

基于声纹识别的蒸汽泄漏报警监控系统

方维岚, 陆正卿, 曹 鑫, 朱晓燕, 马 骏

(上海烟草集团有限责任公司, 上海 200082)

摘要:蒸汽管道在工业企业中随处可见, 并且是相当重要的设备之一。一旦发生蒸汽泄漏, 轻则浪费能源, 影响生产质量, 重则造成安全事故。由于蒸汽管道往往位于工厂隐蔽的位置, 一旦泄漏, 除了日常巡检, 平时很难发现。为了及时发现蒸汽泄漏, 研制了 1 套基于声纹识别的蒸汽泄漏报警监控系统。由于蒸汽管道处于密闭空间, 且周围没有其他设备运行, 几乎不存在声音干扰问题。通过阵列麦克风, 利用声源定位算法对气体泄漏频段内的声音进行识别和定位。实验结果表明, 通过特定频段和声纹识别, 能够有效提高蒸汽泄漏的识别率, 实现蒸汽泄漏的自动识别报警。

关键词:蒸汽泄漏; 声源定位; 声纹识别; Mel 倒谱系数; 长短时记忆神经网络

中图分类号: TB52 + 6

文章编号: 1000 - 0682(2024)04 - 0032 - 04

文献标识码: A

DOI: 10. 19950/j. cnki. CN61 - 1121/TH. 2024. 04. 007

A system for monitoring steam leakage alarm based on voiceprint recognition

FANG Weilan, LU Zhengqing, CAO Xin, ZHU Xiaoyan, MA Jun

(Shanghai Tobacco Group Co., Ltd., Shanghai 200082, China)

Abstract: Steam pipelines are ubiquitous in industrial enterprises and are one of the most important equipment. Once a steam leak occurs, it can waste energy, affect production quality, and even cause safety accidents. Due to the fact that steam pipelines are often located in hidden locations in factories, once a leak occurs, no one will notice it except for daily inspections. So, in order to detect steam leaks in a timely manner, a steam leak alarm and monitoring system based on voiceprint recognition has been developed. Due to the steam pipeline being in a closed space and no other equipment operating around it, there is almost no problem of sound interference. Identify and locate the sound within the gas leakage frequency band using a sound source localization algorithm through a whole array of microphones. The experimental results show that through specific frequency bands and voiceprint recognition, the recognition rate of steam leaks can be effectively improved, and automatic recognition and alarm of steam leaks can be achieved.

Keywords: steam leakage; sound source localization; voiceprint recognition; MFCC; LSTM

0 引言

蒸汽管道在工业企业中随处可见, 并且是相当重要的设备之一。一旦发生蒸汽泄漏, 轻则浪费能源, 影响生产质量, 重则造成安全事故。目前对于蒸汽泄漏的发现, 大多由每天定期巡检的巡检人员在工作中发现。由于蒸汽管道通常都处于夹层位置,

人员需要攀爬才能够到达, 并且管道很长, 有时还会跨越多个房间。巡检人员需要花费大量的时间才能完整地巡查完蒸汽管道。所以采用一种无人化的自动识别蒸汽泄漏系统, 对泄漏进行报警是完全有必要的。

人耳听觉的最高频率是 20 kHz, 气体泄漏如果漏孔较小, 产生的声波频率往往超过 20 kHz, 最高可达到 30 kHz。针对这些超声频段的音源提取, 使用阵列麦克风及定位算法解决声音定位; 然后通过 Mel 频率倒谱系数 (MFCC) 和线性预测倒谱系数 (LPCC) 等算法^[1-3]将声源的声信号特征参数提取

收稿日期: 2024 - 03 - 04

第一作者: 方维岚 (1969—), 男, 上海人, 工业电气自动化专业硕士, 高级工程师, 主要研究方向为工业自动控制与建筑智能化等。

E-mail: dengliao79855995@163.com

出来,将声纹数据化。使用深度学习神经网络,建立分类模型,并生成气体泄漏的模型样本。基于监测点声纹的实时采集、分析气体泄漏的模型样本,达到一定相似阈值后,判断为气体泄漏。

该文针对烟厂蒸汽管道层的数据采集与分析,根据蒸汽管道夹层的环境特点以及蒸汽泄漏的声纹特征,提出了 1 种将 MFCC 与 LSTM^[11-15] 相结合的蒸汽泄漏识别方法,并研制了相应的基于声纹识别的蒸汽泄漏报警监控系统。实验结果验证了该方法的有效性,可以用于蒸汽发生泄漏时的及时报警。

1 MFCC 特征提取原理及 LSTM 神经网络

声音识别技术一般分为语音识别技术和非语音识别技术两类。目前,语音识别技术已经相当成熟,而非语音识别技术起步相对较晚,但也越来越受到国内外学者的重视,其应用正在逐步向工业、医学、国防科技等领域扩展。

关于声音信号的识别,一般使用 Mel 倒谱系数 MFCC) 特征提取和混合高斯模型 (GMM) 模式识别的方法。并结合人工神经网络 (BP, RNN 和 LSTM) 等形成声纹识别模型。

在大量研究的基础上,使用一种基于 Mel 倒谱系数 MFCC 与长短时记忆 LSTM 神经网络的声纹识别方法,用于识别在密闭安静的空间内,蒸汽管道泄漏发出的高频声音。

1.1 MFCC 特征提取原理

MFCC 常被用于声信号的识别领域,其原理基于声音频率带宽的变化,通过对以 Mel 频率刻度上线性间隔的三角带通滤波器组输出的对数能量进行去相关计算得到。MFCC 特征提取流程如图 1 所示。

(1) 信号预处理

预加重: 声信号在传输过程中通常会出现低频

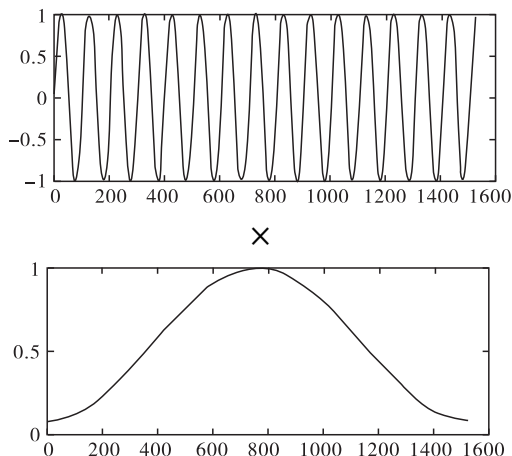


图 2 帧信号叠加 Hamming 窗示意图

信号衰减小、高频信号衰减大的特性,这种传输过程中表现出来的低通滤波特性会使得整个声信号出现与原始信号频率特性不符的现象,导致采集到的声信号大部分能量集中在低频段,为此,要对采集到的声信号进行提升高频成分的预加重处理:

$$H(z) = 1 - az^{-1} \quad (1)$$

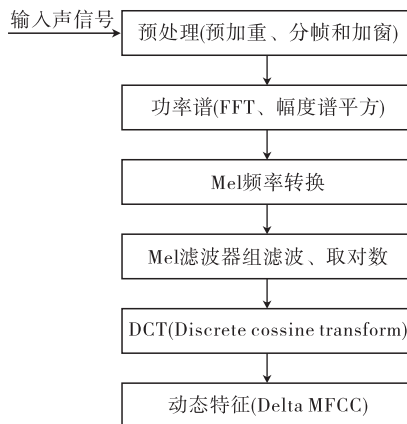


图 1 MFCC 特征提取流程图

分帧和加窗: 声音信号具有准平稳特性,在一个短时间范围(一般为 10 ~ 30 ms)内,可以认为是一个稳定的过程,具有短时平稳性,可以用平稳过程的分析方法进行分析,因此,可以将声音信号划分为一个个的短时段来对其中可能隐含的特征参数进行分析,其中每一短时段称为一帧。由于直接分段截取后的每一帧信号在起始点和终点处可能发生突变,会导致 Gibbs 现象的出现,因此对截取后的每一帧信号需要用窗函数(通常用 Hamming 窗函数)乘以原来的信号,以便减小起始点和终点处的不连续性(如图 2 所示)。进行这样的操作之后,对整个声信号而言,最后分析出的就是由每一帧特征参数组成的整个特征参数的二维时间序列。

(2) 求功率谱

对每一帧信号做快速傅里叶变换 (fast Fourier transform, FFT), 并对求得的幅度谱取平方得到离散功率谱。

(3) Mel 频率转换

将求得的功率谱的线性频率 f 按式(2)转换为 Mel 频率 m_f , 转换示意图如图 3 所示。

$$m_f = 2595 \times \log_{10}(1 + f/700) \quad (2)$$

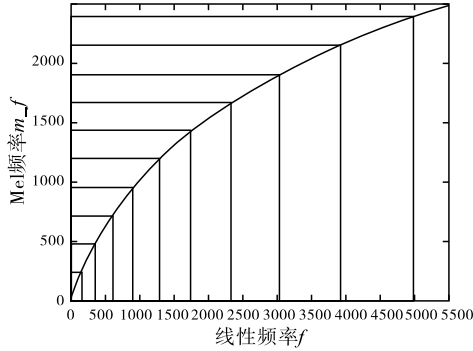


图 3 Mel 频率转换示意图

(4) Mel 滤波器组滤波、取对数

Mel 滤波器组为在 Mel 频率轴上均匀分布的数个带通滤波器 $H_m(k)$, 其中, $0 \leq m < M$ (M 为滤波器总数), $0 \leq k < N$ (N 为每帧的采样点个数):

$$s(m) = \ln\left(\sum_{k=0}^{N-1} |Xa(k)|^2 H_m(k)\right) \quad (0 \leq m < M) \quad (3)$$

(5) 离散余弦变换

$$C(n) = \sum_{m=0}^{N-1} s(m) \cos \frac{\pi n(m-0.5)}{M} \quad (n = 1, 2, \dots, L) \quad (4)$$

式中: L 为 MFCC 系数的阶数 ($L \leq M$)。

(6) 动态特征

上式求得的 $C(n)$ 即为声信号参数的静态特征, $C(n)$ 的一阶差分 $d(n)$ 的表达式为:

$$d(n) = \begin{cases} C(n+1) - C(n) & n < K \\ \frac{\sum_{k=1}^K k[C(n+k) - C(n-k)]}{\sqrt{2 \sum_{k=1}^K k^2}} & \text{Other} \\ C(n) - C(n-1) & n \geq L-K \end{cases} \quad (5)$$

将一阶差分的计算结果, 再进行一次一阶差分计算, 就获得动态特征。

1.2 LSTM 神经网络

长短期记忆人工神经网络, 简称 LSTM 神经网络

络是基于时间序列的预测算法, 是为了解决循环神经网络存在的长期依赖问题而设计出来。

在实际应用中, 选择蒸汽泄漏时声纹特征的关键参数的历史数据, 利用基于时间序列的预测算法来预测参数在未来某个或某些时刻的值, 通过预测模型将关键参数数据与识别信息数据结合, 对蒸汽泄漏的声纹进行识别和持续预测, 判断是否存在泄漏的情况。

1.2.1 LSTM 基本结构单元 (LSTM cell)

LSTM 的基本结构单元如图 4 所示。图中包含了遗忘门、输入门、输出门以及细胞状态更新层, 它们以一定的方式进行数据交互, 从而避免发生梯度消失和梯度爆炸现象。

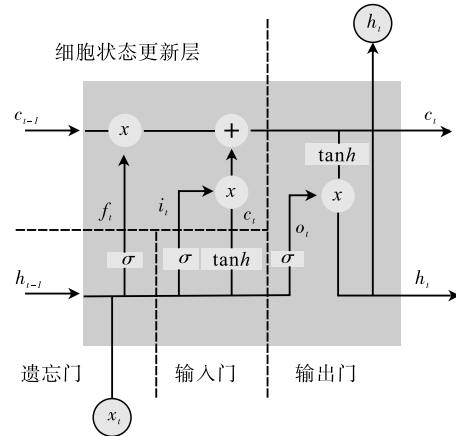


图 4 LSTM 基本结构单元示意图

(1) 遗忘门

遗忘门的结构表达式为:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (6)$$

式中: f_t 为 t 时刻遗忘门的输出; σ 为 sigmoid 函数; W_f 为遗忘门权重; h_{t-1} 为 $t-1$ 时刻的输出值; x_t 为 t 时刻的输入; b_f 为遗忘门的偏置项。

(2) 输入门

输入门结构表达式为:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (7)$$

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (8)$$

式中: i_t 为 t 时刻输入门的输出; W_i 为输入门权重矩阵; b_i 和 b_c 为输入门偏置项; \tilde{C}_t 为 t 时刻的 1 个新的备选值向量。

(3) 输出门

输出门结构表达式为:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (9)$$

$$h_t = o_t \cdot \tanh(C_t) \quad (10)$$

式中: o_t 为 t 时刻输出门的输出; W_o 为输出门权

重矩阵; b_o 为输出门的偏置项。

(4) 细胞状态更新层

将 $t-1$ 时刻的旧细胞状态 C_{t-1} 更新到 t 时刻的细胞状态 C_t 。细胞状态更新层的结构表达式为:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \quad (11)$$

1.2.2 LSTM 神经网络内部工作流程

(1) 遗忘阶段

该阶段是为了从上个节点获取到的输入信息中,提取重要信息并输出,过滤掉无用信息,以提高之后的计算效率。

(2) 选择性记忆阶段

该阶段将重要信息根据一定分类、筛选,并结合 x_t 进行关联记忆,来进行下一步的分析。

(3) 输出阶段

该阶段将当前需要有效信息来进行输出。

1.2.3 LSTM 神经网络总体结构

该文设计了一个由输入层、隐藏层、全连接层、分类层和输出层构成的蒸汽管道泄漏诊断模型,其结构如图 5 所示。

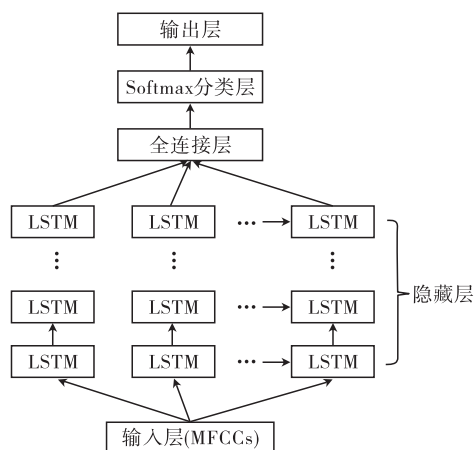


图5 LSTM神经网络总体结构示意图

表1 MFCC特征提取参数设置

AD 采样率/kHz	滤波器组/个	每帧长度	每帧位移	分段数据长度	重叠数据长度	小信号切除阈值/%	窗函数
16	26	256	80	200	100	5	Hamming

该文提取的 MFCC 特征参数包含静态特性参数和一阶差分动态特性参数 2 部分。

3 结论

该文深入探究了利用蒸汽泄漏时发出的高频声音进行蒸汽泄漏识别的主要方法,根据蒸汽泄漏点声音的特点及识别的需求,提出了基于 MFCC 特征提取及 LSTM 神经网络识别的蒸汽泄漏识别方法。

(1)输入层用于参数输入,输入的参数为定时从现场采集的声纹信号分帧后提取得到的 MFCCs 参数信息;

(2)隐藏层为多层 LSTM 构成的 LSTM 网络,用于对输入的 MFCCs 参数信息进行分类的多层次特征划分比较;

(3)全连接层用于增加模型的非线性表达能力,提高模型的学习能力;

(4)分类层用于接收处理过的特征数据,计算每个类别的可能性并进行分类,Softmax 的输出形式是 ≥ 1 的正整数,用于区分不同的类别;

(5)输出层将结果以标准接口的方式输出给其他系统。

2 实验结果及分析

该文用整列麦克风采集现场模拟蒸汽泄漏时的超声音频,用 Matlab 编制了相关的 MFCC 特征提取及 LSTM 神经网络识别软件,并进行了相关的实验测试。

MFCC 特征提取参数设置见表 1。实验中,声信号的 AD 采样率设为 16 kHz;滤波器组的个数设为 26 个,能够对接收声信号的整个频率范围进行良好的覆盖;每帧长度设为 256 个采样点,在 20 ~ 30 kHz 采样率下对应的时间长度为 16 ms,符合理论上对于声信号具有平稳性短时长度的限制性要求(10 ~ 30 ms);根据计算时使用的 Hamming 窗,每帧位移设为 80 个采样点,能够保证每帧之间有相对平滑的过度;根据每帧长度和每帧位移,分段数据长度设为 200,重叠数据长度设为 100;小信号切除阈值 5% 用来切除声信号数据序列中存在的微小噪声。

MFCC 在提取声音信号的特征方面性能良好,可以有效地提取故障信号与正常信号之间不同的时频分布特征,而 LSTM 神经网络可以克服长期以来传统神经网络面临的梯度消失等问题,与堆栈稀疏自编码器分类和 BP 神经网络等相比具有更高的分类准确率,因此,采用 MFCC 特征提取及 LSTM 神经网络方法可以有效的搭建蒸汽泄漏声音模型,实现现场蒸汽管道发生泄漏的准确诊断。

(下转第 51 页)

- 方法研究[J]. 中国安全科学学报, 2010, 20(7): 71.
- [5] 马龙. 抗电磁干扰技术在煤矿监控系统中的应用[J]. 煤矿安全, 2019, 50(11): 113 - 115.
- [6] 吴玉峰, 田彦文, 韩元山. 气体传感器研究进展和发展方向[J]. 计算机测量和控制, 2003, 11(10): 731 - 734.
- [7] 魏继涛. 煤气传感器材料的研究进展[J]. 传感器与微系统, 2015, 34(03): 4 - 7.
- [8] 石发强. 矿用产品金属外壳的电磁屏蔽效能分析[J]. 工矿自动化, 2014, 40(12): 18 - 21.
- [9] 朱前伟. 煤矿安全监控系统及组成设备抗干扰设计[J]. 工矿自动化, 2017, 43(6): 18 - 21.
- [10] 孙继平, 张宏伟, 张龙. 矩形弯曲隧道中电磁波的传输特性分析[J]. 工矿自动化, 2007(4): 4 - 6.
- [11] 石发强. 矿用分站外壳的电磁屏蔽效能分析[J]. 煤炭与化工, 2014, 37(2): 92 - 94.
- [12] 莫军. 滤波连接器在电磁兼容设计中的应用[J]. 国外电子测量技术, 2010, 29(6): 66 - 67.
- [13] 陈玉明, 吴克寿, 李向军. 一种基于信息熵的异常数据挖掘算法[J]. 控制与决策, 2013, 28(6): 867 - 872.
- [14] 龙英, 何怡刚, 张镇, 等. 基于信息熵和 Haar 小波变换的开关电流电路故障诊断新方法[J]. 仪器仪表学报, 2015, 36(3): 701 - 711.
- [15] 邹哲强, 庄捷, 屈世甲. 煤矿井下中低频段电磁干扰测量与分析[J]. 工矿自动化, 2013, 39(5): 1 - 5.
- [16] 赵北涛. 含分布式电源的配电变压器稳态电磁干扰抑制方法[J]. 电子设计工程, 2023, 31(21): 137 - 140.

(上接第 35 页)

参考文献:

- [1] 周萍, 沈昊, 郑凯鹏. 基于 MFCC 与 GFCC 混合特征参数的说话人识别[J]. 应用科学学报, 2019, 37(1): 24 - 32.
- [2] Kasiprasad M, Panyam N S, Maloji S. MFCC - GMM based accent recognition system for Telugu speech signals[J]. International Journal of Speech Technology, 2016(19): 87 - 93.
- [3] Ankur M, Divya K, R. K. Agarwal. Speaker recognition for hindi speech signal using MFCC - GMM approach[C]. 6th International Conference on Smart Computing and Communications (ICSCC), 2017: 880 - 887.
- [4] GREFF K, SRIVASTAVA R K, KOUTNIK J, et al. LSTM: A search space odyssey[J]. IEEE Transactions on Neural Networks & Learning Systems, 2017, 29(10): 2222 - 2232.
- [5] Himadri M, Sk Md Obaidullah, K. C. Santosh, et al. A lazy learning - based language identification from speech using MFCC - 2 features[J]. International Journal of Machine Learning and Cybernetics, 2020(11): 1 - 14.
- [6] Sakshi G, Ravi S S, Rajesh K. S, et al. Deep learning bidirectional LSTM based detection of prolongation and repetition in stuttered speech using weighted MFCC[J]. International Journal of Advanced Computer Science and Applications, 2020, 11(9): 345 - 456.
- [7] Samia A El - Moneim, M. A. Nassar, Moawad I. O, et al. Text - independent speaker recognition using LSTM - RNN and speech enhancement[J]. Multimedia Tools and Applications, 2020(79): 24013 - 24028.
- [8] 王鑫, 吴际, 刘超, 等. 基于 LSTM 循环神经网络的故障时间序列预测[J]. 北京航空航天大学学报, 2018, 44(4): 772 - 784.
- [9] 龚安, 马光明, 郭文婷, 等. 基于 LSTM 循环神经网络的核电设备状态预测[J]. 计算机技术与发展, 2019, 29(10): 41 - 45.
- [10] 李俊卿, 陈雅婷. 基于 LSTM 网络的同步电机励磁绕组匝间短路故障预警[J]. 电力科学与工程, 2020, 36(6): 37 - 42.
- [11] 牟俊杰, 姚刚, 孙涛. 基于 CNN - LSTM 神经网络的声纹识别系统设计[J]. 电子技术应用, 2021, 47(3): 75 - 78.
- [12] 王志成, 张玉一, 巴天星, 等. 面向 CIM 和动态交通分析的多源异构数据融合技术研究[J]. 电子设计工程, 2024, 32(8): 68 - 72.

欢迎投稿！ 欢迎订阅！ 欢迎刊登广告！

国内邮发代号: 52 - 49 国际发行代号: BM529 定价: 18.00 元/期 108.00 元/年
地址: 西安市高新区沣惠南路 8 号 邮编: 710075 电话: 029 - 81871277
网址: <http://yb-zdh.shaangu-group.com> 电子邮箱: gyybbjb@126.com