

Orthogonal Decomposition of Gas Sensor Array Signals Based on a Quasi-Legendre Basis *

ZHANG Wenna^{*}, QIN Guojun, HU Niaoqing

(College of Mechatronics Engineering and Automation, National University of Defense Technology, Changsha 410073, China)

Abstract: A new decomposition method based on an orthogonal quasi-Legendre basis is proposed, which decomposing the gas sensor array signals in the time domain to an alternative space spanned by a set of orthogonal functions. An electronic nose is applied to detect the volatile gas in the headspace of gear oil and machine oil. The signals are decomposed into five orthogonal quasi-Legendre functions, which preserve the meaningful information coming from the whole dynamic response. Accordingly, the resulted coefficients are used as fingerprints of the gas being tested. They are applied to build the PLS-CA discrimination models and PLS quantification models. The results show that two kinds of gases are clearly discriminated by the classifier, and the different concentrations of the gases are correctly predicted by the PLS calibration models. The feature extracting method based on quasi-Legendre basis is accurate and feasible, and it needs no any preprocesses, which provides a new way for the processing of sensor array signals.

Key words: gas sensor array; feature extraction; orthogonal decomposition; Quasi-Legendre basis

EEACC:7230

doi:10.3969/j.issn.1004-1699.2012.011.002

基于类 Legendre 基的气体传感器阵列信号的正交分解 *

张文娜^{*}, 秦国军, 胡萑庆

(国防科技大学机电工程与自动化学院, 长沙 410073)

摘要: 提出了一种基于类 Legendre 基的气体传感器阵列响应信号的正交分解方法, 将时域信号转换到由类 Legendre 基张开的空间。使用电子鼻系统对齿轮油和机油两类挥发气体进行了检测, 对响应幅值进行正交分解, 应用前 5 个基函数建立的模型能准确地描述整个动态响应过程。提取展开系数为特征值建立 PLS-CA 定性分类和 PLS 定量分析的模型。结果表明, PLS-CA 分离器实现了对两类样本的有效分类, PLS 定量模型可以准确地预测挥发气体的浓度。说明基于类 Legendre 基提取特征参数的方法准确可行, 在应用时不需任何预处理, 为气体传感器阵列信号的处理提供了一种新方法。

关键词: 气体传感器阵列; 特征提取; 正交分解; 类 Legendre 基

中图分类号: TP212.6

文献标识码: A

文章编号: 1004-1699(2012)11-1479-05

自从 1962 年 Seiyama 等首次提出应用半导体传感器检测气体以来, 气体传感器在气体检测中的发挥了重要的作用^[1-3]。由于单个氧化物半导体传感器存在选择性差和交叉敏感特性, 实际应用中选择多个敏感特性有差异的传感器组成阵列。在进行气体识别和定量分析时, 通常提取传感器阵列响应的特征数据值, 建立一个多元校正模型, 进行分析^[4-5]。特征提取是将传感器阵列的高维响应信号映射到低维空间, 同时保留了原始信号有用信息^[6]。因此, 提取有价值的特征信息值, 已成为该领域研究的一个热点。

通常的方法是提取尽可能多的参数值组成特征值集合, 在此基础上选择一个合适的子集。目前有多种提取参数值的方法: 如提取稳态值和初始响应值的差异^[7-8]; 从整个动态响应过程中提取均值、时间段的积分值等不同的参数^[9-10]; 应用系统识别理论^[11-12]; 以及利用其他信号处理方法如随机共振信噪比谱法^[13]、小波包分解法^[14]等。

由于气体传感器响应机理及其动力学特性复杂, 目前还缺乏通用的理论数学模型, 尤其是对于金属氧化物半导体气体传感器, 其响应特性不仅与接

项目来源: 国家自然科学基金项目(50975279)

收稿日期: 2012-07-03 修改日期: 2012-10-25

触气体的种类和浓度有关,还受环境的温湿度、工作温度、敏感层性质(传感器类型)的影响。对于这类问题,目前一种有效方法就是利用测量得到的响应信号对被分析气体进行描述,即对响应信号进行分解,得到若干个基本组成成分,即

$$x(t) = \sum_{i=1}^L c_i \phi_i(t) \quad (1)$$

式中, $\phi_i, i = 1 \sim L$ 表示基函数, c_i 是展开系数或称权值。

这种参数化表示方法有三个目的。一是把时间域的响应曲线变换成其它空间的若干个参数;二是可以用用时域更有效的方法处理变换域的参数^[15];三,由于这种信号变换意味着原信号和目标空间基函数的内积变换,所以展开系数是唯一的。当然,选择基函数非常重要,它决定着传感器响应信号映射的子空间。选择正交基函数对传感器信号变换最方便,从内积的角度来说,正交基使内积的计算变得非常简单。且由于分解的成分互相正交,可有效减少参数冗余和计算量。在气体吸附期和脱附期,金属氧化物半导体传感器的电阻值是朝两个不同方向变化的,响应曲线比较复杂。常用的傅里叶(Fourier)基函数是一种周期函数,其波形和气体传感器响应的复杂波形匹配程度低,所以需要找到一种更为有效的正交基函数。可用非周期函数进行信号变换,如多指数模型、多项式函数、小波(Wavelet)基函数等^[16-20],但多指数函数和多项式函数均非正交基,而 Wavelet 基函数存在与 Fourier 基函数同样的不足。文献[15,21]提出一种基于贝塞尔-傅里叶基函数的正交分解方法,相对 Wavelet 基函数与 Fourier 基函数而言,其基函数波形和传感器响应曲线比较匹配,利用展开系数作为特征参数,对气体进行了有效分析。

本文提出利用类 Legendre 正交基函数来建立气体传感器响应模型。不同于 Fourier 变换和 Wavelet 变换,类 Legendre 正交基函数不是周期函数^[22],且分析的不是局部时间或频率的值,而是随时间而变化的幅值,具有类似和传感器响应值的结构信息,所以完全可以用来描述传感器响应过程中的幅值减小或增加。由于是正交分解,因此可以剔除与响应值无关的噪声信号。

1 理论基础

1.1 类 Legendre 基

定义^[22] 在一般的代数三角函数空间 $\Gamma_n = \text{span}\{1, t, t^2, \dots, t^{n-4}, \text{sint}, \text{cost}, \text{sin}2t, \text{cos}2t\}$ 中的一组正交基 $\{A_n(t)\}$ 被称为类 Legendre 基,其定义式

如下:

$$\begin{cases} A_0(t) = 1 \\ A_1(t) = \sin\left(t - \frac{\alpha}{2}\right) \\ A_2(t) = \alpha \cos\left(t - \frac{\alpha}{2}\right) - 2\sin\left(\frac{\alpha}{2}\right) \\ A_3(t) = \sin(2t - \alpha) - A_1 \sin\left(t - \frac{\alpha}{2}\right) \\ A_4(t) = [\cos(2t - \alpha) - A_2] - A_3 \left[\alpha \cos\left(t - \frac{\alpha}{2}\right) - 2\sin\left(\frac{\alpha}{2}\right) \right] \\ \vdots \\ A_k(t) = \frac{d^{k-4}}{dt^{k-4}} M_{k-2}(t), \quad k=5, 6, \dots, n \end{cases} \quad (2)$$

式中, $A_1 = \frac{\int_0^\alpha \sin(2t - \alpha) \sin\left(t - \frac{\alpha}{2}\right) dt}{\int_0^\alpha \sin^2\left(t - \frac{\alpha}{2}\right) dt}$,

$$A_2 = \int_0^\alpha \frac{\cos(2t - \alpha)}{\alpha} dt,$$

$$A_3 = \frac{\int_0^\alpha \cos(2t - \alpha) \left[\cos\left(t - \frac{\alpha}{2}\right) - \frac{\alpha}{2} \sin \frac{\alpha}{2} \right] dt}{\int_0^\alpha \left[\cos\left(t - \frac{\alpha}{2}\right) - \frac{\alpha}{2} \sin \frac{\alpha}{2} \right]^2 dt}.$$

令 $\beta = s - \frac{\alpha}{2}, U_k^{2k}(s) = u_{k,2k}(s), U_{k-i}^{2k}(s) = u_{k-i,2k}(s) + u_{k+i,2k}(s), i = 1, 2, \dots, k_0$

则 $M_k(t) =$

$$\begin{vmatrix} \int_0^\alpha U_{k-1}^{2k}(s) \cos\beta ds & \int_0^\alpha U_k^{2k}(s) \cos\beta ds & \int_0^\alpha U_{k-2}^{2k}(s) \cos\beta ds \\ U_{k-1}^{2k}(t) & U_k^{2k}(t) & U_{k-2}^{2k}(t) \\ \int_0^\alpha U_{k-1}^{2k}(s) \cos 2\beta ds & \int_0^\alpha U_k^{2k}(s) \cos 2\beta ds & \int_0^\alpha U_{k-2}^{2k}(s) \cos 2\beta ds \end{vmatrix} \quad (3)$$

$A_1(t) \sim A_4(t)$ 分别如图 1(a) ~ (d) 所示。

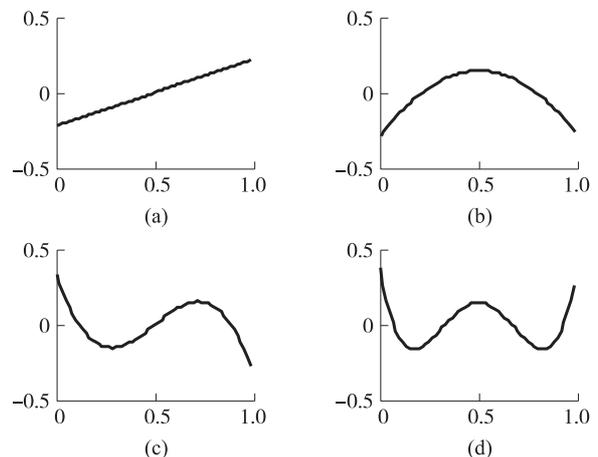


图 1 类 Legendre 基函数示意图

1.2 Γ_n 空间上传感器响应信号特征提取

对于实际传感器响应信号的采样值 $x(t_k)$, $t_k \in (0, a)$, $k = 1, 2, \dots, N$ 。应用类 Legendre 基函数将 $x(t_k)$ 进行正交分解为

$$x(t_k) = \sum_{i=1}^L c_i A_i(t_k) \quad (4)$$

式中, c_i 是类 Legendre 展开系数。

为了获得最优逼近, 采用最小二乘法, 由下式求出 c_i

$$c = (\phi\phi^T)^{-1} \phi x \quad (5)$$

式中, $x = [x(t_1), x(t_2), \dots, x(t_N)]^T$; 矩阵 ϕ 的元素为 $\phi_{ik} = A_i(t_k)$; 向量 $c = [c_1, \dots, c_L]^T$ 中 L 个元素为展开系数, 且唯一确定, 可用作分析气体的指纹图。相当于用代数三角函数空间 Γ_n 的 L 个值 (即为展开系数) 描述了在时域上的传感器对气体的响应。本文取前 5 个类 Legendre 基函数进行分解, 因为第 6 项以后基函数的求解很复杂, 且由前 5 项基已获得了很佳的传感器响应曲线拟合效果。

2 实验系统与实验材料

2.1 电子鼻系统

实验采用德国 AIRSENSE 公司 PEN3 型电子鼻。该电子鼻包含 10 个不同的金属氧化物半导体传感器, 各个传感器的名称及性能描述如表 1 所示。传感器工作在高温下 ($200\text{ }^\circ\text{C} \sim 500\text{ }^\circ\text{C}$), 气体在半导体材料颗粒表面的吸附和脱附可改变半导体元件的电导率。

表 1 PEN3 所用传感器及其性能描述

序号	传感器名称	性能描述
1#	W1C	对芳香物族化合物灵敏
2#	W5S	灵敏度大、对氮氧化物很灵敏
3#	W3C	对氨等芳香物族化合物灵敏
4#	W6S	主要对氢气有选择性
5#	W5C	对烷烃、芳香物族化合物灵敏
6#	W1S	对甲烷灵敏
7#	W1W	对硫化物灵敏
8#	W2S	对乙醇灵敏
9#	W2W	对芳香成分、有机硫化物灵敏
10#	W3S	对高浓度烷烃灵敏

2.2 实验材料

利用注射器分别抽取 100 mL 齿轮油 (G) 和机油 (M) 两种液状油品, 注入到 500 mL 带橡胶塞的玻璃瓶中, 在室温下 ($(25 \pm 1)\text{ }^\circ\text{C}$) 密封静置 2 h, 顶空富集油液挥发成分作为原始气体初样。采用顶空抽气方法, 用注射器从柴油和齿轮油初样中分别抽取 12 种不同容积 (1 mL, 5 mL, 10 mL, 20 mL, 30 mL, 40 mL, 50 mL, 60 mL, 70 mL, 80 mL, 90 mL, 100 mL) 的气体, 转

注到 250 mL 的空瓶中密封作为实验样本, 这样共制备 24 个样本。对样本进行编号: 柴油为 1~12, 齿轮油为 13~24。

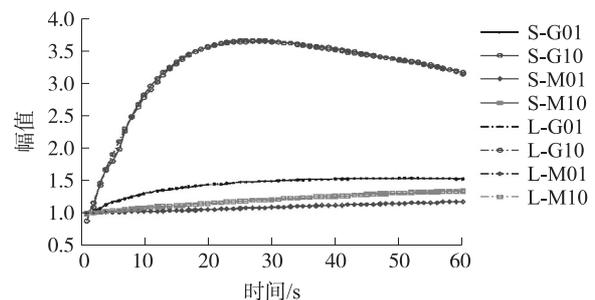
实验时, 保持温度和湿度恒定, 电子鼻经过 0.5 h 的预热稳定后, 进行“采样吸入气体—清洗传感器阵列—稳定系统”的循环过程, 采样频率为 1 Hz, 每个样本采集 60 个点。

3 实验结果与分析

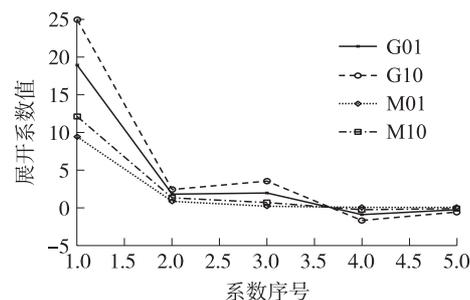
3.1 传感器响应信号正交分解和重构

10 个气体传感器对油液挥发气体样本都有响应, 但响应幅度不同, 对传感器阵列的响应幅值应用前 5 个类 Legendre 基函数进行正交分解。则对时域上的每条响应曲线的 60 个采样点, 可由在类 Legendre 基张开的空间上 Γ_n 的 5 个展开系数 ($c = [c_1, \dots, c_5]^T$) 来表征。当用气体传感器阵列对 24 个样本进行测试时, 测试结果可由一个 24×50 的矩阵表示。

图 2 是 1# 传感器对 1 mL、10 mL 的柴油和机油挥发气体样本 (分别记为 G01、G10、M01、M10) 的响应曲线的正交分解结果。图 2(a) 中实线为传感器响应的实测曲线 (用前缀 S 表示), 虚线为正交变换后的重构曲线 (用前缀 L 表示), 可知, 经过类 Legendre 基分解后的重构曲线和原始响应曲线拟合程度比较高, 且由图 2(b) 四个样本响应信号的正交分解展开系数可知, 两类不同的油液的展开系数有一定程度上的区别。



(a) 传感器响应曲线及类 Legendre 基重构曲线



(b) 正交分解展开系数

图 2 传感器响应信号类 Legendre 基正交分解

3.2 基于 PLS-CA 的油液定性分析

偏最小二乘法(PLS)是化学计量中常用的分析方法,PLS 能克服数据间的共线,并且它不仅利用了自变量即观测矩阵 X 的信号,同时还利用了目标测量值 Y 的信息,相当于同时对 X 和 Y 进行主成分分解,并在分解时考虑两者之间的相互关系,加强对应计算关系,从而保证获得最佳的校正模型。

本实验数据分析中,主成分数的确定如下:取主成分数从 1 到 12 依次增加,进行 PLS 计算,取校正均方根误差值(RMSEC)最小值对应的主成分数。

$$RMSEC = \sqrt{\sum_{i=1}^n (y_i - y_{ii})^2 / (n - 1)} \quad (6)$$

式中, y_i 为目标值, y_{ii} 为预测值。

在对两种油液的定性分析中,设置目标值为

$$Y = [y_G; y_M];$$
$$y_G = [1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0];$$
$$y_M = [0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1].$$

y_G 、 y_M 分别表示传感器对齿轮油和机油挥发气体的测定值。对于每一个测试样本的输入值会得到一个对应的行向量,行向量中值最大的元素的位置 i ($i=1,2$),表示这个样本属于第 i 种气体。RMSEC 最小对应的主成分数为 12。由 PLS 分析得到两类油液 24 个样本的预测值 y_{iG} 、 y_{iM} 如图 3 所示。对于齿轮油样本, y_{iG} 大, y_{iM} 小;对于机油样本则反之,可知两类样本明显区分开,并近似按直线分布。

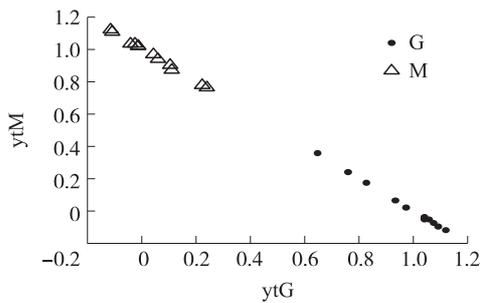


图 3 24 个样本 PLS 定性分析结果

把 PLS 计算结果进行聚类分析(CA),样本间的距离采用欧式距离,分析得到的树形图如图 4。从聚类结果可知,齿轮油(1~12 号)和机油(13~24

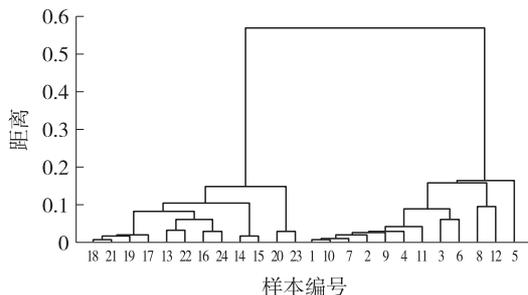


图 4 24 个样本的 CA 树形图

号)两类样本很好分离开来,正确率为 100%。

3.3 基于 PLS 的油液定量分析

为进一步验证传感器响应信号类 Legendre 基正交分解的有效性,下面分别对不同浓度的油液进行定量分析。利用 T_n 空间上 24×50 的矩阵作为 PLS 预测模型的输入 X ,目标浓度值为向量

$$Y = [1 \ 5 \ 10 \ 20 \ 30 \ 40 \ 50 \ 60 \ 70 \ 80 \ 90 \ 100]$$

计算得到主成分数为 10。PLS 定量分析结果如表 2,油液浓度很低(1 mL)时,传感器阵列的响应值小,测得实验数据误差相对较大,模型预测误差大,浓度增大到 10 mL 时,预测精度达到 95% 以上。为了更好地评估预测模型的有效性,进一步建立实际值 Y 和预测值 Y_i 之间的线性回归方程,齿轮油和机油的回归方程分别为

$$Y_{iG} = 0.9999Y + 0.0015$$

$$Y_{iM} = 0.9995Y + 0.0024$$

相关系数分别为 $R_G = 0.9999$, $R_M = 0.9997$ 。

表 2 两种油液样本浓度预测结果

实际值 Y/mL	齿轮油			机油		
	样本 编号	预测值 Y_{iG}/mL	预测误 差/%	样本 编号	预测值 Y_{iM}/mL	预测误 差/%
1	1	1.1246	12.4553	13	1.1546	15.4619
5	2	4.5712	-8.5767	14	4.6335	-7.3303
10	3	10.2857	2.8573	15	10.4275	4.2755
20	4	20.1438	0.7191	16	20.6036	3.0215
30	5	30.0359	0.1197	17	29.1293	-2.2901
40	6	39.8128	-0.4680	18	40.2899	0.7247
50	7	49.9360	-0.1280	19	50.5892	1.1785
60	8	60.0328	0.0546	20	59.6678	-0.5537
70	9	70.1311	0.1873	21	69.7479	-0.3602
80	10	79.9002	-0.1248	22	79.6736	-0.4080
90	11	90.0307	0.0341	23	90.4599	0.5110
100	12	99.9961	-0.0039	24	100.0963	0.0963
		$Y_{iG} = 0.9999Y + 0.0015$			$Y_{iM} = 0.9995Y + 0.0024$	
		$R_G = 0.9999$			$R_M = 0.9997$	
		$RMSEC_G = 0.0465$			$RMSEC_M = 0.0990$	

4 结论

本文采用电子鼻系统对齿轮油和机油两类挥发气体各 12 个不同浓度的样本进行了检测,对响应幅值分别进行基于类 Legendre 基的正交分解,应用前 5 个正交基函数建立的模型能较好地重构原始响应曲线。提取前 5 个展开系数作为描述传感器响应的特征值,即把传感器响应的 60 个参数的时域信息转

换到仅有 5 个参数所表征的类 Legendre 基所张成的空间中。对特征值进行 PLS-CA 定性分类和 PLS 定量分析, 结果表明, PLS-CA 分离器实现了对两类样本的有效分类, 正确率为 100%, PLS 定量模型可以准确地预测挥发气体的浓度, 浓度较高时预测的精度达到 95% 以上。

本文研究表明, 基于类 Legendre 基正交分解提取的参数, 能准确的表征传感器的响应信息, 且在应用时不需任何预处理, 结合模式识别方法即可取得很好的定性分类和定量分析的效果。

参考文献:

- [1] Rudneva V A, Boichenkob A P, Karnozhytskiya P V. Classification of Gasoline by Octane Number and Light Gas Condensate Fractions by Origin with Using Dielectric or Gas-Chromatographic Data and Chemometrics Tools[J]. *Talanta*, 2011, 84:963-970.
- [2] Wilson A, Baietto M. Applications and Advances in Electronic-Nose Technologies[J]. *Sensors*, 2009, 9(7):5099-5148.
- [3] Ryan M A, Buehler H Z, Manatt M G, et al. Monitoring Space Shuttle Air Quality Using the Jet Propulsion Laboratory Electronic Nose[J]. *IEEE Sensors Journal*, 2004, 4(3):337-347.
- [4] Pavlou A K, Turner A P. Sniffing out the Truth: Clinical Diagnosis Using the Electronic Nose[J]. *Clinical Chemistry and Laboratory Medicine*, 2000, 38(2):99-112.
- [5] Martín Y G, Oliveros M C C, Pavón J L P, et al. Electronic Nose Based on Metal Oxide Semiconductor Sensors and Pattern Recognition Techniques: Characterizations of Vegetable Oils[J]. *Analytical Chimica Acta*, 2001, 449(1-2):69-80.
- [6] Gutierrez-Osuna R, Nagle H T. A Method for Evaluating Data-Preprocessing Techniques for Odor Classification with an Array of Gas Sensors[J]. *IEEE Trans. Syst. Man Cybern*, 2011(B29):626-632.
- [7] Pardo M, Sberveglieri G. Comparing the Performance of Different Features in Sensors Arrays[J]. *Sensors and Actuators B*, 2007, 123(2):437-443.
- [8] 邹小波, 赵杰文. 基于小波去噪和支持向量机的苹果品种识别法[J]. *仪器仪表学报*, 2007, 3(28):534-538.
- [9] Skov T, Bro R. A New Approach for Modeling Sensor Based Data[J]. *Sensors and Actuators B*, 2005, 106(2):719-729.
- [10] 于慧春, 王俊. 电子鼻技术在茶叶品质检测中的应用研究[J]. *传感技术学报*, 2008, 21(5):748-752.
- [11] Holmberg M, Davide F, Di Natale C, et al. Drift Counteraction in Odour Recognition Applications[J]. *Sensors And Actuators B-Chemical*, 1997(42):185-194.
- [12] Vergara A, Llobet E, Brezmes J, et al. Quantitative Gas Mixture Analysis Using Temperature-Modulated Micro-Hotplate Gas Sensors: Selection and Validation of the Optimal Modulating Frequencies[J]. *Sensors And Actuators B-Chemical*, 2007(123):1002-1016.
- [13] 惠国华, 陈裕泉. 基于电子鼻的低温贮藏草鱼品质预测方法研究[J]. *传感器技术学报*, 2012, 25(2):157-161.
- [14] 张宏顺, 张仲欣, 殷勇. 食醋分类中气体传感器阵列的一种优化方法[J]. *传感器与微系统*, 2008, 27(8):75-78.
- [15] Vergara A, Martinelli E, Huerta R, et al. Orthogonal Decomposition of Chemo-Sensory Cues[J]. *Sensors and Actuators B-Chemical*, 2011(159):126-134.
- [16] Yang Y S, Ha S C, Kim Y S. A Matched-Profile Method for Simple and Robust Vapor Recognition in Electronic Nose(E-Nose) System[J]. *Sensors and Actuators B*, 2005, 106(1):263-270.
- [17] Leone A, Distante C, Ancona N, et al. A Powerful Method for Feature Extraction and Compression of Electronic Nose Responses[J]. *Sensors and Actuators B*, 2005, 105(2):378-392.
- [18] Martinelli E, Falconi C, Amico A D, et al. Feature Extraction of Chemical Sensors in Phase Space[J]. *Sensors and Actuators B-Chemical*, 2003(95):132-139.
- [19] Gutierrez-Osuna R, Nagle H T, Schiffman S S. Transient Response Analysis of an Electronic Nose Using Multi-Exponential Models[J]. *Sensors and Actuators B*, 1999:170-182.
- [20] 殷勇, 周秋香, 于慧春, 等. 有机磷农药气敏传感阵列检测信号小波包降噪方法[J]. *农业机械学报*, 2011, 42(4): - .
- [21] Vergara A, Martinelli E, Huerta R, et al. Orthogonal Decomposition of Chemo-Sensory Signals: Discriminating Odorants in a Turbulent Ambient[J]. *Procedia Engineering*, 2011(25):491-494.
- [22] 周志敏, 汪国昭. 一种类 Legendre 基及其应用[J]. *浙江大学学报(理学版)*, 2006, 33(4):398-402.



张文娜(1976-), 女, 副教授, 在职博士, 现于国防科技大学机电工程与自动化学院任教。研究方向为新型传感器技术、模式识别与信号处理技术, zwna@nudt.edu.cn;



胡葛庆(1967-), 男, 教授, 博士生导师, 现于国防科技大学机电工程与自动化学院任教。研究方向为机器状态监控与故障诊断、先进信号处理技术等, hngq@nudt.edu.cn。