

# 基于KFDA的食醋电子鼻鉴别方法\*

殷勇 薛俊莉 于慧春 陈红友 孟森娟

(河南科技大学食品与生物工程学院, 洛阳 471023)

**摘要:** 为了提高6种食醋的鉴别正确率,引入了基于核变换的Fisher判别分析(KFDA)方法,以及基于矩阵相似性的核参数确定方法。在选取径向基函数(RBF)为核函数,并提取食醋样本电子鼻检测信号的积分值和相对稳态平均值2种特征参量的基础上,优化确定了对应于2种特征参量的RBF核参数值,分别为5.770 0和5.387 8。对比分析了Fisher判别分析(FDA)与KFDA的鉴别结果,积分值与相对稳态平均值2种特征参量的鉴别结果分别由FDA的93.3%、90.6%提高到KFDA的98.3%、98.3%,说明合适的KFDA方法可有效提高6种食醋样品的鉴别结果。

**关键词:** 食醋 电子鼻 鉴别 核 Fisher 判别分析 Fisher 判别分析

**中图分类号:** TP212.2; TS207.5 **文献标识码:** A **文章编号:** 1000-1298(2014)09-0236-05

## 引言

用电子鼻<sup>[1-2]</sup>检测食品、农产品时,当鉴别样品品种增多、鉴别难度加大时,传统的线性模式识别方法,如主成分分析(PCA)<sup>[3-4]</sup>、Fisher判别分析(FDA)<sup>[5]</sup>等已不能很好地完成模式分析工作,需要一种改进的或新的模式识别方法。随着传统模式识别方法中核技术的引入,通过非线性映射将输入数据映射到一个高维的核空间中,然后在此空间中实施线性判别分析,可望较好地解决了这一问题<sup>[6]</sup>。目前,国内外的相关学者对基于核变换的线性判别方法进行了一些研究,如核主成分分析(KPCA)<sup>[7-9]</sup>、核Fisher判别分析(KFDA)<sup>[10-12]</sup>等。但多数都是KPCA和KFDA的理论研究以及在故障检测、人脸识别等方面的一些应用,在电子鼻检测食品中的应用报道尚不多见。为了较好地实现多种食醋样品的电子鼻鉴别,本文引入KFDA方法,同时选取RBF作为核函数,并采用基于距离测度的矩阵相似性度量方法优化确定RBF核函数参数。

## 1 试验材料与方法

### 1.1 试验仪器

试验使用的电子鼻系统由实验室自行研制。该系统主要由气敏传感器阵列、测量箱、A/D数据采集装置、稳压电源及计算机等几部分组成。传感器阵列由14只SnO<sub>2</sub>型金属氧化物气敏传感器组成,分别为TGS800、TGS812、TGS813、TGS816、TGS821、

TGS822、TGS824、TGS825、TGS826、TGS830、TGS831、TGS832、TGS880、TGS842。各传感器的型号和对应的敏感特征如表1所示。由预试验已知,这些传感器对不同食醋样品的挥发物有着不同程度的响应信号,同时考虑到这类气敏传感器具有广谱响应特性,因此这样的阵列信号可以用来表征检测样品的差异,以实现食醋样品的鉴别。

表1 传感器与典型敏感气体对应表

Tab.1 Typical detection gas of 14 sensors

传感器型号	敏感特征
TGS800	对通风设备控制系统中的烟、汽油蒸汽具有很高的敏感度
TGS812	对氢气、一氧化碳、甲烷、乙烷敏感
TGS813	对天然气、液化气敏感
TGS816	对烷类等可燃气体敏感
TGS821	对氢气有高选择和敏感性
TGS822	对酒精、甲苯、二甲苯等有机溶剂较敏感
TGS824	对氨气具有高敏感度
TGS825	对硫化氢有高度敏感性
TGS826	对氨气有高敏感度,对低浓度的氨气响应时间短
TGS830	对三氯三氟乙烷、二氟一氯甲烷、一氟三氯甲烷、二氟二氯甲烷的响应较快
TGS831	对一氟二氯甲烷、二氟一氯甲烷的响应较快
TGS832	对一氟三氯甲烷、二氟二氯甲烷、三氯三氟乙烷有高敏感度
TGS880	对食品中易挥发的气体和蒸汽,及烹调过程中的气体敏感
TGS842	对甲烷、丙烷、丁烷敏感,对杂质气体具有较低的敏感度

收稿日期: 2013-10-15 修回日期: 2013-12-01

\* 国家自然科学基金资助项目(31171685)

作者简介: 殷勇,教授,博士生导师,主要从事农产品、食品品质检测技术研究, E-mail: yinyong@haust.edu.cn

试验过程中,气敏传感器测量回路电压控制在 $(10 \pm 0.01)$  V 范围内,加热电压为 $(5 \pm 0.05)$  V。此外,该系统还配有温、湿度传感器各一个以测量环境温度、湿度的变化,用于补偿环境对气敏传感器的影响。

## 1.2 试验材料

为了探讨不同品牌、不同质量食醋的鉴别,试验选取建洛袋装与桶装米醋、来福瓶装与桶装陈醋、紫林瓶装与桶装陈醋等3个类别6种样品(如表2所示)进行了测试,后续的鉴别分析中,各种样品分别用JLA、JLB、LFA、LFB、ZLA、ZLB表示。

表2 试验材料的基本信息

Tab.2 Basic information of the vinegar samples

食醋名称	原料	总酸质量浓度/(g·L <sup>-1</sup> )	发酵方式	产地
建洛袋装	水、大米、大曲	≥35	固态发酵	河南
建洛桶装	水、小米、麸皮、大曲	≥35	液态发酵	河南
来福瓶装	水、高粱、麸皮、大麦、豌豆	≥40	液态发酵	山西
来福桶装	水、高粱、麸皮、大麦、豌豆	≥40	固态发酵	山西
紫林瓶装	水、高粱、麸皮、谷壳、稻壳、大麦、豌豆	≥45.1	固态发酵	山西
紫林桶装	水、高粱、麸皮、大麦、谷壳、稻壳、豌豆	≥35	固态发酵	山西

## 1.3 试验方法与试验结果

每种食醋样品采集30个样本,总计得到180个样本。测试前,先采集传感器对环境响应20 s内的电压(空载响应值),然后进行样本动态响应电压测试。数据采集时,无论是对环境测试还是对样本测试,均每隔1 s测量一次电压。样本测量时,用移液管量取5 mL样本于表面皿中,然后迅速将表面皿放入电子鼻测量箱进行数据采集。预试验可知,各传感器测试至1 300~1 500 s阶段时基本上处于相对稳定的状态,如图1给出了14个传感器对1个建洛桶装样品去基准处理后的响应曲线。所以,试验中每个样本采集时间定为1 500 s,即每个样本采集1 500个数据。样品采集结束后,进行传感器恢复,以便下一个样品的测试。同样,预试验可知,传感器在自然通风条件下恢复960 s可至初始状态,所以恢复时间设定为960 s。

为了减少环境温、湿度的影响,每个样本1 500个响应值都减去空载响应值的平均值,得到样本1 500个试验结果。

## 2 KFDA 简介与核参数确定

### 2.1 KFDA 简介

KFDA 方法,首先是将低维空间数据通过一非

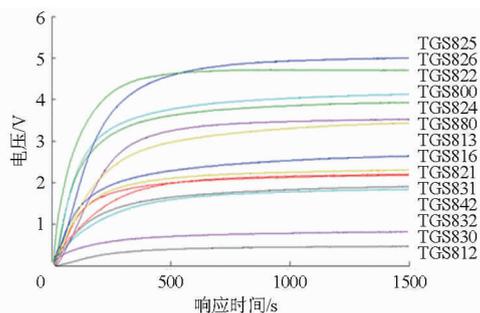


图1 14个传感器对建洛桶装样品的响应曲线

Fig.1 Response curves of 14 sensors to one barrelled Jianluo sample

线性核函数变换至高维数据空间。在实际变换中,并不需要对原空间数据进行直接的非线性映射,而是借助数据向量间的内积计算实现非线性变换<sup>[13]</sup>,形成核矩阵。然后在变换后的高维数据空间中(即核矩阵中)进行PCA数据融合。最后用FDA分析基于PCA的数据矩阵<sup>[14]</sup>,即可得到KFDA鉴别结果。

## 2.2 核矩阵与核参数确定方法

### 2.2.1 核矩阵

常用的核函数有多项式核函数、径向基核函数、Sigmoid核函数等,RBF核函数应用最为广泛,无论是低维、高维、小样本、大样本等情况,RBF核函数均适用,具有较宽的收敛域,是较为理想的分类依据函数<sup>[15]</sup>。故选取RBF核函数作为核变换函数,其表达式为

$$k(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{\eta}\right) \quad (1)$$

式中  $\mathbf{x}_i$ ——第*i*个数据向量

$\mathbf{x}_j$ ——第*j*个数据向量

$\eta$ ——待定核参数

不同的 $\eta$ 值会得出不同的核变换结果。所以, $\eta$ 参数的确定非常重要。

在确定RBF核函数为核变换函数后,假设 $\Phi$ 为一非线性映射函数,则原数据空间中两个数据向量 $\mathbf{x}_i, \mathbf{x}_j$ 的核变换结果可定义为

$$k(\mathbf{x}_i, \mathbf{x}_j) = \Phi(\mathbf{x}_i)^T \Phi(\mathbf{x}_j) = k_{ij} \quad (2)$$

$k(\mathbf{x}_i, \mathbf{x}_j)$ 称为数据向量 $\mathbf{x}_i, \mathbf{x}_j$ 之间的核函数值。对于原始数据矩阵,矩阵中所有行向量(这里1个行向量对应于1个样本)间的核函数值就构成了对应原始数据矩阵的核矩阵 $\mathbf{K} = \{k_{ij}\}$ 。

### 2.2.2 核参数确定方法

根据文献[16],称 $\mathbf{K}$ 矩阵为基于核函数变换的实际样本的Gram矩阵。该矩阵反映了在核函数导出的高维核矩阵空间中样本之间的相近关系。而一个理想的核函数可使同类的样本相互靠近,异类的样本相互远离,因此可定义一个理想Gram矩阵,其

构成元素为

$$k_{ij} = \begin{cases} 1 & (x_i = x_j) \\ 0 & (x_i \neq x_j) \end{cases} \quad (3)$$

定义这个矩阵的意义在于同类样本在高维核矩阵空间中的夹角都为 $0^\circ$ ,异类样本的夹角都为 $90^\circ$ 。当实际样本的 Gram 矩阵与此矩阵越相近,确定的 $\eta$ 核参数就越好。因此引入基于距离测度的矩阵相似性度量方法。计算公式为

$$S(A', B') = (A' - B')_{1 \times m^2} (A' - B')_{m^2 \times 1} \quad (4)$$

式中  $A'$ ——实际 Gram 矩阵对应的  $m^2$  维向量

$B'$ ——理想 Gram 矩阵对应的  $m^2$  维向量

当  $S(A', B')$  越小时,2 个向量越接近,即矩阵  $A$  和矩阵  $B$  越相似。此时,确定核参数的优化问题就变为:在式(3)前提下,确定核参数  $\eta$  使实际的 Gram 矩阵能让式(4)取得极小值。

### 3 结果与分析

#### 3.1 FDA 鉴别分析

特征参量鉴别能力的差异,影响分类的正确率,因此需要选取合适的特征参量。文献[17]中指出在食醋电子鼻鉴别中,特征参量为积分值时鉴别能力最高,特征参量为相对稳态平均值时鉴别能力次之。因此选择积分值和相对稳态平均值 2 个特征参量进行鉴别分析,由此形成了 2 个  $180 \times 14$  维的特征矩阵。另外,为了进一步补偿环境温、湿度的影响,把每个样本测试过程中的温度平均值、湿度平均值也分别作为模式识别系统的特征值输入,所以用于下文分析的特征矩阵变为 2 个  $180 \times 16$  维矩阵。特征参量的计算公式为

$$I = \sum_{i=1}^N x_i \Delta t \quad (5)$$

$$A = \frac{\sum_{i=t_0}^N x_i}{N - t_0} \quad (6)$$

式中  $I$ ——积分值  $A$ ——相对稳态平均值

$x_i$ ——1 个传感器对 1 个样本的第  $i$  个响应值

$\Delta t$ ——相邻 2 个响应值的间隔时间( $\Delta t = 1$  s)

$t_0$ ——稳态响应阶段初始点对应的时间( $t_0 = 1300$  s)

$N$ ——稳态响应阶段的终点对应的时间( $N = 1500$  s)

图 2、3 分别为积分值和相对稳态平均值 2 种特征参量下 6 种食醋样品的 FDA 分析结果,FD1、FD2 为 FDA 的第 1 投影轴和第 2 投影轴。

由图 2、3 可以看出,6 种样品大部分能够分开,但同品种样品较为分散,建洛桶装和来福瓶装、来福

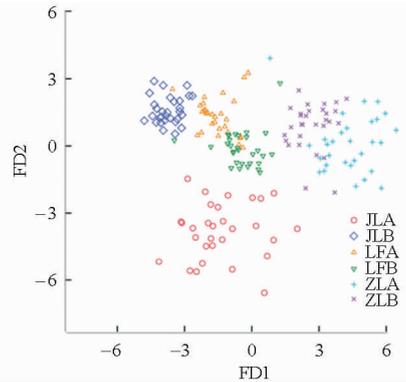


图 2 基于积分值的 FDA 鉴别结果

Fig. 2 FDA result of the integral value

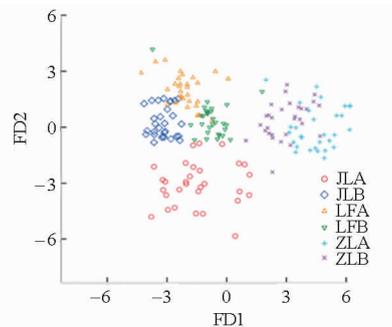


图 3 基于相对稳态平均值的 FDA 鉴别结果

Fig. 3 FDA result of average value in relative steady-state

桶装和来福瓶装、紫林瓶装和紫林桶装样品均有少量重叠。2 个特征参量的判别正确率分别是 93.3% (积分值) 和 90.6% (相对稳态平均值),交叉验证正确率分别为 83.3%、85.0%。

#### 3.2 KFDA 鉴别分析

在对特征矩阵实施 KPCA 后,进行 FDA 分析,具体方法为:

(1) 运用式(1)、式(3)分别计算实际 Gram 矩阵与理想 Gram 矩阵中的各元素,形成相应的 Gram 矩阵。

(2) 运用式(4)计算实际 Gram 矩阵与理想 Gram 矩阵间的欧氏距离,对含有核参数  $\eta$  的距离表达式进行极值计算,即可给出对应于积分值和相对稳态平均值特征矩阵下的 RBF 核参数  $\eta$ ,分别为 5.770 0 和 5.387 8。然后再运用式(1)计算对应 2 个特征核矩阵  $K_1$  (积分值)、 $K_2$  (相对稳态平均值),2 个  $K$  矩阵均为  $180 \times 180$  维的对称阵。

(3) 对核矩阵  $K_1$  和  $K_2$  分别进行中心化得到矩阵  $L_1$  和  $L_2$ 。

(4) 分别对  $L_1$  和  $L_2$  进行 PCA 计算。

(5) 当累积贡献率为 0.9 时,则分别提取了前 53 个主分量(积分值)和前 49 个主分量(相对稳态平均值),形成了 2 个 FDA 分析矩阵  $F_1$  和  $F_2$ 。

(6) 对  $F_1$  和  $F_2$  进行 FDA 分析,即实现了 KFDA

分析。

图4、5分别为积分值和相对稳态平均值2种特

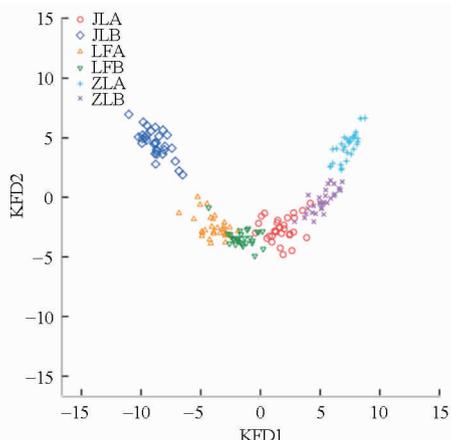


图4 基于积分值的KFDA鉴别结果

Fig.4 KFDA result of integral value

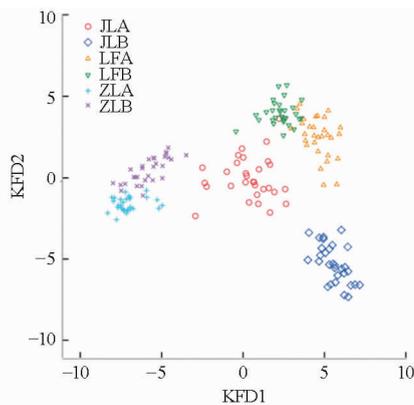


图5 基于相对稳态平均值的KFDA鉴别结果

Fig.5 KFDA result of average value in relative steady-state

征参量下6种食醋样品的KFDA分析结果,KFD1和KFD2为KFDA的第1投影轴和第2投影轴。

由图4、5可以看出同种样本相对比较聚集。尽管仍有少量样本重叠,但它们的判别正确率已分别达到98.3%(积分值)和98.3%(相对稳态平均值),与FDA的鉴别分析结果相比较,较为明显地提高了食醋的鉴别正确率。另外,交叉验证正确率也有了提高,分别为85.6%、89.4%。这说明KFDA方法确实可以提高食醋的鉴别正确率。

## 4 结论

(1)在选择RBF为核函数以及提取电子鼻检测食醋样品信号的积分值和相对稳态平均值2个特征参量的基础上,运用基于矩阵最佳相似性的方法优化确定了RBF核参数,分别为5.7700(积分值)和5.3878(相对稳态平均值)。

(2)分别运用FDA和KFDA对6种食醋进行了鉴别分析,鉴别正确率由FDA的93.3%(积分值)、90.6%(相对稳态平均值)提高到KFDA的98.3%(积分值)、98.3%(相对稳态平均值),交叉验证正确率也由FDA的83.3%(积分值)、85.0%(相对稳态平均值)提高到KFDA的85.6%(积分值)、89.4%(相对稳态平均值)。说明基于KFDA的食醋电子鼻鉴别方式是有效的。

## 参 考 文 献

- Eun Jeunghong, Sue Jeepark, Jin Yongchoi, et al. Discrimination of plam olein oil and palm stearin oil mixtures using a mass spectrometry based electronic nose [J]. Food Science and Biotechnology, 2011, 20(3): 809-816.
- 伍世元, 骆德汉, 邓炳荣, 等. 不同产地和采收期的中药材电子鼻鉴别研究[J]. 传感技术学报, 2011, 24(1): 10-13.  
Wu Shiyuan, Luo Dehan, Deng Bingrong, et al. Discrimination of Chinese herbal medicines from different growing areas and harvesting dates based on E-nose[J]. Chinese Journal of Sensors and Actuators, 2011, 24(1): 10-13. (in Chinese)
- Yin Y, Tian X. Classification of Chinese drinks by a gas sensors array and combination of the PCA with Wilks distribution[J]. Sensors and Actuators B, 2007, 124(2): 393-397.
- 史波林, 赵镭, 支瑞聪, 等. 西湖龙井茶品质的智能嗅觉识别[J]. 农业机械学报, 2012, 43(12): 130-135.  
Shi Bolin, Zhao Lei, Zhi Ruicong, et al. Quality recognition of Xihu Longjing tea based on intelligent olfactory [J]. Transactions of the Chinese Society for Agricultural Machinery, 2012, 43(12): 130-135. (in Chinese)
- 郭文川, 王铭海, 岳绒. 基于近红外漫反射光谱的损伤猕猴桃早期识别[J]. 农业机械学报, 2013, 44(2): 142-146.  
Guo Wenchuan, Wang Minghai, Yue Rong. Early recognition of bruised kiwifruit based on near infrared diffuse reflectance spectroscopy[J]. Transactions of the Chinese Society for Agricultural Machinery, 2013, 44(2): 142-146. (in Chinese)
- 孔锐, 张冰. 基于核 Fisher 判决分析的高性能多类分类算法[J]. 计算机应用, 2005, 25(6): 1327-1329.  
Kong Rui, Zhang Bing. New perfect performance multiclass classification algorithm based on KFDA [J]. Journal of Computer Applications, 2005, 25(6): 1327-1329. (in Chinese)
- 杜卓明, 屠宏, 耿国华. KPCA 方法过程研究与应用[J]. 计算机工程与应用, 2010, 46(7): 8-10.  
Du Zhuoming, Tu Hong, Geng Guohua. KPCA method research and application process [J]. Computer Engineering and Applications, 2010, 46(7): 8-10. (in Chinese)
- Giorgio M S. Accuracy of suboptimal solutions to kernel principal component analysis [J]. Computational Optimization and Applications, 2009, 42(2): 265-287.

- 9 胡金海,谢寿生,陈卫,等. 基于核函数主元分析的航空发动机故障检测方法[J]. 推进技术,2008,29(1):79-83.  
Hu Jinhai, Xie Shousheng, Chen Wei, et al. An aeroengine fault detection method based on kernel principal component analysis [J]. Journal of Propulsion Technology, 2008, 29(1):79-83. (in Chinese)
- 10 Li Junhong, Cui Peiling. Improved kernel fisher discriminant analysis for fault diagnosis[J]. Expert Systems with Applications, 2009, 36(2):1423-1432.
- 11 Xu Yong, Yang Jingyu, Lu Jianfeng. An efficient renovation on kernel Fisher discriminant analysis and face recognition experiments[J]. Pattern Recognition, 2004, 37(10):2091-2094.
- 12 裴昱,刘海林. 核 Fisher 鉴别分析在掌纹识别中的应用[J]. 量子电子学报,2009,26(6):647-653.  
Pei Yu, Liu Hailin. Kernel Fisher discriminant analysis used in palmprint recognition [J]. Chinese Journal of Quantum Electronics, 2009,26(6):647-653. (in Chinese)
- 13 徐勇,张大鹏,杨健. 模式识别中的核方法及其应用[M]. 北京:国防工业出版社,2010.
- 14 Yang Jian, Jin Zhong, Yang Jingyu, et al. Essence of kernel Fisher discriminant: KPCA plus LDA[J]. Pattern Recognition, 2004, 37(10):2097-2100.
- 15 林升梁,刘志. 基于 RBF 核函数的支持向量机参数选择[J]. 浙江工业大学学报,2007,35(2):163-167.  
Lin Shengliang, Liu Zhi. Parameter selection in SVM with RBF kernel function[J]. Journal of Zhejiang University of Technology, 2007, 35(2):163-167. (in Chinese)
- 16 李晓宇,张新峰,沈兰荪. 一种确定径向基核函数参数的方法[J]. 电子学报,2005,33(12A):2459-2463.  
Li Xiaoyu, Zhang Xinfeng, Shen Lansun. A selection means on the parameter of radius basis function[J]. Acta Electronica Sinica, 2005, 33(12A):2459-2463. (in Chinese)
- 17 于慧春,褚冰,殷勇. 食醋电子鼻检测中一种特征评价方法研究[J]. 农业工程学报,2013,29(3):1-8.  
Yu Huichun, Chu Bing, Yin Yong. Evaluation method of feature vector identification ability based on correlation coefficient in vinegar identification by electronic nose [J]. Transactions of the CSAE, 2013, 29(3):1-8. (in Chinese)

## Identification Method of Electronic Nose Based on KFDA for Different Vinegar Samples

Yin Yong Xue Junli Yu Huichun Chen Hongyou Meng Miaojuan

(School of Food & Bioengineering, Henan University of Science and Technology, Luoyang 471023, China)

**Abstract:** In order to enhance the correct rate of identification result of six kinds of vinegars, a kernel Fisher discriminant analysis (KFDA) method is introduced. And a measuring method of matrix similarity based on distance discrimination was presented to define the radial basis function (RBF) characteristic parameter, where RBF was selected as the kernel function. The measuring method is that, firstly, an ideal Gram matrix is defined, and the actual kernel Gram matrix is calculated by RBF; secondly, Euclidean distance can be employed to measure the degree of approximation between the actual kernel Gram matrix and the ideal Gram matrix; finally, the optimal kernel parameter can be obtained by extremal solution of the distance. When two kinds of feature vectors, whose were integral value and average value in relative steady-state, were extracted from the E-nose signals of vinegar samples, the corresponding characteristic parameters were 5.770 0 (integral value) and 5.387 8 (average value in relative steady-state). Comparing and analyzing the results of Fisher discriminant analysis (FDA) and KFDA, their identification correct rates were respectively from 93.3% and 90.6% (FDA) up to 98.3% and 98.3% (KFDA). This indicates that the suitable KFDA method can effectively improve the identification results of the six kinds of vinegar samples.

**Key words:** Vinegar Electronic nose Identification KFDA FDA