

基于主元分析的电子舌味觉检测

蒋行国¹, 侯洪彬¹, 陈真诚²

1. 桂林电子科技大学信息与通信学院, 广西桂林 541004

2. 桂林电子科技大学生命与环境科学学院, 广西桂林 541004

摘要 基于光寻址(LAPS)型电子舌的基本结构及检测原理,提出了一种基于主元分析(PCA)的味觉识别算法。通过对 LAPS 型电子舌提取到的信号特征进行主元分析,在其第一主元和第二主元得分图上,5 种基本味道样本溶液的落点在各自区域内互不干扰,5 种基本味道能够很好地区分。进一步采用最小距离法对其分类,表明该方法优势突出,对 5 种基本味道分类的平均识别准确率达到 96.54%。试验表明,将主元分析应用于 LAPS 型电子舌的味觉检测分析方面有很好的效果,为研制小型化 LAPS 型电子舌提供了科学依据。

关键词 光寻址;味觉检测;主元分析;最小距离法

中图分类号 TP391.4

文献标识码 A

doi 10.3981/j.issn.1000-7857.2011.23.009

Electronic Tongue Taste Detection Based on Principal Component Analysis

JIANG Xingguo¹, HOU Hongbin¹, CHEN Zhencheng²

1. School of Information & Communication, Guilin University of Electronic Technology, Guilin 541004, Guangxi Province, China

2. School of Life and Environmental Sciences, Guilin University of Electronic Technology, Guilin 541004, Guangxi Province, China

Abstract According to the fundamental structure and the detection principle of the Light Addressable Potentiometric Sensor (LAPS) type electronic tongue, a taste recognition algorithm based on Principal Component Analysis (PCA) is designed. The signal detected by LAPS type electronic tongue is processed by PCA. The scores pictures of the first and the second principal components show that the spots of five basic tastes do not interfere with each other. Therefore, the five basic tastes could be more accurately recognized by PCA. Meanwhile, minimum distance method is employed to classify the principal components and the results of the experiments show that the method is prominent and the average rate of identification could reach at 96.54%. Final test results indicate that applying PCA to LAPS type electronic tongue has the huge potential for analyzing the taste of liquid, and also provides a scientific proof for developing the miniature LAPS type electronic tongue.

Keywords LAPS; taste detection; PCA; minimum distance method

0 引言

随着人们对食品安全及品质的要求不断提高,仅依靠带有强烈主观性的品评专家难以满足社会的需求,因而必须寻找一种科学、有效、快捷、方便、无损的评品手段。自 20 世纪 80 年代 Hafeman 等首次提出味觉传感器阵列以来^[1],国内外

研究者在味觉传感器方面做了大量研究,如 Legin 和 Vlasov 研究小组^[2]于 1995 年研究开发了硫属玻璃膜传感器系统,并首次以“电子舌”这个更形象的名称进行命名。

光寻址(LAPS)型电子舌系统主要由传感器阵列、信号处理和模式识别 3 部分组成。模式识别系统是电子舌的重要组

收稿日期:2011-07-31;修回日期:2011-08-05

基金项目:国家高技术研究发展计划(863 计划)项目(2009AA04Z214)

作者简介:蒋行国,副教授,研究方向为信号与信息处理,电子信箱:tonny_jiang@guet.edu.cn;陈真诚(通信作者),教授,研究方向为传感器与智能仪器、医学图像处理,电子信箱:chenzhcheng@163.com

成部分,主要是对传感器阵列采集的数据进行预处理、特征提取的自动分析与综合,以完成对待测样品定性定量的决策、评估和分类^[9]。主元分析(Principal Component Analysis, PCA)是电子舌模式识别方法中最常用的一种,区分应用非常广泛。浙江大学王平等选用主元分析法研究出了一台用于海水中衡量金属检测的电子舌^[4];王俊等^[5]使用 PCA 法对葡萄酒进行分类识别研究;西班牙 Moreno 等^[6]采用 PCA 法准确辨别了矿泉水的品牌,为矿泉水的快速分类和质量控制提供了一个非常可行的方法。

LAPS 型电子舌采集的数据量非常大,增加了计算量和分析问题的复杂性。另外,LAPS 型电子舌交互感应特性导致数据之间存在一定的相关性以及数据信息之间的冗余性,而 PCA 则是在保持原有信息属性不变的基础上,通过降维技术从多个变量中提取出几个主元的多元统计方法。因此,PCA 技术应用于 LAPS 型电子舌系统中成为必然。

鉴于 PCA 在处理多数据上的优越性,本文介绍了 PCA 原理以及在 LAPS 型电子舌系统中的应用,利用主元图比较直观地对 5 种基本味道进行了分类,在主元分析的基础上,用最小距离法对 PCA 识别能力进行验证,实现了对 5 种基本

味道较为快速、准确、无损的识别。

1 LAPS 型电子舌原理结构

LAPS 型电子舌是模拟人类味觉感知及识别系统的仪器^[7],其工作原理如图 1 所示,光寻址电位传感器是基于半导体技术的传感器,基本结构属于 EIS(电解质-绝缘层-硅)型^[8]。通过 AVR 单片机控制 LED 照射溶液池中与硅底相连的敏感膜时,由于敏感膜对离子的吸附作用形成膜电位和半导体的本征吸收作用产生的电子-空穴对,硅胶半导体和绝缘层两端的电压发生偏移。在偏置电压的作用下,LAPS 产生与待测溶液中离子相对应的光电流信号,由于光电流信号非常小,另外电压信号比电流受干扰的程度小且容易检测,光电流先经 I/V 转换电路后对微小的光电压信号进行放大、噪声滤波,再依次通过锁相放大器、A/D 转换后,最后把采集到的数据通过 USB 接口传输到电脑中保存,通过 Matlab 软件编程对数据进行模式识别并输出结果。因为敏感膜对不同离子的吸附作用不同,而味觉信号检测对象是多成分的,所以在传感器的绝缘层上需镀上多种敏感膜。用光束照射不同的敏感区可以检测到敏感膜在该区域所吸附离子的浓度,也就依次得到了味

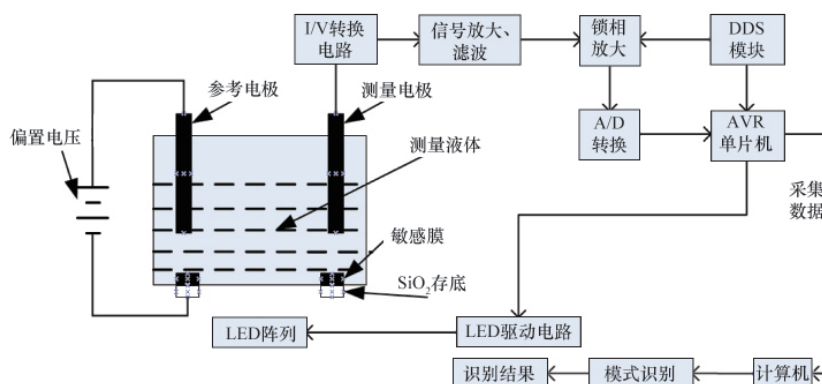


图 1 LAPS 型电子舌系统工作原理

Fig. 1 Operating principle of LAPS type electronic tongue

觉溶液中不同成分所对应的光电流信号。另外,采用的光源是阵列形式,可以同时对多块敏感膜照射,通过对频率的设定实现了检测时间短、多参数多点测量。

2 主元分析过程

PCA 法是在 Karhunen-Loeve(K-L)变换基础上,由 Turk 和 Pentlad 提出的一种无监督的多元统计模式识别方法。PCA 的主要思想是通过降维的方法把多个变量转化为几个主元,简化问题,使数据能够实现可视化。而且,只需要少量的几个主元就能够反映原始变量的大部分信息,这些主元为原始变量的线性组合^[9-11]。此方法得到的主成分轴是该数据矩阵方差最大的方向,且相互正交。第一主元含有的原始信息量最多,第二主元、第三主元次之,以此类推。

设有 n 个样品,每个样品有 k 项指标: $X_1, X_2, \dots, X_k$,得到

$$\text{样本矩阵 } X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1k} \\ x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix} = (X_1, X_2, \dots, X_k), \text{ 其中 } X_i = (x_{1i}, x_{2i}, \dots, x_{ni})^T (i=1, 2, \dots, k)。$$

主元分析算法对矩阵 X 进行特征提取的步骤^[12-15]如下。

- (1) 对矩阵 X 里的数据进行均值化处理,即 $z_{ij} = \frac{x_{ij} - \bar{x}_i}{\sigma_i} (j=1, 2, \dots, k)$, 其中, $\bar{x}_i = \frac{1}{k} \sum_{j=1}^k x_{ij} (i=1, 2, \dots, k)$, 由 z_{ij} 组成 1 个 Z 矩阵。

- (2) 求矩阵 Z 的相关系数矩阵,相关系数矩阵的公式为

$$C_Z = (ZZ^T)/n \quad (1)$$

得到的 C_Z 是 $n \times n$ 的矩阵。相关系数矩阵反应的是不同指标间

数据的相关性程度。

(3) 根据奇异值分解原理计算相关系数矩阵 C_Z 的特征值与特征向量,并对非零特征值由大到小进行排序 $\lambda_1 > \lambda_2 > \dots > \lambda_k (0 \leq k \leq n-1)$,对排序好的非零特征值以下公式求 n 值:

$$\frac{\sum_{k=1}^k \lambda_k}{\sum_{k=1}^{n-1} \lambda_k} \geq \beta \quad (2)$$

其中, β 为方差的累积贡献率,得到特征值所对应的特征向量即为主元(用 U 表示)。当 $\beta=0.85$ 时,表示前 k 个主元综合了原始信息 85% 的信息量。

(4) 对原始数据矩阵 X 在主元方向投影,得到矩阵 G

$$G=U^T X \quad (3)$$

其中, G 的每一列就是一种味道的压缩值。

(5) 求 5 组不同时间测得数据的压缩值,画出主元得分图,实现 5 种基本味道的样本溶液的分类识别。

(6) 为了能更充分、更直观地反映主元分析法的准确性,采用最小距离法,通过直方图更进一步说明。最小距离法^[6]公式为

$$d_i = \min \|x_i - m_p\| \quad (4)$$

其中, x_i 是酸、甜、苦、咸、鲜 5 种基本味道在主元方向上的投影向量; m_p 是待测的基本味道经过在主元方向上投影得到的向量。选定 1 个待测液体样本进行 LAPS 型电子舌味觉特征提取,将数据按列排列并在训练样本的主元方向上进行投影,得到矩阵 H ,并分别计算 H 与式(3)中得到的矩阵 G 中各列向量的距离。从各个距离中选择出最短距离后,即可知待测溶液属于哪种味道。

3 试验与结果分析

3.1 味觉检测

酸、甜、苦、咸、鲜 5 种基本味道样本溶液均是在实验室中分别通过柠檬酸、蔗糖、盐酸奎宁、氯化钠、管氨酸钠配置所得,浓度均为 0.02mol/L,现配现用,样本在试验前未经过任何特殊处理。每种味道溶液重复配制次数根据后面测试数据次数决定。

本文试验均采用 LAPS 型电子舌系统。测量之前,LAPS 型电子舌系统可以对采样频率进行设置以及自检,以确保采集数据的可靠性。每次检测前先用蒸馏水对容器进行洗涤,再将 100mL 的样本溶液倒入 LAPS 型电子舌系统的溶液池中。设置每个样本的采集时间为 10s,采样频率为 100Hz,样本溶液在相同测量环境和温度下进行通电测量,通过示波器进行观察。刚开始时感应强度有不同程度的起伏变化,待 LAPS 传感器响应强度稳定后,把检测到的数据以文本类型保存在计算机中,以待进行下步主元分析。

3.2 PCA 处理与分析

本文中光寻址型电子舌传感器阵列为 2×5 ,每次测量都可得到 10 组数据,通过试验分别测得 5 种基本味道溶液的

相关数据,其维数为 $1 \times n$ 。在图 2 中,提取的是某一照射区的 1000 个采样点。光照强度、温度等的微小变化对 LAPS 型电子舌检测到的数据可能有不同影响;此外,由于味觉检测时到达稳定状态所需的时间不同,需要对数据进行预处理,具体方法为根据测量周期以及每个周期的采样点数,通过在 Matlab 中对读入的每组数据进行自动检验,以使每个测量值与采样周期中的相应点对应,从而组成式(1)中的 X 矩阵。

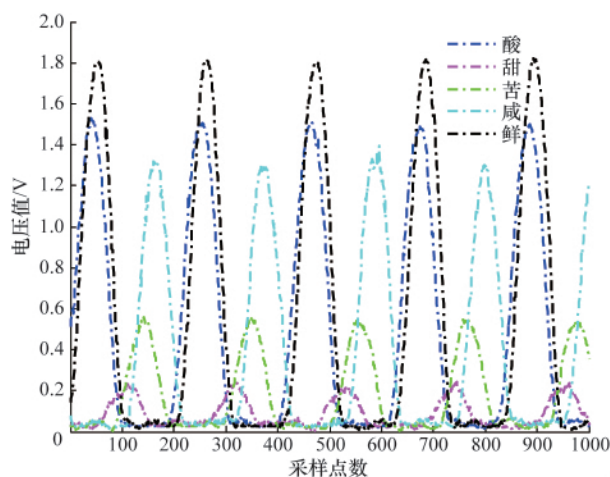


图 2 光寻址型电子舌测得的 5 种液体数据

Fig. 2 Five kinds of liquids' data detected by the light addressable electron tongue

同时,由于检测得到的电压值比较小,对 LAPS 型电子舌采集的 5 组样本数据组成的矩阵先放大 100 倍,增加数据之间的差异,然后根据主元分析步骤进行处理。求得第一主元的贡献率占 75.19%,第二主元的贡献率为 12.24%,累积贡献率为 87.43%。可见,这 2 个主元可以比较充分地反映原始数据的信息。

根据 PCA 识别算法同样可以求得 5 种基本味道主元的得分分布,如图 3 所示。其中,数据点代表样本的第一主元和第二主元得分,同种味道测了 8 个点。

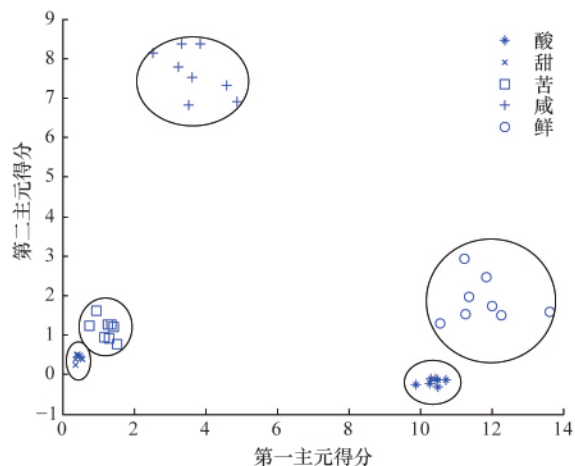


图 3 第一主元与第二主元得分

Fig. 3 First and second principle components scores

由图 3 可以看出,同种样本的 8 个点离散度较小,而不同种样本之间没有重叠,存在着不同程度的差异。其中,苦和甜、酸和鲜分布区域比较近,但没有互相干扰,而咸味与它们之间的距离较远,由此可知 LAPS 型电子舌的敏感膜对 5 种味道的反应程度,从而可以对 LAPS 型电子舌味觉信号进行识别。

为了能更准确地进行味觉识别,进行了 20 次试验并求取了每组中,5 种基本味道的主元,并对各种味道数据求平均值,得到一个新矩阵,再对其进行 PCA 处理,得到式(3)中的矩阵 G 作为识别模板。分别取 5 种味道不同的样本溶液 A、B、C、D、E 进行测试,根据上述方法求得矩阵 G, H , 求出二者

之间的距离。通过 Matlab 设定最小距离为红色。由图 4 可看出 5 种味道不同的样本溶液与 5 种已知溶液之间的差别程度(即纵坐标所表示的大小),红色所代表种类即为样本溶液 A(图 4(a))、B(图 4(b))、C(图 4(c))、D(图 4(d))、E(图 4(e))所属的味道。

为进一步验证算法和提高准确率,根据上述方法又进行了 200 次试验测量及味觉识别,识别结果统计表明,酸味的识别准确率为 95.6%,甜味的识别准确率为 96.1%,苦味的识别准确率为 96.3%,咸味的识别准确率为 98.7%,鲜味的识别准确率为 96.0%。

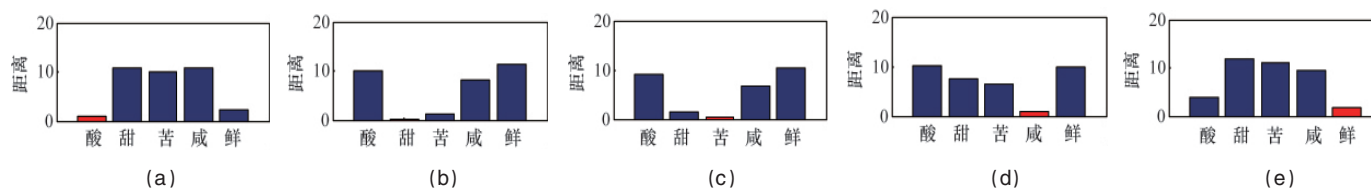


图 4 未知溶液与样本溶液之间的距离

Fig. 4 Distance between unknown solution and sample solution

4 结论

本文根据 LAPS 型电子舌原理,对酸、甜、苦、咸、鲜 5 种基本样本溶液味道进行信号特征提取。通过主元分析法对数据进行处理,能够消除数据的相关性和信息冗余性,从而判断出味觉之间的差异,由此可以对未知液体的味觉识别提供依据。从试验得到的第一、第二主元得分图中可以得出,LAPS 型电子舌可以将 5 个基本味道明显识别开。另外通过最小距离法,可以更直观地看到各味觉之间的差异程度。

参考文献 (References)

- [1] 张志勇, 陈兴梧, 牛文成, 等. 光寻址电位传感器测量系统的研究[J]. 仪表技术与传感器, 2003(12): 34-54.
Zhang Zhiyong, Chen Xingwu, Niu Wencheng, et al. *Instrument Technique and Sensor*, 2003(12): 34-54.
- [2] Vlasov Y G, Rudnitskaya A M, D'Amico A, et al. "Electronic tongue"-new analytical tool for liquid analysis on the basis of non-specific sensors and methods of pattern recognition [J]. *Sensors and Actuators B: Chemical*, 2000, 65(1-3): 235-236.
- [3] 程秀娟, 朱虹, 郑丽敏, 等. 电子舌技术在饮料识别中的应用 [J]. 食品科技, 2006, 31(8): 163-167.
Cheng Xiujuan, Zhu Hong, Zheng Limin, et al. *Food Science and Technology*, 2006, 31(8): 163-167.
- [4] 李毅, 王银瓶, 吝涛, 等. 船载海洋环境重金属检测的电子舌研究[J]. 传感器技术学报, 2007, 20(1): 32-37.
Li Yi, Wang Yiping, Lin Tao, et al. *Chinese Journal of Sensors and Actuators*, 2007, 20(1): 32-37.
- [5] 王俊, 姚聪. 基于电子舌技术的葡萄酒分类识别研究 [J]. 传感技术学报, 2009, 22(8): 1088-1093.
Wang Jun, Yao Cong. *Chinese Journal of Sensors and Actuators*, 2009, 22(8): 1088-1093.
- [6] Moreno L, Merlos A, Abramova N, et al. Multi-sensor array used as an "electronic tongue" for mineral water analysis [J]. *Sensors and Actuators B: Chemical*, 2006, 116(1-2): 130-134.

- [7] Bousse L, Mostarshed S, Hafeman D, et al. Investigation of cation transport through silicon wafers by photocurrents [J]. *J Appl Phys*, 1994, 75(8): 4000-4008.
- [8] 梁卫国, 韩经鸿. 光寻址电位传感器的机理研究[J]. 半导体光电, 2001, 22(3): 184-210.
Liang Weiguo, Han Jinghong. *Semiconductor Optoelectronics*, 2001, 22(3): 184-210.
- [9] 余映, 王斌, 张立明. 一种面向数据学习的快速 PCA 算法[J]. 模式识别与人工智能, 2009, 22(4): 567-573.
Yu Ying, Wang Bin, Zhang Liming. *Pattern Recognition and Artificial Intelligence*, 2009, 22(4): 567-573.
- [10] 范杰. 主成分分析法的数值实现算法 [J]. 河北工程大学学报, 2007, 24(4): 103-105.
Fan Jie. *Journal of Hebei University of Engineering*, 2007, 24(4): 103-105.
- [11] 帅斌, 杜文. 物流产业结构的 DEA/PCA 评价 [J]. 西南交通大学学报, 2006, 41(5): 599-602.
Shuai Bin, Du Wen. *Journal of Southwest Jiaotong University*, 2006, 41(5): 599-602.
- [12] 纪荣芳. 主元分析法中数据处理方法的改进 [J]. 山东科技大学学报, 2007, 26(5): 95-98.
Ji Rongfang. *Journal of Shandong University of Science and Technology*, 2007, 26(5): 95-98.
- [13] Palit M, Tudu B, Duta P K, et al. Classification of black tea taste and correlation with tea taster's mark using voltammetric electronic tongue [J]. *IEEE Transaction on Instrumentation and Measurement*, 2010, 59(8): 2230-2239.
- [14] Guru D S, Suraj M G, Manjunath S. Fusion of covariance matrices of PCA and FLD[J]. *Pattern Recognition Letters*, 2011, 32(3): 432-440.
- [15] 滕炯华, 王磊, 袁朝辉. 基于电子舌技术的果汁饮料识别 [J]. 测控技术, 2004, 23(11): 4-5.
Teng Jionghua, Wang Lei, Yuan Chaohui. *Measurement & Control Technology*, 2004, 23(11): 4-5.
- [16] Kranz H G. Diagnosis of partial discharge signals using neural networks and minimum distance classification [J]. *IEEE Trans on EI*, 1993, 28(6): 1016-1024.

(责任编辑 孙秀云)