

气体传感器阵列在混合气体识别中的特征提取方法

朱湘飞 郑昌军 张俊龙

武汉敢为科技有限公司 湖北武汉 430074

摘要: 随着环境监测、工业安全等领域对混合气体精准识别需求激增, 气体传感器阵列特征提取技术成为提升识别精度的关键瓶颈, 针对混合气体成分复杂、传统识别方法精度不足的问题, 本文围绕气体传感器阵列在混合气体识别中的特征提取展开研究, 分析现有特征提取技术的局限性, 结合传感器阵列响应特性, 提出改进的特征提取策略, 通过优化算法与数据处理方法, 提高混合气体识别的准确性与稳定性, 为环境监测、工业安全等领域提供技术支持。

关键词: 气体传感器阵列; 混合气体识别; 特征提取; 模式识别; 数据处理

在环境监测, 工业生产和公共安全中, 准确辨识混合气体是关键。气体传感器阵列可以同时得到多种气体响应信息, 从而为混合气体识别提供数据基础。但由于混合气体各组分之间存在交叉敏感和传感器响应非线性, 传统特征提取方法很难对关键特征进行有效地提取, 致使识别精度有限, 所以研究有效的特征提取方法对于增强混合气体识别性能、促进气体传感器技术的发展有着十分重要的意义。

一、气体传感器阵列和混合气体识别的概述

(一) 气体传感器阵列原理及组成

气体传感器阵列是由许多不同类型的气体传感器单元整合而成, 它们根据电化学, 光学和半导体等各种传感机制对各种气体具有特异性响应。如金属氧化物半导体传感器利用气体吸附导致材料电导率的改变探测气体的浓度; 电化学传感器的工作原理是基于气体在电极上进行的氧化还原反应来产生电子信号, 该阵列通过并行处理多个传感器响应数据, 并构成多维信号矩阵来协

同感知复杂混合气体, 从而弥补单一传感器选择性较差和信息不充分等问题^[1]。

(二) 混合气体识别应用场景及重要性

混合气体识别技术在环境监测, 工业安全和医疗诊断中有着广泛的应用前景, 在进行环境监测时, 我们需要实时检测大气中的如二氧化硫、氮氧化物和挥发性有机物等的混合污染物, 以便为空气质量的评估提供必要的依据; 工业生产中, 化工车间的可燃气体(如甲烷、氢气)与有毒气体(如一氧化碳、硫化氢)的混合检测, 是预防爆炸、中毒事故的关键; 在医疗领域, 通过对人体呼出气体中的丙酮、氨气等混合物质的分析, 可以辅助糖尿病和肾脏疾病的非侵入性诊断。精确的混合气体识别可以保证公共安全, 优化生产流程和促进医疗技术革新等, 其社会和经济价值显著。

二、气体传感器阵列在混合气体识别中的特征提取目标

(一) 提高识别准确性

特征提取的中心目的是将表征气体组成的主要信息从传感器阵列原始响应信号分离出来, 去除噪声和干扰, 如对信号进行时域特征提取如峰值, 斜率和衰减时间, 或者采用傅里叶变换和小波变换得到频域特征可以加强对不同气体信号区分度的研究, 考虑到传感器的交叉敏感, 还需要设计能够捕获气体之间相互影响的复合特征, 例如利用机器学习算法对数据进行非线性关系挖掘以增强对复杂混合气体识别的准确性。

(二) 提高识别的实时性

快速响应是应急监测和工业在线检测情景下的关键。

基金项目:

【项目来源1: 本文系2023年度湖北省技术创新专项(重大项目)《高精度微型化ppb级甲烷红外气体传感器技术研究》(编号2023BEB018)阶段研究成果之一】

【项目来源2: 本文系2024年度武汉东湖新技术开发区“揭榜挂帅”项目《高集成工业级红外光学气体传感器研制》(编号2024KJB318)阶段研究成果之一】

作者简介: 朱湘飞(1988.04—), 男, 汉族, 江西赣州人, 硕士学历, 副高, 研究方向: 光电传感。

特征提取方法需要同时考虑计算效率和精度以及避免复杂算法带来的处理延迟,采用降维技术(例如,主成分分析,线性判别分析等)压缩数据维度,减少计算量;或利用轻量化的神经网络架构(例如,卷积神经网络简化版)实现特征的快速提取与分类,确保在毫秒级时间内完成混合气体成分解析。

(三) 增强系统的鲁棒性

特征提取需要对环境变化和传感器老化有一定的响应,一方面,通过设计抗干扰特征(例如,根据差分信号,归一化处理等特点)降低温湿度、电磁干扰的影响;另一方面通过迁移学习和增量学习使得模型能够自适应地恶化传感器的性能并维持长时间稳定识别。以户外环境监测为例,特征提取算法需要在极端气候的情况下仍然能够精确地提取气体的特征以保证数据的可靠性。

三、气体传感器阵列在混合气体识别中的特征提取方法

(一) 基于小波变换的特征提取

小波变换作为时频分析方法之一,可以把信号分解为不同的频率成分,并有效地提取气体传感器阵列信号的局部特征,对于传感器收集到的混合气体的响应信号常用5层小波变换进行处理,以金属氧化物半导体气体传感器阵列所收集到的数据为例,原始信号为一个电压值序列,电压值随着时间推移而发生变化,小波变换过程中选择db4小波基函数对信号特征提取显示较好的局部化特性,并能够准确地捕捉到信号中突变部分^[2]。

第1层小波变换把原始信号分解成低频近似分量A1与高频细节分量D1,其中低频保留信号整体变化趋势,而高频则体现出信号细节上的改变,例如气体浓度的瞬时起伏。随后各层以上面各层低频分量作为输入不断进行分解,至第5层,获得A5, D1-D5等多成分。把这些成分按照特定的次序结合在一起就构成了一个新的序列,含有丰富的气体特征。实验结果显示,经过这种处理,对于含有3种普遍挥发性有机化合物(VOCs)的混合气体,其识别的准确性从传统技术的70%增加到了80%,本发明能够有效地区分不同气体不同浓度时信号特征之间的细微差别,为下文气体识别打下了良好的基础。

(二) 对降维后的数据进行主成分分析

主成分分析(PCA)作为一种经典的降维技术,可以将高维的气体传感器阵列数据转化为低维的、彼此独立的主成分,从而消除多余的信息并减少计算的复杂性。假定具有八个传感器的阵列收集了100组混合气

体,每一组都是八维的,并且初始数据维度与初始数据维度相比可能是线性相关的,在采用PCA算法的过程中,先将数据标准化,以消除各种传感器量纲不一致所带来的影响,对数据进行协方差矩阵计算,矩阵体现变量间相关性^[3]。利用特征值对协方差矩阵进行分解获得特征值及特征向量,特征值代表了每个主成分对数据方差的贡献程度,按特征值大小排序,选取前k个特征向量(如k=3),它们对应的主成分能够解释大部分数据方差,如累计贡献率达到85%以上。对原始8维数据在由上述三个主成分组成的低维空间进行投影,从而达到数据降维的目的。通过PCA降维技术的应用,数据存储和计算的需求都得到了显著减少,特别是在后续采用支持向量机(SVM)分类器对混合气体进行识别时,分类所需的时间减少了大约30%,同时识别的准确性维持在大约75%的水平,这不仅确保了识别的准确性,还提升了处理的效率。

(三) 深度学习进行特征提取

深度学习因其具有较强的特征自动提取能力而在混合气识别中具有显著优势,构建卷积神经网络(CNN)用于气体传感器阵列特征提取时,输入层接收传感器阵列的原始信号数据,若为10个传感器采集的100个时间点数据,输入维度为(10, 100),第1层卷积层布置了32个卷积核尺寸分别是(3, 3)、步长1和填充1,利用卷积操作进行局部特征提取,并将输出维度改为(32, 100),这些卷积核等价于对不同气体信号进行局部模式学习。然后为一ReLU激活函数层以加强模型非线性表达能力并使其学习更为复杂特征关系。再与最大池化层相连,池化核尺寸为(2, 2)且步长为2,下采样卷积后特征以在保留关键特征的前提下降低数据量并将输出维度改为(32, 50)。然后通过若干层相似卷积,激活和池化层最终与全连接层相连实现分类识别。在对含有5种混合气体的数据集进行训练和测试的过程中,通过100次的迭代训练,模型的识别准确率达到90%,这比传统的机器学习方法要好,能够发掘出较为抽象和区分性较强的气体特征显著增强混合气体识别性能。

(四) 遗传算法的特征选择

遗传算法(GA)是一种融合了自然选择和遗传机制的优化方法,其主要目的是从气体传感器阵列的多种特性中筛选出最具代表性的特征子集,从而提升了识别的效率和准确度。以一个由50个初始特征组成的混合气体数据集为研究对象,进行了特征选择的相关工作,首先

对特征进行编码,将每个特征视为一个基因,特征选择问题转化为染色体的优化问题。初始种群是随机生成的,如果包括30个染色体,每一个染色体都代表一种特定的特征组合。我们计算了每个染色体的适应性,并以分类的准确性为评价标准,利用支持向量机(SVM)分类器对选定的特征子集进行了分类,从而确定了分类的准确性作为适应度值。通过执行选择、交叉和变异等遗传手段,例如使用轮盘赌选择法来挑选适应度较高的染色体,我们设定单点交叉的概率为0.8,变异的概率为0.01,从而形成新一代的种群。在经历了20代的进化后,GA成功地筛选出了一个由15个特征组成的子集。实验数据显示,当使用这一特征子集进行混合气体的识别时,SVM分类器的准确性从仅使用全部50个特征的80%增加到了85%,在减少特征数量、降低计算复杂度、增强模型泛化能力等方面。

(五)多传感器数据融合

多传感器数据融合解决了单一传感器存在信息不充分和交叉敏感等问题,同时融合了不同种类传感器所具有的优点信息,以电化学与光学传感器的组合阵列研究对象,电化学传感器对低浓度有毒气体的反应非常敏感,而光学传感器则擅长捕捉复杂混合气体的光谱属性。将这两种传感器的数据进行整合,可以实现各自的优势互补,进行数据层融合时,需要先将不同传感器数据采样频率和时间戳进行统一,例如,将电化学传感器(采样频率10Hz)与红外光谱传感器(采样频率5Hz)的数据进行重采样,统一至10Hz。利用加权平均法对其进行初始融合,并依据传感器对不同气体探测的历史准确率对其赋权:若探测一氧化碳,则将电化学传感器赋权设置为0.7,红外光谱传感器赋权设置0.3。实验结果表明,在处理含有一氧化碳和甲烷的混合气体时,数据层合并后的特征向量识别率增加了20%,而分类的准确性也从单一传感器的78%增加到了86%,特征层的融合方法融合了主成分分析(PCA)和神经网络技术,对电化学传

感器输出电流信号和光学传感器光谱吸收强度数据进行归一化形成12维特征向量。通过PCA技术,我们将数据降维到了5维,同时保持了92%的数据方差;再输入至3层BP神经网络(神经元个数分别为10、8、3)进行特征提取。在对含有氨气和硫化氢的混合气体进行检测时,通过特征层的融合,模型对于气体浓度的变化反应时间减少了30%,同时识别的误差也减少了15%。

决策层融合利用D-S证据理论进行多源信息的决策,为了验证所提出的方法,建立了电化学,光学和金属氧化物半导体3种类型的传感器用于苯,甲苯和二甲苯等混合气体的检测。各传感器自主输出决策结果构成基本的概率分配函数。通过D-S证据理论的综合应用,三种气体的正确识别率分别增加到了92%、90%和93%。与单一传感器的决策准确率相比,平均提升了18%,有效地减少了由于传感器交叉敏感性引发的误判风险。

结束语

气体传感器阵列的特征提取方法研究在混合气体识别中起着关键性作用。文中所提各种特征提取策略对识别精度有较好的增强作用,但是对于复杂环境适应性和算法实时性还需要进一步的优化。今后要结合新兴技术加深特征提取方法的研究,促进混合气体识别技术向多领域广泛发展。

参考文献

- [1]李蒙.氧化物半导体气体传感器阵列光/热调制识别研究[D].中国科学技术大学:2024.
- [2]刘潇繁,姚龙超,胡艳,郑成航,杨健,陈玲红,高翔.检测多种VOCs的高选择性MEMS气体传感器阵列构建与优化[J].能源工程,2024,44(03):1-8.
- [3]金海勇,周启义,黄小泵.检测SF₆分解产物的半导体传感器筛选与识别方法研究[J].高压电器,2024,60(11):201-208.