

基于 FastICA 算法的大气监测电子鼻设备的研究探讨

胡文韬, 李志晨

(哈尔滨工业大学(威海), 山东威海 264209)

摘要:电子鼻是一种由具有部分选择性的化学传感器阵列和适当的模式识别系统组成,能识别简单或复杂气味的仪器,它模拟人的嗅觉系统,检测、分析、识别气味成分,广泛应用于现场污染源排放监测、突发事件现场的应急检测、作业场所安全监测和分析、污染气监测和分析。本设计由气体传感器阵列进行实时实地测量,由 DSP 芯片进行数据处理,采用盲源分离(BSS)原理,利用独立分量分析(ICA)算法分离出混合前的多个独立信号。便携式电子鼻由于其成本低、易携带、通用性好、能现场即时检测等特点在大气监测领域将会有广阔的应用前景。

关键词:电子鼻;盲源分离;DSP;ICA

中图分类号:X851

文献标识码:B

文章编号:1006-8759(2010)04-0054-05

STUDY ON AIR MONITORING ELECTRONIC NOSE BASED ON FASTICA

HU Wen-tao, LI Zhi-chen

(Harbin Institute of Technology at Weihai, Weihai 264209, China)

Abstract: Electronic nose is a part by a selective chemical sensor array and an appropriate pattern recognition system, to identify the odor of simple or complex, it simulates the human olfactory system, detection, analysis, identification of odor components, identification of odor components. Widely used in electronic nose detection and scientific research environment and many other occasions. The gas sensor array designed for real-time field measurements, and data processing using the DSP chip. Distinguish between the smell of the data processing using blind source separation (BSS) theory, the use of independent component analysis (ICA) algorithm to solve such problems, to isolate a number of independent pre-mixed signal. Portable electronic nose because of its low cost, portable, versatile, and can immediately detect the characteristics of the scene at the market will be a wide application prospect.

Keywords: Electronic nose, BSS, DSP, ICA

0 前言

随着工农业生产的发展,环境污染现象变得越来越严重。有毒性气体和可燃性气体不仅污染环境,而且有产生爆炸、火灾和使人中毒的危险。大气污染监测是监测和检测大气中的污染物及其含量,目前已认识的大气污染物有 100 多种,因此

气味监测在日常生活及工业中占有重要的地位,尤其在一些环境恶劣或有危险物质存在的环境中,单纯地依靠人类嗅觉将无法给出定量指标并且具有很大的危险性。在此背景下,电子鼻这一技术应运而生,此技术能够在若干种混合气体复杂的环境中快速、准确地识别出各种气体的浓度,并

作出定量分析。电子鼻主要由气体传感器阵列和信号处理系统组成,其识别气味的主要机理是由于在阵列中的每个传感器对被测气体都有不同的灵敏度,所以整个传感器阵列中各种型号传感器对不同气体的响应输出是不同的,正是这种区别,基于模式识别算法,才使系统能根据传感器的响应输出来识别气味。传感器观察得到的气味信号是多个独立信号的瞬时线性混合,气味种类及混合方式均为未知量。仅从观测信号中恢复所有的原始独立分量的过程称为盲源分离(Blind source separation, BSS)。独立分量分析(Independent

Component Analysis, ICA) 是近年来在解决这一问题过程中发展起来的一种新的技术。

1 系统设计

本设计以 DSP 芯片为核心实现了电子鼻功能。整个系统包括软硬件两部分:硬件部分实现气味信号的检测、调理处理、同步采集、DSP 基本工作平台等,使 DSP 器件获取到气味信息,类似人的鼻子;软件部分利用 DSP 器件的数据处理能力基于盲源分离原理在独立变量分析算法的基础上从混合信号中分辨出其中的成分,类似人的大脑。

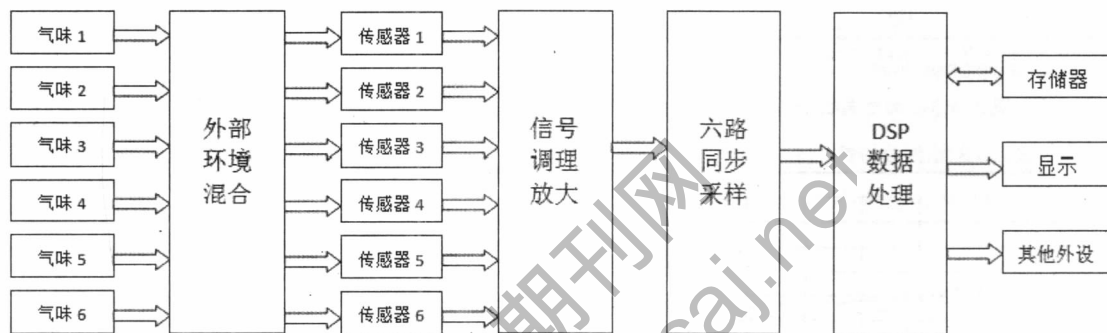


图 1 系统方案框图

2 信号拾取设计

传感器阵列就像人的鼻子,被气味分子激励产生电信号,由具有广谱响应特性、高灵敏度、对不同气味灵敏度不同的气味元件组成。单个气味传感器与传感器阵列在性能上有质的区别,单个气味传感器对气味的响应可用强度来表示,而气味传感器阵列除了各个传感器的响应强度外,还形成了多维空间中的响应模式,在环境条件一定的情况下,传感器阵列的响应模式与激励是一一对应的,而这正是电子鼻系统能对多种气体或气味进行辨识的关键所在。显然,传感器阵列的选取并不过于追求高选择性,相反利用了气味传感器所固有的交叉性敏感特性,使传感器阵列所确定的多维空间能够蕴含更多有用的信息。传感器阵列中各气味器件之间的响应特性应该有差别,并且这种差别越大,识别的效果越好,理论上讲,也就是观察信号的矩阵的行向量线性无关;器件响应时间和恢复时间快。

本设计开发的电子鼻所选用的传感器是 SnO₂ 金属氧化物型半导体传感器(MOS)。当它们与气味分子反应后,通常会引起金属氧化物半导体

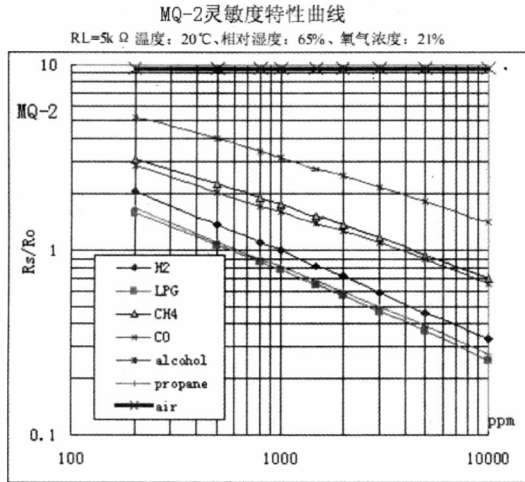
的电阻发生变化,通过传感器应用电路产生电信号,是目前生产量大应用广泛的气味传感器。其优点是气味选择范围宽、灵敏度高、稳定性好、能耗小、寿命长、耐腐蚀性强、制作简单等。鉴于本次设计中采用的采样芯片具有 6 路同时采样通道,所以选用了 3 种传感器(每种 2 个)组成传感器阵列。其特性和常用电路如表 2.1 所示。

表 2.1 传感器基本特性表

传感器型号	性能	检测范围
MQ-2	甲苯、乙醇、硫化氢、氢气	300~10 000 ppm
MQ-5	液化石油气	碳氢化合物
MQ-4	天然气	500~10 000 ppm

举例说明传感器的特性。MQ-2 气体传感器的特征气体为 1 000 ppm 异丁烷、探测范围为 300~10 000 ppm、灵敏度 gas ≥ 5、敏感体电阻 1~20KΩ、响应时间 ≤ 10s、恢复时间 ≤ 30s、加热电阻 31 ± 3Ω、加热电流 ≤ 180mA、加热电压 5.0V ± 0.2V、加热功率 ≤ 900mW、测量电压 ≤ 24V、工作温度: -20℃~+55℃。其特性曲线和温湿度曲线如图 2 和图 3 所示。

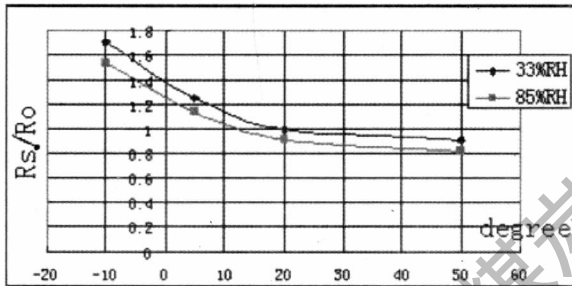
输入数据为传感器检测的混合信号，在源信号未知、信号传输通道特性未知的情况下，来获得源信号及信号传输通道特性。BSS原理示意图如图4所示。



Rs: 元件在不同气体, 不同浓度下的电阻值。
R0: 元件在洁净空气中的电阻值。

图2 MQ-2 特性曲线

MQ-2型气敏元件的温湿度特性



Ro: 20°C, 33%RH条件下, 1000ppm氢气中元件电阻。
Rs: 不同温度, 湿度下, 1000ppm氢气中元件电阻。

图3 MQ-2 温湿度曲线

3 盲分离原理、算法及验证

电子鼻识别气味的主要机理是在阵列中每个传感器对被测气体都有不同的灵敏度。例如，一号气体可在某个传感器上产生高响应，而在其它传感器上则产生低响应；同样，二号气体产生高响应的传感器对一号气体则不敏感。

本设计的根本问题是从混合信号中分离出未混合前的信号种类及性质。事先并不知道有几个独立信号，也不知道这些信号混合的方式，但还是要分离出这些独立信号，这个问题就是盲源分离(BSS)。盲源分离主要包括两大任务：一是构造目标函数，二是应用合适的数值优化方法使得目标函数达到最值。盲源分离目前常用的算法有基于高阶统计量的FASTICA算法、基于信息理论的Infomax算法和基于四阶统计量的JADE算法。对比这三种算法，采用FASTICA算法有可能解决此次设计的问题。

盲源分离算法 (Blind Source Separation, BSS)

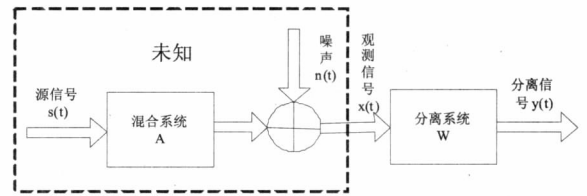


图4 BSS原理示意图

上图中， $S=[s_1(t), \dots, s_N(t)]^T$ 是未知的 N 维源信号向量， A 是未知的混合矩阵， $n=[n_1(t), \dots, n_N(t)]^T$ 是 M 维噪声向量， $X=[x_1(t), \dots, x_M(t)]^T$ 是传感器输出的 M 维观测信号向量，有 $X=AS+n$ ，盲源分离算法要求只知道 X 来确定 S 或 A 。独立分量分析 (Independent Component Analysis, ICA) 是 BSS 的一种，其基本含义是把信号分解成若干个互相独立的成分。如上图，ICA 的目标就是寻找一个分离矩阵 W ，使 X 经过变换后得到的新矢量 $Y=[y_1(t), \dots, y_N(t)]^T$ 的各分量尽可能的独立。 $Y=WX$ 为待求的分离信号向量，也就是源信号 S 的估计值。

ICA 的主要依据和前提是假设源信号是独立的，因此我们自然可以设想 ICA 算法的第一步是建立目标函数来表征分离结果的独立程度。目标函数确定后，通过各种不同的优化算法进行优化确定分离矩阵 W ，主要有基于峭度、基于负熵及近似负熵、基于最小互信息的三种目标函数，人们从最小互信息、最大似然函数和负熵最大化等角度提出了不同的算法。代表性的算法有：基于负熵的 FastICA 算法、基于信息理论的 Infomax 算法和基于四阶统计量的 JADE 算法。下面，仅对此次设计所用的 FastICA 算法进行说明。

FastICA 算法的信号预处理包括信号的去均值和白化。通过信号的预处理，可以使 FastICA 的工作量大为减少，提高盲源分离的效率，也能使问题更符合约束条件。当信号是 m 行 n 列的矩阵 S 时，对其每一行求均值并写成 m 行 1 列的均值矩阵 S_1 ，将 S 的每一列减去 S_1 后，就完成去均值处理。白化就是使信号通过白化矩阵变换后的自相关为 0，可以证明，经过白化后的数据盲分离矩阵 W 为正交矩阵，这样可以减少需要估计的参数个数。

FastICA 基于负熵 $J(x)=H_c(x)-H(x)$, 实际应用中由于概率密度函数难以确定, $J(y)$ 也不能直接计算, 一般取近似 $J(y) \approx \frac{1}{12} E\{y^3\}^2 + \frac{1}{48} kurt(y)^2$, 但该式的第一项在实际存在的大量的对称分布中为 0, 使负熵的近似值等于峭度的平方, 会出现非鲁棒性问题。为此使用一个任意的非二次函数 G , 做相应的近似变换 $J(y) \propto [E\{G(y)\} - E\{G(v)\}]^2$, 其中 v 是均值为 0、方差为 1 的高斯随机变量。通过选择 G 可以得到更好的负熵近似值, 若选择随自变量增长不是太快的函数 G 则能得到更好的鲁棒性。这样我们得到了能在峭度和负熵两个经典的非高斯性度量间取得很好折中的负熵近似估计, 其概念简单、计算量小, 且具有良好的统计特性和鲁棒性。

FastICA 的准则是负熵最大化。因为 $E\{G(v)\}$ 为常数, 为求负熵的最大值, 相当于求 $E\{G(y)\}$ 的极值点。根据 Lagrange 定理和约束条件, 构造代价函数:

$$F(w_i) = E\{G(w_i^T z)\} + \lambda [E\{(w_i^T z)^2\} - 1]$$

要求 w_i 使 F 取得极大值, 先求

$$f(w_i) = \frac{\partial F(w_i)}{\partial w_i}$$

令其等于零, 再利用牛顿迭代法求解 w_i , 得到迭代公式:

$$w_i(k+1) = w_i(k) - \frac{f(w_i(k))}{f'(w_i(k))}$$

可化为:

$$w_i(k+1) = E\{zg(w_i^T(k)z)\} - E\{g'(w_i^T(k)z)\}w_i(k)$$

这样就可以算出一个独立分量, 但每次迭代完成后, 应对 w_i 进行归一化处理。当计算 n 个独立分量时, 需要 n 个列矢量, 并且每次迭代后需进行线性组合去掉相关性。

以上就是处理气味分离的 FastICA 算法的基本原理。下面, 总结一下采用负熵的固定点算法的步骤:

- 1) 对数据进行中心化, 使其均值为零;
- 2) 然后进行白化, 得 z ;
- 3) 选取一个具有单位范数的初始化向量 w (可随机选取);
- 4) 更新 $w \leftarrow E\{zg(w^T z)\} - E\{g'(w^T z)\}w$, 其中 g 为选定的函数, 一般取 g 为 $G_2, G_2 = -\exp(-\frac{y^2}{2})$;
- 5) 标准化 w ;

6) 如不收敛返回步骤 4。

FastICA 算法和别的 ICA 算法相比, 有着许多我们期望的特性: 收敛速度快; 和梯度算法不同, 无需选择步长参数, 易于使用; 该算法能适用于任何非高斯信号; 可以通过使用不同的非线性函数 G 使算法获得不同的性能; 独立分量可被逐个估计出来, 这在仅需要估计几个独立分量的情况下, 能减小计算量。FastICA 算法还有许多神经算法的优点: 并行、分布、计算简单, 内存要求小。

FastICA 算法的实现目前多在 MATLAB 环境下。我们在 MATLAB 中产生一个余弦波和一个方波作为信源信号, 然后利用一个随机矩阵混合生成的信号作为混合信号, 最后利用 FastICA 算法进行分离, 源信号、混合信号和分离信号如下图所示。结果显示, 对于完全满足假设条件的信号, 算法 FastICA 的分离效果是比较理想的, 这就验证了算法的可行性。

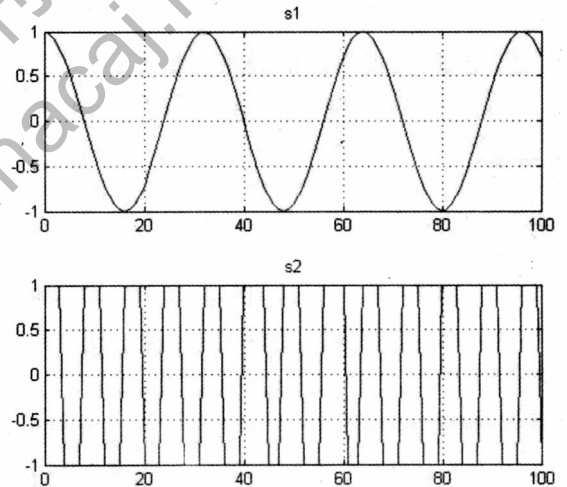


图5 源信号—余弦波、方波

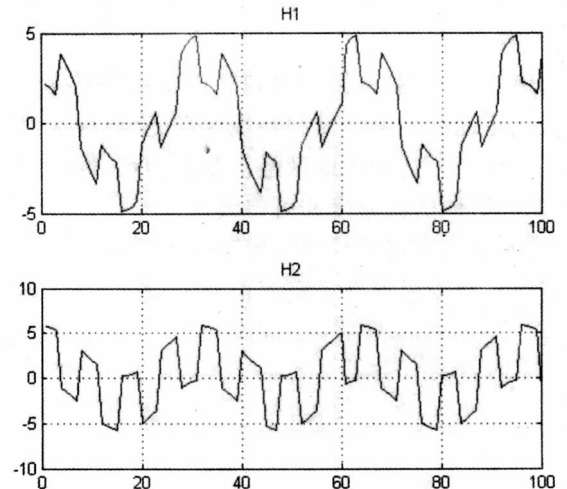


图6 余弦波、方波混合后信号

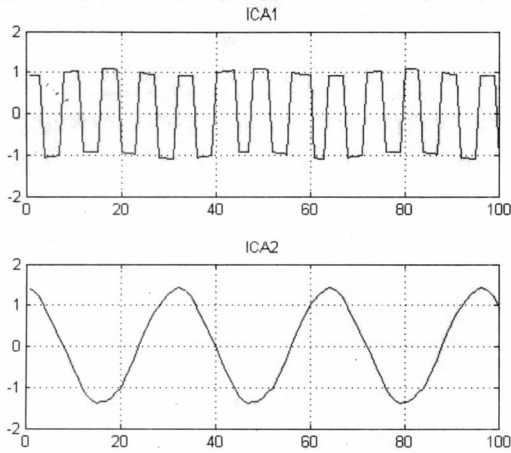


图7 分离后的方波、余弦波

4 结语

电子鼻通过识别气体表象图谱的方法，具有扩展可检测气体的能力，以达到识别多种气体的功能。而且对不同的气体识别准确，同时对所测气体进行浓度测量，测量下限可达 ppm 级，精度在 5% 以内。与传统大气监测设备相比，具有快速、定性、定量、便携以及准确探测抗干扰能力强的优点，在大气监测领域将会有广阔的应用前景。

参考文献

[1] 许燕文.基于 DSP 的便携式电子鼻系统的研究[D].西安理工大学硕士学位论文.2008:2~5.
 [2] 马建龙,牛奕龙,陈海洋.盲信号处理[M].国防工业出版社.2006:

85~98

[3] 朱世东.盲信号分离的 DSP 实现[D].安徽大学硕士学位论文.2007:9~13.
 [4] 张发启.盲信号处理及应用 [M].西安电子科技大学出版社.2006:1~42.
 [5] 杨福生,洪波.独立分量分析的原理与应用[M].清华大学出版社.2006:91~110.
 [6] 徐毓龙,徐玉成.电子鼻的研究和开发现状 [J].传感器世界.1999:3~5.
 [7] 张红梅,何玉静.电子鼻技术的历史、研究现状及发展前景[J].科技信息.2008:1~2.
 [8] 徐后坤,胡木林,谢长生.手持式电子鼻的实现[J].传感器世界.2006:1~3.
 [9] 祝佳婧.电子鼻传感器技术的研制进展[J].北京生物医学工程.2002:1~3.
 [10] 梁坚,辛德暖,陆毅新.电子鼻-智能气体传感器[J].传感技术学报.1995:1~2.
 [11] 周宗潭,董国华,徐昕,胡德文.Aapo Hyvarinen,Juha Karhune,Erkki Oja.独立成分分析[M].电子工业出版社.2007:93~135.
 [12] Seungjin Choi,Andrzej Cichocki,Hyung -Min Park,Soo -Young Lee.Blind Source Separation and Independent Component Analysis. Neural Information Processing - Letters and Reviews.2005:3~40.
 [13] Andrzej Cichocki,Shun -ichi Amari.Adaptive Blind Signal and Image Processing,John Wiley & Sons.2002:37~62.
 [14] Aapo Hyvarinen,Juha Karhune,Erkki Oja.Independent Component Analysis. John Wiley & Sons,2001:93~135.
 [15] Malaya K.Nath,Student Member,IEEE.Independent Component Analysis of Real Data.2009 Seventh International Conference on Advances in Pattern Recognitio.2009:1~4.

(上接第 53 页)

4 结论和建议

(1)资源型城市和地区，在资源产业发展过程中引发的环境问题，是资源诅咒现象的一个重要诱发因素；

(2)在地区产业结构调整 and 制定新经济复兴计划时，注重地区生态修复和环境保护，会给地区经济注入新的动力，加快地区经济复苏步伐，同时提高了新经济体的健康度和可持续性；

(3)引起资源诅咒现象的原因是多方面多层次的，资源环境问题促发了资源诅咒现象的发生，环

境问题的解决应从政策措施层面解决，即制度层面上解释和解决资源环境问题可以有效的帮助走出经济发展困境。

参考文献

(1)刘贞,程永军,杨武,任丽芸.从演化制度经济学角度分析资源诅咒现象[J].能源环境保护,2009,2:1~6.
 (2)王文行.资源诅咒问题研究新进展.经济动态[J],2008,5.
 (3)张刚柱.探讨煤炭开采引起的环境问题及对策[J].中国期刊网,2009,6.
 (4)徐康宁,邵军.自然禀赋与经济增长对资源诅咒命题的再检验[J].世界经济,2006,29.