

基于人工嗅觉系统的油液渗漏检测方法*

张文娜, 秦国军, 胡葛庆, 张 纪

(国防科技大学 机电工程与自动化学院, 湖南 长沙 410073)

摘要:机电设备油液渗漏是一种典型的多发性故障,利用人工嗅觉技术对渗漏油液挥发气体进行测试从而进行故障诊断是一种新的无损检测方法。使用人工嗅觉系统对模拟柴油、齿轮油和机油渗漏产生的挥发气体进行检测,结果为三维数据阵列(样本 \times 时间 \times 传感器)。应用二维主成分分析法和三维平铺主成分分析法、平行因子分析方法对三种油液挥发气体样本进行分类,结果表明平行因子分析法由于利用了数据集的三维结构信息,所以分类效果更佳;应用主成分回归方法实现了机油挥发气体样本的定量确定,说明使用人工嗅觉系统实现设定阈值报警是可能的。

关键词:人工嗅觉系统;油液渗漏;主成分分析;平行因子分析;主成分回归

中图分类号:TP212 **文献标志码:**A **文章编号:**1001-2486(2012)06-0175-06

Oil leakage detection based on the artificial olfactory system

ZHAGN Wenna, QIN Guojun, HU Niaoqing, ZHAGN Ji

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

Abstract: Oil leakage from cracks is a kind of common fault in mechatronics systems. Responses of artificial olfactory system to the volatile organic compounds (VOCs) emitted by leakage oil can be used for leak detection, which is a novel non-destructive method for fault diagnosis. An artificial olfactory system was applied to detect three different kinds of leakage oil such as diesel oil, machine oil and gear oil. The collected dataset was arranged in a three-dimensionality matrix (sample \times time \times sensor). Three methods as two-way principal component analysis (PCA), three-way unfolding PCA and parallel factor analysis (PARAFAC) were adopted to distinguish the VOCs. The results showed that PARAFAC was superior to the other two methods because PARAFAC took into account the true three-dimensionality structure of the dataset. The multivariate calibration method, principal component regression (PCR) was applied in the prediction of different concentrations of diesel oil. The results indicate that it is possible to use an artificial olfactory system to give an alarm by setting a threshold.

Key words: artificial olfactory system; oil leakage; principal component analysis (PCA); parallel factor analysis (PARAFAC); principal component regression (PCR)

机电设备在制造中难免存在各种缺陷。随着服役年限的增加,构件固有的微小缺陷以及在使用中产生的疲劳和应力腐蚀裂纹,将会造成燃油和滑油等油液微渗漏,这是一种典型的重复多发性故障。不仅会带来经济损失和环境污染,还可能发生火灾和爆炸,造成人员伤亡事故。油液渗漏的检测和定位一直是机械设备使用、维护和故障诊断的一个难点问题。目前国内外对机电设备的油液渗漏诊断主要采用差压检测、光学、热学、超声、涡流及射线等检测方法^[1-4]。但由于渗漏造成的油压变化小、可能渗漏点多,所以在灵敏度、定位能力、响应速率等方面不能满足实时监测的要求。例如,利用声的发射及反射原理对渗漏的部位进行检测,结果受外界的影响大,误报率高;利用X光等探测法进行无损检测,在线性较

差,不仅检测时间长,且设备的价格昂贵,不易推广。同时,这些方法不能具体检测出是何种油液发生了渗漏,无法定位故障源。

一般来讲,渗漏的油液必然会挥发一定的气体,因此,可望通过人工嗅觉技术,利用气体传感器阵列对挥发气体进行识别,达到油液渗漏无损检测的目的。人工嗅觉系统是受生物嗅觉机理的启发,将传感器阵列技术、电子技术和模式识别技术等工程手段紧密结合研制成的新颖仿生检测仪器^[5-6]。气体传感器阵列由一具有组交叉敏感特性的气体传感器组成,用于感受气体。信号预处理功能主要实现响应信号的调理,而模式识别系统则相当于大脑,对嗅觉信息进行判断、分析及识别。基于人工嗅觉系统的方法检测速度快、对要检测的样品不要复杂的预处理、操作简单,还可以

* 收稿日期:2012-04-28

基金项目:国家自然科学基金资助项目(50975279)

作者简介:张文娜(1976—),女,河南信阳人,副教授,在职博士研究生,E-mail:zwna@nudt.edu.cn

完成实时在线监测。目前已在食品、环境、医疗卫生、安保、军事等方面得到了广泛应用^[7-11]。

本文通过人工嗅觉系统获取模拟泄漏的汽油、柴油和机油三种油液挥发气体的“指纹图”，应用多种方法对三维测量数据阵列进行处理，实现了三种油液的定性判别，并利用主成分回归方法实现油液的定量检测。

1 实验方法

1.1 实验样品制备

用注射器分别抽取 100ml 柴油(D)、机油(M)、齿轮油(G)三种液状油品，注入 250ml 带橡胶塞的玻璃瓶中，密封在室温下静置 1h，待玻璃瓶内油液充分挥发达到稳定后作为初样。再采用顶空抽气的方法，从初样中抽取一定容积的气体后转注到 250ml 的实验空瓶中密封，作为一种浓度的实验样本。

1.2 实验系统及实验方法

实验采用德国 AIRSENSE 公司 PEN3 型便携式电子鼻，主要包含传感器阵列、采样及清洗通道、数据采集系统及计算机三个部分，系统结构与检测原理如图 1 所示。传感器阵列包含 10 个不同的掺有杂质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详见文献[12]。根据传感器接触到样品挥发物后的电导率 G 与在经过标准活性炭过滤的洁净空气中的电导率 G_0 的比值 G/G_0 进行数据预处理和模式识别。

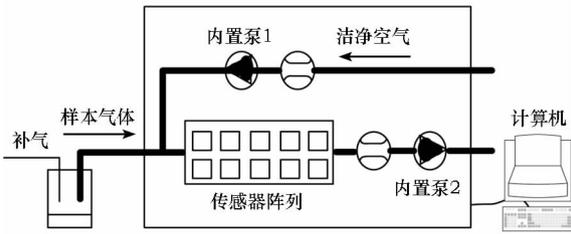


图 1 PEN3 电子鼻的结构

Fig. 1 Schematic diagram of PEN3

实验时通过内置泵 1 样品气体从进口处被吸附到传感器阵列反应腔中，每次测量前后，过滤后的洁净空气通过内置泵 2 流到反应腔中，对传感器阵列进行清洗和标准化，以保证电子鼻测量数据的稳定性和精度，有效消除累积效应。电子鼻测定条件为，采样时间间隔为 1s，传感器自清洗时间为 100s，测试时间为 60s。由此整个实验过程得到的响应数据是一个三维数据阵列 $X(I \times J \times K)$ ， I 为气体样本个数， J 为采样时间点数， K 为传感器只数。

1.3 数据预处理

根据模式识别理论，只有各个独立变量间的协方差很高，才能取得很好的分类效果。为了均衡各变量对样本总体方差的贡献，使各种气味能显著地分开，首先对样本数据经过一些预处理。

(1) 平滑处理

数据采集过程中，不可避免地受到随机噪声的影响，本文采用 6 点平滑处理，算子为

$$G'(i) = \frac{1}{h-2} \cdot \left[\sum_{j=0}^{h-1} G(i+j) - \max_{0 \leq j \leq h-1} G(i+j) - \min_{0 \leq j \leq h-1} G(i+j) \right] \quad (1)$$

其中， $h = 6$ 。

(2) 基线修正

传感器在初始状态下输出为 G_0 而并不是 0，一般把这时的输出值称为传感器的基线值。采用差分法对传感器响应进行基线修正，即 $(G - G_0)/G_0$ ，同时有助于补偿传感器的温度、湿度灵敏度漂移。

(3) 归一化处理

为了提高对不同气体的分类辨识效果，对传感器阵列的输出信号进行归一化预处理，使每个传感器的输出处于某个较小的区间，以消除或降低气味浓度对传感器输出的影响。本文采用对样本模式进行归一化处理如下

$$x_{ijk} = \frac{x_{ijk}}{\sqrt{\frac{1}{JK} \sum_{j=1}^J \sum_{k=1}^K x_{ijk}^2}} \quad (2)$$

式中， x_{ijk} 为采样时间 j 时，第 k 只传感器对第 i 个样本的响应值。

1.4 模式识别方法

(1) 主成分分析法和平铺主成分分析法

主成分分析(PCA)是一种双线性非监督模式识别方法，已成功应用于人工嗅觉系统中^[8]。该方法的实质是寻找在最小均方意义下最能代表原始数据的投影方法，实现对数据的转换和降维，用新的 $F(F < K)$ 个主成分变量来表示原来 K 只传感器的响应信号。平铺主成分分析法(Unfolding PCA)把传感器阵列响应的的时间轴的所有数据进行拼接，将三维数据阵转换为二维 $I \times (J \times K)$ ，进行 PCA 分析^[13]。

(2) 平行因子分析法

平行因子分析法(PARAFAC)最早于 1970 年被提出并应用于心理学方面的研究^[14]，在处理三维数据时，以数据分离部分代替化学分离部分，显示出其快速、通用和可靠的特点，在化学计量领域

中得到了广泛应用^[15-16]。

PARAFAC 维持了三维数据结构,其模型为

$$x_{ijk} = \sum_{f=1}^F a_{if} b_{jf} c_{kf} + e_{ijk} \quad (3)$$

式中, F 为成分数; a_{if}, b_{jf}, c_{kf} 分别为矩阵 $A(I \times F)$, $B(J \times F)$, $C(K \times F)$ 的元素; e_{ijk} 为测量误差。

该方法将传感器阵列响应的三维数据集($I \times J \times K$) 通过特征值分解为 3 个二维矩阵,由于三线性分解后解的唯一性,所得结果具有明确的物理意义,可以对样本、响应过程和传感器贡献等方面信息进行详细分析。

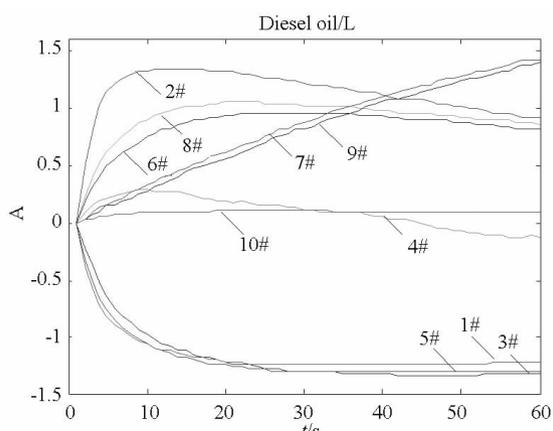
(3) 主成分回归

由于气体传感器响应机理比较复杂,很难用单一的信号值或理论数学模型来表示传感器响应和气体浓度的关系。利用主成分回归(PCR) 分析,可以实现单一气体浓度的检测,进一步达到设定阈值门限和报警目的。主成分回归分析是一种线性多元校正分析方法^[8],通过对参与回归的主成分的合理选取可以去掉噪声,建立主成分变量和气味浓度之间的回归模型,从而利用这个回归方程来对未知气体的浓度进行定量预测。

2 实验结果分析与验证

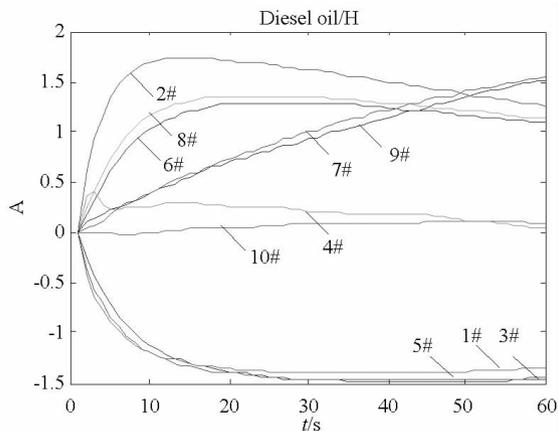
2.1 传感器阵列对油液挥发气体的响应指纹图

图 2 给出进行数据预处理后传感器的响应曲线。每条曲线代表一只传感器响应,横轴为采样时间 t ,纵轴为预处理后的响应值。由图可知,传感器阵列从初始洁净空气到接触样品气体的过程中相对电导率都有变化。每一个传感器的响应各不相同:有的响应值变化很快,并达到了最大值,随着玻璃瓶中气体不断被抽出,浓度下降而响应值相应减小,有的响应值在测量的 60s 内一直是上升的;传感器阵列对不同种类油液以及对不同浓度



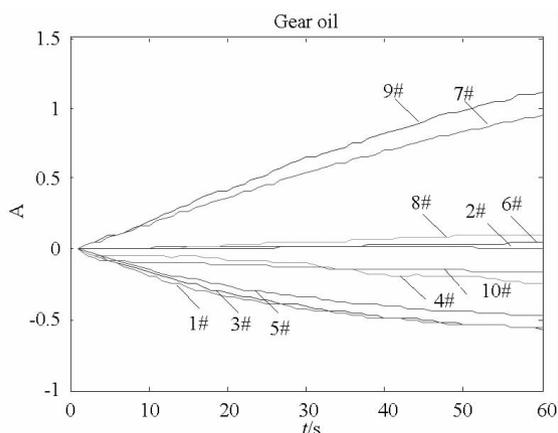
(a) 低浓度柴油

(a) diesel oil with low concentration



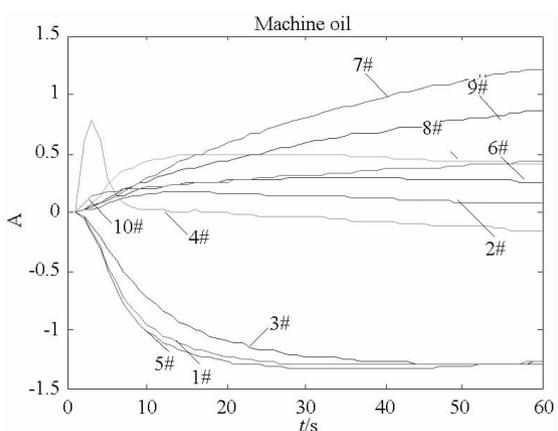
(b) 高浓度柴油

(b) diesel with high concentration



(c) 齿轮油

(c) gear oil



(d) 机油

(d) machine oil

图 2 人工嗅觉系统对三种不同油液挥发气体的响应图(预处理后)

Fig. 2 Time response curves of the artificial olfactory system to VOCs (preprocessed)

的同一种油液的响应指纹图也各异。这意味着利用人工嗅觉系统测量油液挥发物是可行的。经分析,发现 4#传感器的测量误差较大,所以在下面

的分析中,去掉了 4#传感器的响应数据。

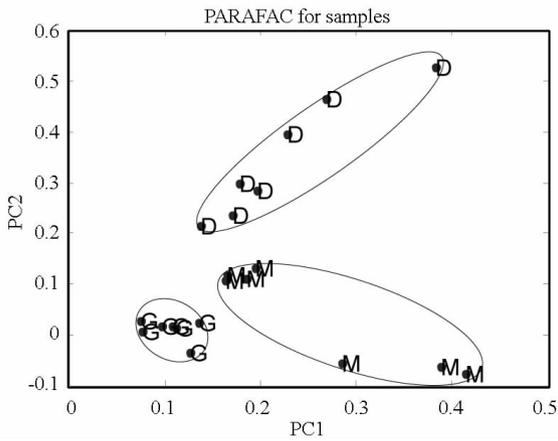
2.2 不同种类油液的定性识别

制备不同浓度的汽油、机油和齿轮油的挥发气体样本各 7 种,对其进行测试,得到三维数据集 (21 × 60 × 9)。对数据集进行平滑处理、基线处理以及归一化处理。

表 1 给出了不同模型的分析结果。各个模型前 2 个主成分的累积方差解释率都超过了 85%,取主成分数为 2。应用 PCA 时,分别选取最大值和上升阶段的最大变化率作为特征值,其中后一种只能把柴油挥发气体样本完全分离开,机油和齿轮油挥发气体样本有交集。图 3(a)、(b)、(c) 分别给出了三维模型 PARAFAC、Unfolding PCA 和二维模型 PCA(取最大值作为特征值)分析得到的主成分得分结果。可知三种模型都能很好地

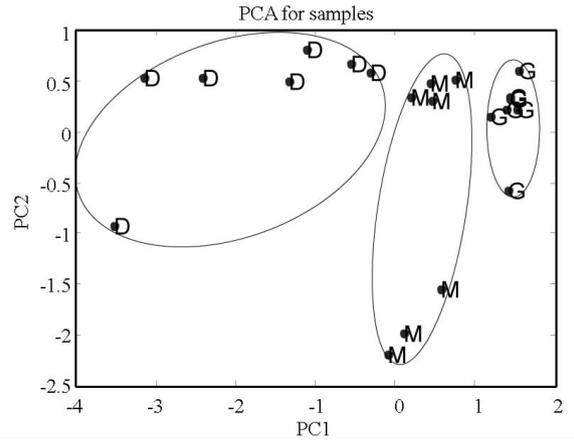
把三种油液挥发气体区分开来。PARAFAC 保持了数据的三维信息,利用了传感器响应过程的全部信息,不同种类油液挥发气体的得分分布更加密集,分离效果最佳。Unfolding PCA 对数据进行拼接,造成二维数据的表达矩阵的维数增加,导致计算量大大增加,同时由于平铺处理,损失了三维数据结构的部份信息。

进一步利用 PARAFAC 模型分析了各个传感器对前 2 个主成分信息贡献情况,如图 3(d)。由图知,1#、3#、5#传感器对 2 个成分的贡献率基本相同,特性基本相同,与图 2 中是一致的。可进一步进行传感器阵列优化,剔除其中 2 个传感器,减小数据冗余。表 1 中给出了优化后(剔除 3#、5#传感器)的阵列应用 PCA 模型的分析结果,表明三种油液气体能完全分离开来。



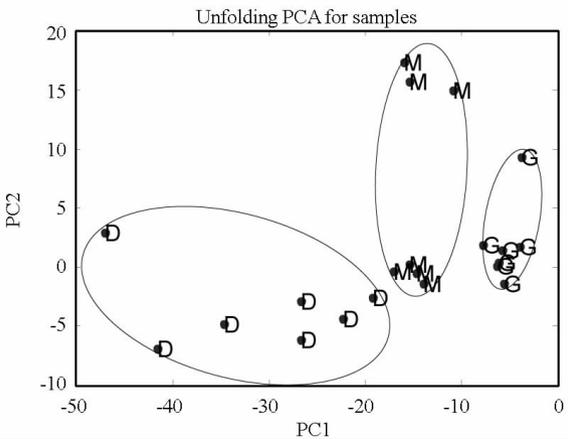
(a) PARAFAC 分析样本

(a) PARAFAC model for samples



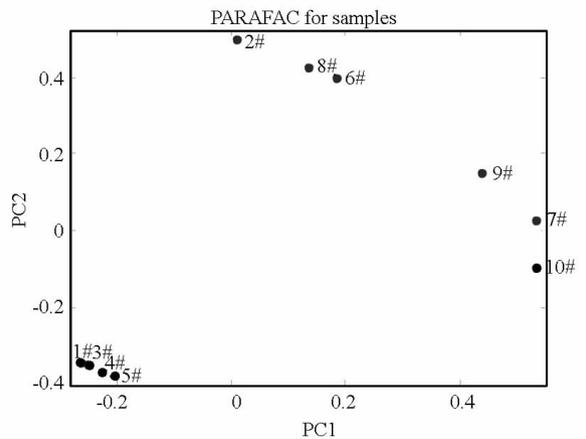
(b) PCA(最大值) 分析样本

(b) PCA (maximum) model for samples



(c) Unfolding PCA 分析样本

(c) Unfolding PCA model for samples



(d) 传感器贡献情况

(d) Loadings plot of PARAFAC model for 9 sensors

图 3 不同模型分析结果

Fig. 3 Scores plots of the first two factors of different models

表 1 不同模型分析的油液气体定性分析结果(主成分数为 2,去掉 4#传感器)

Tab.1 Qualitative analysis results obtained by different models (two components, the 4# sensor eliminated)

提取特征	数据阵大小	模型	累积方差解释率	分离效果	计算量
所有时间点	21 × (60 × 9)	Unfolding PCA	75.1343, 86.4564	DMG ⁽¹⁾	较大
所有时间点	21 × 60 × 9	PARAFAC	76.0271, 89.4970	DMG	大
最大值	21 × 9	PCA	83.0586, 92.5253	DMG	小
最大上升率	21 × 9	PCA	68.1451, 85.4494	D	小
最大值	21 × 6 ⁽²⁾	PCA	78.1574, 92.8574	DMG	小

(1)表示分离效果,DMG:三种油液挥发气体完全分离开来;D:柴油挥发气体完全分离,另 2 种有交叠;

(2)去掉 3#和 5#传感器响应数据后的数据阵。

2.3 不同浓度油液的定量分析

制备 11 种不同浓度(5ml、10ml、20ml、30ml、40ml、50ml、60ml、70ml、80ml、90ml 和 100ml)的机油挥发气体样本。每个样本分别测量 2 次,得到 22 个样本,其中 11 个不同浓度样本用作建模,另 11 个作为测试样本。对测量数据集进行平滑处理和基线处理,取响应最大值作为特征值进行分析。

从对建模样本进行分析获得的各主成分来看,前四个因子的方差解释率所占的比例分别为 93.38% ,5.69% ,0.82% ,0.11% ,取主成分数为 2。图 4 所示为 11 个建模样本的实际值 T 和主成分回归模型预测值 Y 。

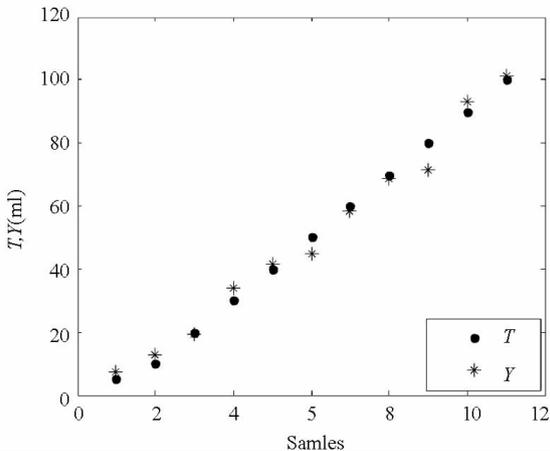


图 4 建模样本的实际值和 PCR 回归估计值比较图

Fig.4 Actual values vs. predicted values of the PCR model for the model data

为了更好评估主成分回归模型的有效性,进一步建立建模样本的实际值 T 和主成分回归的估计值 Y 之间的线性回归方程为 $Y = 0.99T + 2.6$,并可计算相关系数 $R = 0.99355$,相关系数接近 1,表明回归效果好。

表 2 主成分回归分析方法对测试样本估计的相对误差

Tab.2 Relative errors of the PCR model for the test data

序号	实际值 T (ml)	预测值 Y (ml)	相对误差 (%)
1	5	5.80	16
2	10	8.28	-17.2
3	20	18.76	6.2
4	30	32.44	8.13
5	40	43.95	9.87
6	50	50.57	1.14
7	60	59.32	-1.13
8	70	72.21	3.16
9	80	82.84	3.55
10	90	90.72	0.80
11	100	99.68	-0.32

$Y_i = 0.996T_i + 5.4 \quad R = 0.9974$

利用测试样本对回归模型进行验证,表 2 为得到测试样本的回归估计值 Y_i ,以及 Y_i 和测试样本的实际值 T_i 之间的相对误差。浓度比较低时,传感器响应值小,测得实验数据误差大,同时由于储气装置内环境气体的影响,最终使得回归误差比较大;浓度比较高时,主成分回归方法对测试样本取得了很好的估计效果。

3 结论与展望

本文应用人工嗅觉系统对机械设备渗漏可能产生的柴油、齿轮油和机油的挥发气体进行了测量。利用三维 PARAFAC、Unfolding PCA 模型和二维 PCA 模型都能实现对三种气体分类,尤其是 PARAFAC 充分利用数据的三维结构信息,三类油液挥发气体样本的得分分布更为集中,并可以实现传感器阵列的优化。同时,应用 PCR 分析实现了对油液挥发气体的定量确定,在油液浓度较

高时,误差在 10% 以下。说明人工嗅觉系统可用来进行渗漏油液的分析,并设定阈值进行报警,为进行机电设备故障诊断提供一种新方法。下一步的研究将集中在对微量渗漏以及混合油液渗漏的检测。

参考文献 (References)

[1] 张健,谢辉,曹谢东,等. 油气管道漏检测技术及发展趋势[J]. 仪器仪表与分析监测, 2011, 4(2): 39 - 41.
ZHANG Jian, XIE Hui, CAO Xiedong, et al. Inner examination technology on pipeline and the development trend [J]. Instrumentation Analysis Monitoring, 2011, 4(2): 39 - 41. (in Chinese)

[2] 王延年,朱强,赵则祥. 长输油气管道渗漏检测方法研究进展[J]. 石油机械, 2007, 35(3): 49 - 53.
WANG Yannian, ZHU Qiang, ZHAO Zexiang. Advancement in the study of leakage detection method for long-distance oil and gas transportation pipeline [J]. China Petroleum Machinery, 2007, 35(3): 49 - 53. (in Chinese)

[3] 梁伟,张来斌,王朝晖. 声发射检测技术在管道渗漏信号识别中的应用[J]. 科学技术与工程, 2007, 7(8): 1596 - 1601.
LIANG Wei, ZHANG Laibin, WANG Zhaohui. Application of acoustic emission technique in pipeline leak detection [J]. Science Technology and Engineering, 2007, 7(8): 1596 - 1601. (in Chinese)

[4] 庄须叶,王浚璞,邓勇刚,等. 光纤传感技术在管道渗漏检测中的应用与进展[J]. 光学技术, 2011, 37(5): 543 - 550.
ZHUANG Xuye, WANG Junpu, DENG Yonggang, et al. Optical fiber sensing technologies for pipeline leakage detection [J]. Optical Technique, 2011, 37(5): 543 - 550. (in Chinese)

[5] Gardner J W, Bartlett P N. A brief history of electronic noses [J]. Sensors and Actuators B, 1994, 18(1): 211 - 220.

[6] Wilson A, Baietto M. Applications and advances in electronic-

nose technologies [J]. Sensors, 2009, 9(7): 5099 - 5148.

[7] Röck F, Barsan N, Weimar U. Electronic nose: current status and future trends [J]. Chemical Reviews, 2008, 108(2): 705 - 725.

[8] Pearce T C, Schiffman S S, Nagle H T, et al. Handbook of machine olfaction: electronic nose technology [M]. Weinheim: Wiley-VCH, 2003.

[9] Brezmes J, Fructuoso M L L, Llobet E, et al. Evaluation of an electronic nose to assess fruit ripeness [J]. IEEE Sensors Journal, 2005, 5(1): 97 - 108.

[10] Ryan M, Homer M, Buehler M, et al. Monitoring space shuttle air quality using the jet propulsion laboratory electronic nose [J]. IEEE Sensors Journal, 2004, 4(3): 337 - 347.

[11] Matthes J, Groll L, Keller H B. Source localization by spatially distributed electronic noses for advection and diffusion [J]. IEEE Transaction on Signal Processing, 2005, 53(5): 1711 - 1719.

[12] 鲁小利,王俊,海铮. 基于遗传优化神经网络的电子鼻对可乐的检测 [J]. 传感技术学报, 2007, 20(6): 1211 - 1214.
LU Xiaoli, WANG Jun, HAI Zheng. Detection of cola using electronic nose based on GA-BP network [J]. Chinese Journal of Sensors and Actuators, 2012, 20(6): 1211 - 1214. (in Chinese)

[13] Kroonenburg P M. Three-mode principal component analysis: theory and applications [M]. Leiden, The Netherlands: DSWO Press, 1983.

[14] Harshman R A. Foundations of the parafac procedure: models and conditions for an "explanatory" multimodal factor analysis [Z]. UCLA working papers in phonetics, 1970, 16: 1 - 84.

[15] Huy Phan A, Cichocki A. Parafac algorithms for large-scale problems [J]. Neurocomputing, 2011, 74(11): 1974 - 1980.

[16] Burian C, Brezmes J, Vinaixa M, et al. Ms-electronic nose performance improvement using the retention time dimension and two-way and three-way data processing methods [J]. Sensors and Actuators B, 2010, 143(2): 759 - 768.