A 280、A 380、A 600	0.299	86.383	0.000

2.1.3 不同时刻的方差分析

4 种不同等级茶叶在 250 mL 容积烧杯中静置 45 m i n 条件下, 对检测的数据分别进行单因素方差分析, 考察不同采样时刻下电子鼻响应信号的差异。

电子鼻对一个茶叶样品的响应曲线如图 1 所示。用电子鼻对其他样品进行检测, 都可以得到类似的响应曲线图。分别提取 15 s、30 s、45 s 和 60 s 处的信号值作为分析数据(如图 1 中圆点标记采样点), 进行方差分析。分析结果如表 3 所示(重复 15 次)。

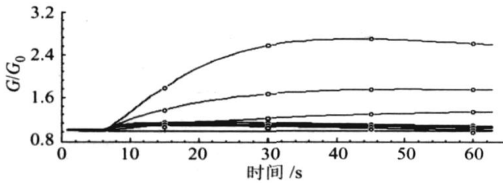


图 1 电子鼻对密封在 250 mL 烧杯中静置 45 min 的一个茶叶样品的响应曲线

Fig 1 Response curves of the e-nose to a tea sample sealed into 250 mL vial and headspace generated time 45 min

表 3 不同时刻对 4 个等级的茶叶影响的方差分析

Tab 3 Multivariate analysis for 4 tea groups at different times (15 duplicators for each tea group)

时刻/s	样本	Wilks' Lambda	F	显著水平
15	A 120, A 280, A 380, A 600	0.001	133.01	0.000
30	A 120, A 280, A 380, A 600	0.000	49.62	0.000
45	A 120, A 280, A 380, A 600	0.000	21.37	0.007
60	A 120, A 280, A 380, A 600	0.001	146.57	0.000

由表 3 可知, 4 个不同时刻 4 个等级龙井茶的均值比较结果差别很大。尤其是 60 s 时差异最为显著($F = 146.57$), 其次是 15 s 时的数据。因此, 采用 60 s 时的数据作为特征值来进行模式识别。

2.2 模式识别

2.2.1 主成分分析

对 A 280 等级茶叶静置 45 min 的条件下, 在 4 种容积烧杯中的响应值进行主成分分析。用 2 个主成分来反映原来的 10 个原始指标。并用主成分值计算各样品间的距离(马氏距离)进行样品的聚类分析, 聚类结果如图 2 所示。第 1 主成分反映了原始变量 74.07% 的信息, 第 2 主成分反映了原始变量 21.31% 的信息。

图 2 中结果可以分成 2 组, A 组是采用 50 mL 和 150 mL 烧杯时的电子鼻响应信号, 由于烧杯容积较小, 顶空茶叶挥发性成分浓度较大, 电子鼻信号响应较高; B 组则采用 250 mL 和 500 mL 的较大容积烧杯, 因此顶空茶叶挥发性成分浓度较小, 电子鼻响应信号相应较小, 但是在整个采样过程中信号较稳定。2 组内部的子组由于信号较接近, 聚类的结果分别出现部分重叠。还可以看出, 采用 250 mL 和

500 mL 的烧杯时信号均比较集中, 因此, 相对于同一等级茶叶样品, 250 mL 烧杯的信号更为集中, 所以以后的试验中均采用 250 mL 的烧杯进行。

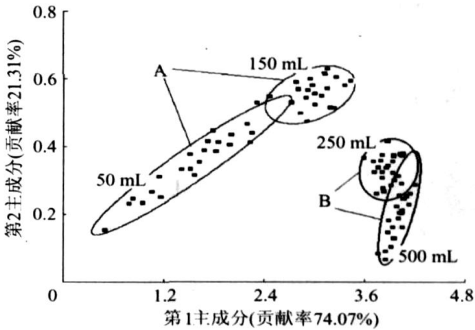


图 2 A 280 等级茶叶用不同容积烧杯顶空静置 45 min 的 PCA 图

Fig 2 PCA plot for tea A 280 by using vials with different volume and headspace generated time 45 min

对各个等级的茶叶进行判别, 分别取各等级茶叶第 60 s 时的数据作为特征值进行主成分分析, 得到前 2 个主成分的贡献率分别为 60.22%、26.94%, 累积贡献率为 87.16%。所以取前 2 个主成分对应的特征向量所决定的两维子空间就能够充分保存原始数据的信息, 如图 3 所示。

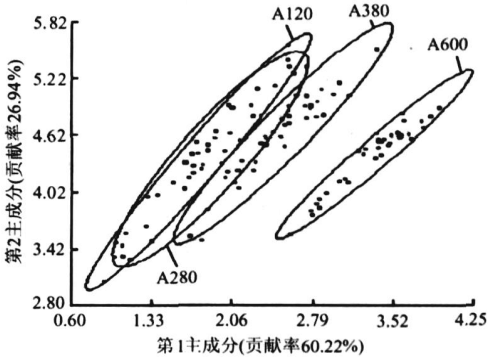


图 3 4 个等级茶叶的 PCA 分析结果

Fig 3 PCA plot for 4 tea groups at the 60th second

由图 3 可以看出, 由于 A 600 级的龙井茶和其他 3 种茶等级区别较大, 所以其分布区域和其他 3 种茶的分布区域分离较远, 差别很明显。A 120、A 280、A 380 等级比较接近, 电子鼻检测时数据分布比较集中, 所以在 PCA 分析图中分布比较集中。A 120 和 A 380 只有小部分重叠, 而比较突出的是 A 280, 分别与 A 120 和 A 380 交错重叠。这是因为其等级正好介于 A 120 和 A 380 之间, 且与这 2 个等级的茶叶级别相差不大, 相比之下, A 120 和 A 380 之间的差别则比较大。这说明传感器对不同等级茶叶挥发物质的响应区别相当明显。通过 PCA 分析可以将等级差别较大的茶叶较好地区分开, 但是对于等级差别较小的茶叶区分还存在着局限。

2.2.2 线性判别函数分析

线性判别分析结果如图 4 所示。通过线性判别,选择合适的判别式,可以较好地将 4 个等级的龙井茶区分开。说明不同品质等级的茶叶之间,挥发物成分存在一定差异,可以被电子鼻检测到。从图 4 可以看出,LDA 分析结果明显优于 PCA 分析,可以将类样品完全分开。而且等级差别越大,分离效果越好。

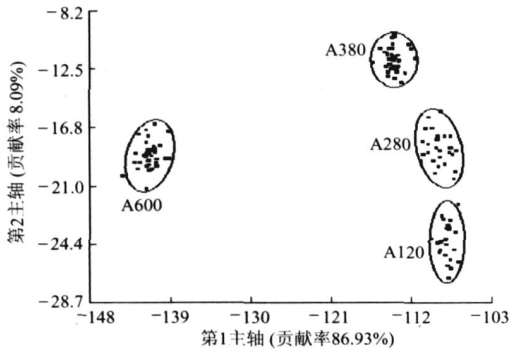


图 4 4 个等级茶叶的 LDA 分析结果

Fig 4 LDA plot for 4 tea groups at the 60th second

2.2.3 人工神经网络

如图 1 所示,从每个传感器的响应曲线上分别提取 60 s 时的响应值、信号最大值、全段平均值作为其特征向量。每个样本中提取的特征向量包括 30 个特征值,作为神经网络的输入矢量。本文分析了 4 个等级的龙井茶,每种等级的茶叶取 45 个样品,共得到 180 个样品(4×45×30=5 400 个特征值)。

用 BP 神经网络对不同等级的龙井茶进行定性分级。180 个样品分成 2 部分,120 个作为训练集,60 个作为测试集。网络输入层神经元数等于每个样品的特征向量的维数,输出层神经元数等于样品的种类数,隐层神经元数通过几次测试结果比较来确定。

选择误差目标参数 0.01,通过对网络训练时间和所需训练步数的比较来确定较合理的隐层神经元数。综合考虑所要达到的精度和网络的训练时间,网络结构设计为:30-12-4,通过测试总体回判率可达到 90%。神经网络分析的结果如表 4 所示。

表 4 神经网络的分析结果

Tab 4 The ANN analysis results

样品	期望输出	实际结果输出	
		训练样本正确率/%	测试样本正确率/%
A 600	[1 0 0 0]	100	100
A 380	[0 1 0 0]	100	80
A 280	[0 0 1 0]	100	86.67
A 120	[0 0 0 1]	100	100

3 结 论

(1)通过 3 因素方差分析表明对于一定质量的样品,试验容器的容积对其响应信号的影响极为显著;对茶叶来说,不同顶空静置时间对其信号影响不大;对 4 个等级茶叶响应值方差分析确定出:在第 60 s 时各等级茶叶的响应值差异最明显,最有益于样品的识别分类。所以选取 60 s 时的数据进行 PCA 和 LDA 分析。

(2)PCA 分析可以将等级差别明显的样品区分开,而对于等级差别较小的样品不能很好地实现分类,存在很大的局限;LDA 分析可以将各个等级的样品完全分开,效果优于 PCA 分析。

(3)从每个传感器响应曲线中提取 3 个特征值,组成包含 30 个特征值的特征向量作为神经网络识别分析的输入矢量,网络测试总体回判率达到 90%。

参 考 文 献

1 于勇, 王俊, 周鸣 电子鼻的研究进展及其在农产品加工中的应用[J] 浙江大学学报, 2003, 29(5): 579~ 584

2 邹小波, 赵杰文, 潘胤飞, 等 基于遗传 RBF 网络的电子鼻对苹果质量的评定[J] 农业机械学报, 2005, 36(1): 61~ 64

3 石志标, 左春桢, 杨长明 仿生鼻系统的试验研究[J] 农业机械学报, 2005, 36(2): 70~ 72

4 Said L, Sandrine B, Sonia C, et al Shelf life determination by electronic nose: application to milk[J] Sensors and Actuators B: Chemical, 2005, 106(1): 199~ 206

5 Rajanaki T, Alakom i H L, Ritvanen T, et al Application of an electronic nose for quality assessment of modified atmosphere packaged poultry meat[J] Food Control, 2006, 17(1): 5~ 13

6 Panigrahi S, Balasubramanian S, Gu H, et al Neural-network-integrated electronic nose system for identification of spoiled beef[J] LWT-Food Science and Technology, 2006, 39(2): 135~ 145

7 Brezmes J, Cabré P, Rojo S, et al Discrimination between different samples of olive oil using variable selection techniques and modified fuzzy artmap neural networks[J] IEEE Sensors Journal, 2005, 5(3): 463~ 470

8 周亦斌, 王俊 新技术在茶叶品质评价中应用的现状及趋势[J] 茶叶科学, 2004, 24(2): 82~ 85