

抗环境音干扰的设备声音故障监测方法^①



李兰村^{1,2}, 廉东本², 毛立爽²

¹(中国科学院大学, 北京 100049)

²(中国科学院 沈阳计算技术研究所, 沈阳 110168)

通讯作者: 李兰村, E-mail: 530468294@qq.com

摘要: 变压器等大型设备在运行过程中发声具有辨识性和平稳性的特点, 但容易受各种环境音的干扰, 针对该问题, 本文利用声音信号处理、特征提取、模式匹配等技术, 提出了一种抗多种环境音干扰的设备声音故障监测方案. 首先对在各种环境声音中变压器的正常和故障声音进行采集和预处理, 然后对其提取出 MFCC 特征并降维, 对变压器正常工作声音特征通过 OPTICS 算法进行训练, 得到一个具有多个分类的标准集, 最后将标准集与包含故障声音的测试样本进行匹配, 若出现不匹配情况但经人工检验为误报, 则将其归为新的分类. 实验结果表明: 该方法不仅能很好的识别样本, 也能在新的正常声音出现时通过标准集增强模块来优化标准集, 从而提高识别准确率并降低误警率.

关键词: 特征提取; 聚类算法; 数据降维; 声音识别; 故障监测

引用格式: 李兰村, 廉东本, 毛立爽. 抗环境音干扰的设备声音故障监测方法. 计算机系统应用, 2019, 28(6): 89-94. <http://www.c-s-a.org.cn/1003-3254/6930.html>

Device Sound Fault Monitoring Method of Anti-Environmental Sound Interference

LI Lan-Cun^{1,2}, LIAN Dong-Ben², MAO Li-Shuang²

¹(University of Chinese Academy of Sciences, Beijing 100049, China)

²(Shenyang Institute of Computing Technology, Chinese Academy of Sciences, Shenyang 110168, China)

Abstract: Large equipment such as transformers has the characteristics of identification and stability during operation, but it is easily interfered by various environmental sounds. To solve this problem, by using sound signal processing, feature extraction, pattern matching, and other techniques, this study proposes a device sound fault monitoring scheme that is resistant to multiple environmental sound disturbances. First of all, the normal and faulty sounds of transformers in various ambient sounds are collected and preprocessed. Then, MFCC features are extracted and dimensionality is reduced. Next, the normal working sound characteristics of the transformer are trained through the OPTICS algorithm to obtain a standard set with multiple clusters. Last, the standard set is matched with the test sample containing the faulty sound. If there is a mismatch, but the manual test is a false positive, it will be classified as a new cluster. The experimental results show that the proposed method can not only identify the sample well, but also optimize the standard set through the standard set enhancement module when the new normal sound appears, thus improving the recognition accuracy and reducing the false alarm rate.

Key words: feature extraction; clustering algorithm; data reduction; voice recognition; fault monitoring

① 收稿时间: 2018-12-10; 修改时间: 2018-12-29; 采用时间: 2019-01-08; csa 在线出版时间: 2019-05-25

1 引言

变压器等大型设备在运行时的声音具有一定的辨识性,目前普遍通过拥有多年设备检修经验的技术人员对变压器运行声音进行监听,并判断设备是否正常运行.但该方式有两个不足之处,一是需要技术人员时刻在设备前监测设备运行声音,若在恶劣环境条件下安全难以保障;二是人耳的听觉具有特定的听觉感知范围,只能感知特定频率范围和特定响度范围内的声音^[1].近年来,多个领域借鉴语音识别技术取得了丰富的研究成果.通过在变压器旁安装拾音器采集声音并对其进行实时监测,不仅不会干扰变压器的正常运行,而且装置安装简单,采集信号方便.但在实际环境中存在复杂环境噪声,可能会对采集的声音信号造成一定影响,从而导致声音识别的准确率下降.

现有的研究中通常采用分类算法建立声音模型,例如文献^[2]中作者提出了一种基于改进的支持向量机分类算法建立声音模型的解决方案.但分类算法只能在分类个数已知的条件下建立声音模型,对于由复杂环境音而导致分类个数不确定的条件下建立声音模型的情形并不适用.此外,现有研究对建立的声音模型采用剥离环境音的方式检测设备是否发生故障,该种检测方式在环境音影响较小的条件下对声音模型的去噪效果理想,但在多种复杂环境音的影响下,对声音模型的去噪将变得困难且效果不理想,从而导致较高的误警率,并降低整体的检测效率.本文提出的方法是对保留环境音的设备工作声音数据进行聚类,并对得到的训练模型进行迭代优化.

2 模型构建

本文采用的模型主要由数据预处理模块、特征提取模块、特征降维模块、特征训练模块、标准集增强模块构成.模型数据流程图如图1所示.

原始声音数据经过预处理、特征提取和降维后分成两部分,一部分用于训练标准集,另一部分用于匹配增强标准集.

2.1 数据预处理模块

采集运行中的设备声音时,由于拾音器参数不同或拾音器所放置的方向和距离的不同会导致采集的声音品质不同,因此需要对采集到的声音进行归一化处理,降低不同样本间因采集因素造成的差异.

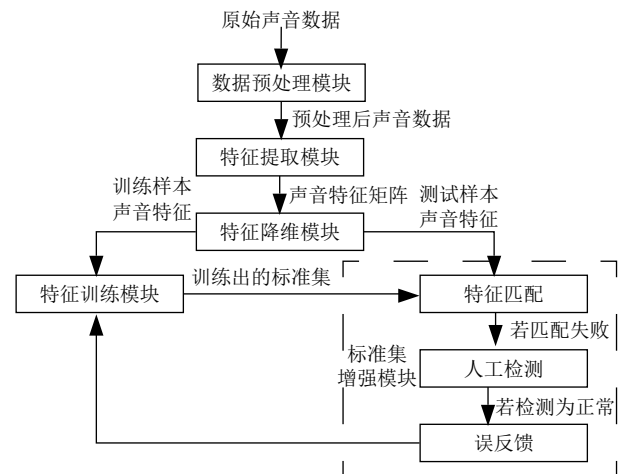


图1 模型数据流程图

处理声音信号时,需使用快速傅里叶变换(FFT)明确声音中各个频率成分的分布.傅里叶变换要求输入的信号是平稳的,虽然声音信号从时间轴上看具有非平稳特性,但在极短的时间内,则呈现平稳信号所具有的特性.因此,将一段声音信号切分成若干小段,就能将其近似地认为是平稳的信号,且该小段被称为帧长.但声音信号具有连续性和相关性,因此不能讲声音直接均分.且为了提高傅里叶变换的分辨率,在变换前需要对每一帧信号进行“加窗”,即将其与一个窗函数相乘,使得该帧信号的两端逐渐趋于零.常见的窗函数有汉宁窗和矩形窗,但汉宁窗的旁瓣衰减较大,具有更平滑的低通特性,能够在较高程度上反应短时信号的频率特性.而矩形窗的旁瓣峰值较大,导致频谱泄露较为严重^[2],因此本文采用汉宁窗函数.由于加窗操作会将一帧信号的两端趋于0,所以需要截取的后一帧向前一帧移动一段距离,称为帧移,一般重叠距离为帧长的1/3至1/2.汉宁窗函数的公式如下:

$$W(n, a) = (1 - a) - a * \cos \left[\frac{2\pi n}{N-1} \right] \quad (1)$$

式(1)中, a 一般取 0.46.

2.2 特征提取模块

通过对声音信号特征的提取,分析和处理,最终为之后的处理得到主要的声音特征数据.选取最优的声音特征参数对于声音特性的识别监测至关重要,只有根据尽量降低计算复杂度和降低特征维数的要求来选择不同的声音特征参数,才能保证故障监测系统的识别正确率和效率.在非语音类声音的模式识别领域,因

为声音信号在时域上变化快速且不稳定,所以通常将其转换到频域上分析特征参数.倒频谱特征参数因为具有分离频谱上的高低频的优点而被广泛使用^[1],常用的倒频谱特征参数有梅尔频率倒谱系数(MFCC)、LPC倒谱系数(LPCC)、线性预测倒谱系数(LPC)等.由于梅尔频率倒谱系数是根据人耳频率的非线性特性而提取出来的倒谱参数,梅尔频率滤波器组模拟人的耳蜗声道模型,实现了模拟人类特定的听觉掩蔽和听觉修复等功能,所以更加适用于需要人耳判别声音的场景,本文选用提取MFCC来转化声音数据,梅尔标度与频率的关系如式(2)所示:

$$Mel(f) = 2595 \times \lg \left(1 + \frac{f}{700} \right) \quad (2)$$

式中 f 代表频率.

提取MFCC系数首先要将预处理后的声音信号进行FFT,计算出每帧数据的频谱参数,以得到在频谱上的能量分布.其次,设置M(24)个带通滤波器,每个滤波器具有三角形滤波器的特性,可以对频谱起到平滑化且消除谐波的作用.然后,计算每个滤波器组输出的对数能量.最后,将对数能量经过离散余弦变换,求出 1×12 的MFCC特征系数矢量.根据MFCC提取声音特征矩阵流程如图2所示.

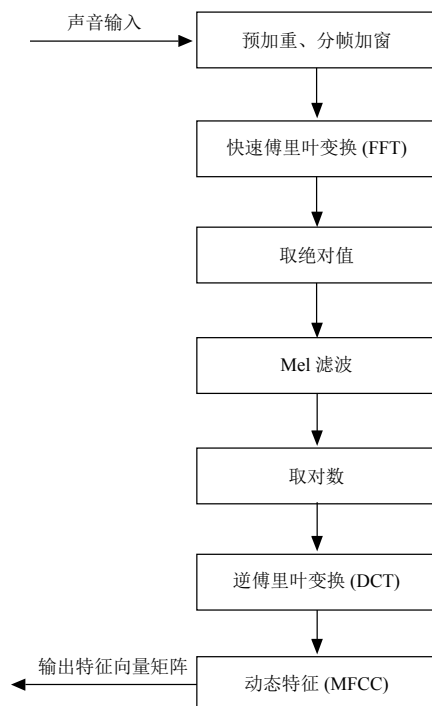


图2 提取MFCC特征矩阵流程

2.3 特征降维模块

提取出的MFCC声音特征参数存在数据维度高、数据有冗余等缺陷从而导致存储空间占用过多和计算量过大等问题.通过降维可以舍弃贡献度不高的维度,从而达到减少计算量和计算时间的目的.此外,提取到的声音中包含复杂的环境音,当数据受到噪声影响时,低贡献度维度数据往往与噪声有关,因此舍弃这部分数据也能减少噪声对数据本身的影响.由于提取到的声音数据特征在不同维度上的方差分布不均匀,因此更适合采用PCA降维算法.PCA通过计算声音数据的协方差矩阵及其特征值分解得到特征值和相应的特征向量划分数据的主成分.PCA降维算法的目标是去掉不同数据维度之间的相关性,并尽量使降维后的数据维度能够很好地保持原始数据的特征.

PCA降维首先要对声音数据标准化,使得每个维度的数据均落入相同的区间;其次,获取声音特征矩阵的协方差矩阵;然后对协方差矩阵进行特征值分解,获取各维度的特征值和对应的特征向量;最后,将特征值降序排列后通过维度选取算法获取降维幅度.

在降维的过程中,对降维幅度的选取至关重要.在现有的研究中,普遍采用交叉验证或逆向重构两种方法来获取最优的降维幅度.为避免额外开销(如学习器等),本文采用逆向重构的方法获取最优降维幅度.根据实际生产经验及研究,本文设定重构阈值 t 的值为95%^[3],然后选取使得式(3)成立的最小维度.

$$\frac{\sum_{i=1}^{d'} \lambda_i}{\sum_{i=1}^d \lambda_i} \geq t \quad (3)$$

其中, d 为原始数据的维度, d' 为降维后的维度, λ_i 为该维度特征向量的特征值.

2.4 特征训练模块

标准集的设计需要能使训练数据获取最优分类,由于设备声音在运行过程中会与环境音混合,设备正常工作声音的种类数量及声音类别标号均未知,因此需要提前确定类别标号的有监督分类学习算法是不可行的^[4-6],综上,本文采用无监督的聚类学习算法.

聚类处理将数据对象集划分成多个簇,使得簇内的对象具有较高的相似性,且与其他簇中对象的相似性较低.通过聚类算法能发现任意形状的簇,且具有处理噪声点数据的能力,这些优点在变压器声音数据进

行模式匹配时具有辅助作用. 在基于密度的聚类算法中比较常见的有 DBSCAN(Density-Based Spatial Clustering of Application with Noise) 和 OPTICS (Ordering Point To Identify the Cluster Structure), 两者较为相似, 均通过核心对象及其邻域发现稠密区域, 该稠密区域被称为簇.

DNSCAN 算法中有两个重要的参数: 分别为 Eps (定义密度时的邻域半径) 和 MinPts(定义核心点时的阈值), 本文将其简化为 ε 和 M . 对于数据集 $X=\{x^{(1)}, x^{(2)}, \dots, x^{(N)}\}$, 引入以下概念:

(1) 邻域

设 $x \in X$, 则

$$N_{\varepsilon}(x) = \{y \in X : d(y, x) \leq \varepsilon\} \quad (4)$$

为 x 的 ε 邻域.

(2) 密度

设 $x \in X$, 则

$$\rho(x) = |N_{\varepsilon}(x)| \quad (5)$$

为 x 的密度.

(3) 核心点

设 $x \in X$, 若 $\rho(x) \geq M$, 则 x 称为 X 的核心点, X 中所有核心点组成的集合记为 X_c , X 中所有非核心点组成的集合为 X_{nc} .

(4) 边界点

若 $x \in X_{nc}$, 且 x 落在某个核心点的 ε 邻域内, 则称 x 为 X 的一个边界点, X 中所有边界点组成的集合记为 X_{bd} .

(5) 噪音点

设 $X_{noi} = X \setminus (X_c \cup X_{bd})$, 若 $x \in X_{noi}$, 则称 x 为噪音点.

(6) 直接密度可达

设 $x, y \in X$, 若 $x \in X_c$, $y \in N_{\varepsilon}(x)$, 则称 y 由 x 直接密度可达.

(7) 密度可达

设 $P^{(1)}, P^{(2)}, \dots, P^{(n)} \in X (n \geq 2)$, 若满足 $P^{(i+1)}$ 是从 $P^{(i)}$ 直接密度可达的, 则称 $P^{(n)}$ 是从 $P^{(1)}$ 密度可达的.

(8) 密度相连

设 $x, y, z \in X$, 若 y, z 都可由 x 密度可达, 则称 y, z 密度相连.

DBSCAN 算法的核心思想是: 从某个随机选定的核心点开始, 不断向密度可达的区域扩张, 最终得到一个包含核心点和边界点的簇, 使得簇中任意两点密度相连. 由于 DBSCAN 算法初始时设定了两个固定的输

入参数 ε 和 M , 这需要提前根据经验来设定, 设置的细微不同可能造成聚类结果的巨大差异. M 的设定可以根据一个原则, 即 $M \geq m+1$ (m 表示数据维度), 但 ε 的设定是比较困难的, 因为对密度差异较大的数据很难界定一个密度, 使得边界点不被遗落且有差异的簇不被归到一个簇中.

OPTICS 算法不需要设置全局参数, 而是为聚类分析生成一个增广簇排序, 它代表了各样本点基于密度的聚类结构, 从该排序中可以获得基于任何半径和任何阈值的聚类^[7]. OPTICS 算法是对 DBSCAN 算法的一种改进, 因此上文提到的关于 DBSCAN 算法中的概念在 OPTICS 算法中依然适用, 但增加了以下两个概念.

(9) 核心距离

设 $x \in X$, 对于给定的参数 ε 和 M , 称使得 x 成为核心点的最小邻域半径为 x 的核心距离, 记为 $cd(x)$.

(10) 可达距离

设 $x, y \in X$, 若 $N_{\varepsilon}(x) < M$, 则

$$rd(y, x) = \max\{cd(x), d(x, y)\} \quad (6)$$

OPTICS 算法过程如下:

算法 1. OPTICS 算法

输入: 数据样本 X , 初始化所有点的可达距离与核心距离均为 MAX, 并初始化邻域半径 ε 和阈值 M .

- 1) 建立两个队列, 有序队列包含核心点及该核心点的直接密度可达点, 结果队列包含存储样本输出及处理次序;
- 2) 若 X 中数据全部处理完毕, 则转到步骤 7), 否则从 X 中选择一个为核心对象且未处理过的点, 将此核心点放入结果队列, 将此核心点的直接密度可达点按可达距离升序排列放入有序队列;
- 3) 若有序队列为空, 回到步骤 2), 否则从有序队列中取出第一个点;
- 4) 若该点不是核心点则回到步骤 3), 若该点是核心点是且该点不在结果队列中则将点放入结果队列;
- 5) 若该点是核心点, 找到该点所有直接密度可达点, 并将这些点放入有序队列中, 且将有序队列中的点按照可达距离重新排序, 若该点已经在有序队列中且新的可达距离更小, 则更新该点的可达距离.
- 6) 重复步骤 3)、4)、5), 直至有序队列为空.
- 7) 算法结束.

至此, 通过算法获取到一个有序数组和每个元素的核心距离和可达距离, 然后利用这些数据生成一个聚类结果, 加入一个邻域半径参数 ζ , $\zeta < \varepsilon$. 由于需要将数据集分成 k 个簇及噪音点集合, 因此引入簇标记数组如下:

$$m_i = \begin{cases} j (j > 0), & \text{若 } x^{(i)} \text{ 属于第 } j \text{ 个簇} \\ -1, & \text{若 } x^{(i)} \text{ 为噪音点} \end{cases} \quad (7)$$

最终仅需生成一个标记数组 $\{m_i\}_{i=1}^N$, 便可生成一个

聚类结果。

2.5 标准集增强模块

将训练样本数据通过特征训练模块训练后得到了具有不同密度的簇,每一簇代表一种噪声混合变压器正常工作声音或仅为变压器正常工作声音。这些簇均属于变压器在正常工作环境下会形成的簇,但簇却是不完整的。由于环境干扰音种类繁多,在前期无法把各种环境干扰音全部采集到,若在对声音特征进行匹配识别时,出现了没有经过训练的环境干扰音,则会出现匹配失败的情况,但该声音的确是正常的声音,因此为了避免多次出现类似的情况,本文引入标准集增强模块,标准集增强模块能将新出现的正常声音加入到标准集当中。

标准集增强模块的流程如下:

(1) 在已经存在标准集的情况下,假设已知参考集为 $C=\{C_1, C_2, \dots, C_n\}$, 其中每一类代表一个簇,假设增加的测试样本数据为 $T=\{t^{(1)}, t^{(2)}, \dots, t^{(s)}\}$, 其中 s 表示样本数据的帧数。当前的测试样本数据中既存在非正常数据帧也存在提前标记完成的正常数据帧;

(2) 将 T 中的每帧与 C 中每个簇按照式 (8) 进行欧式距离的运算。

$$d_{ij} = \sqrt{\sum_{k=0}^n (c_{ik} - t_{ik})^2} \quad (8)$$

式 (8) 中, i 表示样本帧号, j 表示参考集簇的序号, d_{ij} 表示第 i 帧样本数据与第 j 个簇之间的距离, 得到每帧与各个簇之间的距离为: $d_{i1}, d_{i2}, \dots, d_{in}$, 其中 i 为帧号, n 为簇的个数, 取

$$d_{\min} = \text{Min}(d_{i1}, d_{i2}, \dots, d_{in}) \quad (9)$$

若 d_{\min} 小于提前设定好的阈值, 则标记为正常数据, 否则, 标记为异常数据。

(3) 比较数据匹配前后的标记, 若正常数据帧被标记为非正常数据, 说明该正常数据在初始的训练集中并没有出现过, 便将其进行聚类, 形成新的一个簇, 加入到 C 中, 此时簇 $C=\{C_1, C_2, \dots, C_n, C_{n+1}\}$ 。

(4) 至此一次标准集增强处理完毕, 若再有新数据的加入, 则进行增强处理, 训练集就会一直优化迭代。

数据经过标准集增强模块后, 在标准集中没有出现过且的确属于设备正常工作声音种类会被加入到标准集中, 进而在下次与样本声音数据进行匹配时候将该类声音识别为正常声音, 而不会将其识别为异常声音进而误警。本文通过实验, 验证经过标准集增强模

块后的训练集在与测试数据匹配时的误警率随着迭代次数增加而逐渐降低。

3 实验分析

3.1 实验环境

实验运行在 Windows 10 操作系统上, 实验 PC 机主频为 3.6 GHz, 内存 4 GB。由于变压器故障样本获取非常困难, 为了测试算法对故障情况的识别能力, 本文合成了一批故障样本。本文对样本进行筛选并选取了有代表性的声音种类, 包含正常声音种类 20 种, 非正常的故障声音 3 种, 每种声音截取 10 分钟。训练样本从正常声音数据中随机选择 15 种, 测试样本个数为 40 个, 且每一个测试样本均随机从所有声音数据中选择 5 种。

3.2 实验设计和结果

首先将训练样本和参考样本进行预处理和特征提取。在分帧阶段, 设置帧长为 40 ms, 帧移为 20 ms, 故训练样本共有 4.5×10^5 帧, 每个测试样本有 1.5×10^5 帧。在加窗阶段对声音的每一帧乘以汉宁窗函数。对预处理后的数据提取 MFCC 特征参数并将其进行标准化处理后的特征矩阵如图 3 所示。

$$\begin{bmatrix} -1.70352, -2.242446, 0.71540, \dots, 1.91513, -1.79590 \\ -1.78609, -2.13796, 1.02629, \dots, 1.41031, -0.91942 \\ \dots \\ -1.40910, 1.14897, -0.79088, \dots, -2.16831, -0.85796 \\ -2.68992, -0.36141, 0.11810, \dots, -1.52901, -0.64068 \end{bmatrix}$$

图 3 MFCC 特征矩阵

其中特征矩阵共 12 列, 每一列代表每一帧的一个属性值, 特征矩阵共 640000 行, 每一行代表初始声音数据中的一帧。

然后计算该 MFCC 特征矩阵的协方差矩阵, 并将该协方差矩阵进行特征值分解, 得到如图 4 所示的降序排列的 12 个特征值及每个特征值对应的特征向量。

$$\begin{bmatrix} 33.4840, 14.9411, 5.4393, 0.5052, 0.3659, 0.3380, \\ 0.1510, 0.1224, 0.0941, 0.0027, 0.0016, 0.0002 \end{bmatrix}$$

图 4 MFCC 特征矩阵的特征值集合

根据式 (3) 可知, 仅用前三个特征值对应的特征向量便可使重构率达到 97.15%, 满足不小于重构阈值 95% 的要求, 因此可以使用 PCA 算法将原矩阵降维至 3 维矩阵, 降维后的特征矩阵如图 5 所示。

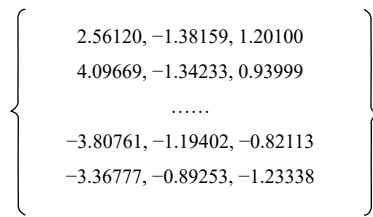


图5 降维后的特征矩阵

对降维后的训练样本声音特征矩阵采用 OPTICS 算法进行聚类, 根据实际开发经验及文献研究, 本实验采用的 OPTICS 算法中邻域半径 ϵ 的值为 0.24, 阈值 M 的值为 16 时, 聚类结果最好, 然后将测试样本依次与标准集进行匹配. 由于变压器正常工作声音数据与标准集偏差小于 0.48, 而变压器故障声音数据与标准集偏差大于 1.5, 因此为减少模型对噪音点的敏感度, 本实验将匹配阈值设置为 0.5, 并记录下每次匹配增强后的标准集的结果. 分别对原始标准集和经过 40 次测试样本迭代增强后的标准集计算正常率和误警率, 得到的结果如表 1 所示.

表 1 模型识别正常率和误警率 (%)

迭代次数	正常率	误警率
5 次迭代后	20.0	40.0
10	20.0	44.4
15	20.0	35.7
20	25.0	29.4
25	36.0	27.8
30	36.7	23.8
35	40.0	21.7
40	42.5	17.9

图 6 为模型的准确率和误警率随着迭代次数的增加而改变的折线图.

3.3 实验结果分析

从表 1 与图 6 可以看出, 模型的正常率随着迭代次数的增加不断升高, 因为设备故障声音数据占总样本数据的 13%, 且每次迭代选择的样本数据比较大, 所以正常率最终只会逼近 $C_{20}^5 / C_{23}^5 = 46.1\%$. 但实际生产中变压器发生故障的情况较少, 所占总样本数据的比例不高, 因此实际生产中正常率会远高于当前数值. 误警率是一个更能反映模型优势的指标, 可以看出, 迭代 10 次后的误警率比迭代 5 次后要高, 因为在迭代前期, 新出现的正常声音数据较多, 出现误报的次数就相对较多. 当新出现的正常声音数据均存入标准集后, 随着迭代次数的增加, 误警率将不断降低. 实验结果说明, 本文提出的模型能准确识别非正常声音, 且随着迭代

次数的增加, 模型的误警率将持续降低, 可信赖程度将不断升高.

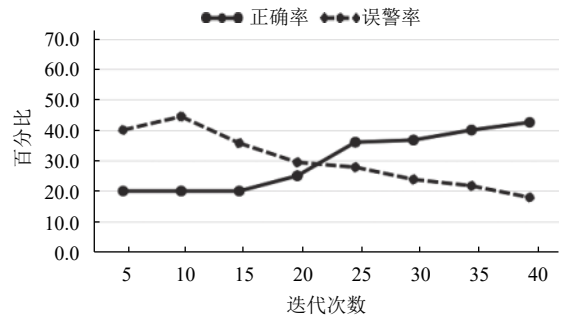


图 6 模型识别正常率和误警率折线图

4 结论与展望

本文基于变压器等大型设备在运行过程中发声具有辨识性和平稳性, 且容易受各种环境音干扰的特点, 提出了一种抗多种环境音干扰的设备声音故障监测方法. 一方面, 将样本数据经过特征提取与 OPTICS 聚类模块得到初始标准集, 另一方面, 将测试样本与已有标准集不断匹配和增强获得更优标准集. 通过实验, 验证了本文提出的方法具有可行性和有效性, 可以提高设备状态监测数据的可靠性, 进而有助于更快推进变压器等大型设备自动监测的进程.

参考文献

- 李晶, 孙农亮, 滕升华. 基于声音识别的设备状态检测算法. 信息技术, 2015, (6): 94-98. [doi: 10.3969/j.issn.1672-9528.2015.06.029]
- 贺玲玲, 周元. 基于改进 MFCC 的异常声音识别算法. 重庆工商大学学报 (自然科学版), 2012, 29(2): 52-57. [doi: 10.3969/j.issn.1672-058X.2012.02.013]
- 李蝉娟. 高维数据降维处理关键技术研究[硕士学位论文]. 成都: 电子科技大学, 2017.
- 柯登峰, 徐波. 互联网时代语音识别基本问题. 中国科学: 信息科学, 2013, 43(12): 1578-1597.
- Banerjee A, Abu-Mahfouz I. Evolutionary clustering algorithms for relational data. Procedia Computer Science, 2018, 140: 276-283. [doi: 10.1016/j.procs.2018.10.319]
- Kumar KM, Reddy ARM. An efficient k-means clustering filtering algorithm using density based initial cluster centers. Information Sciences, 2017, 418-419: 286-301. [doi: 10.1016/j.ins.2017.07.036]
- 张强, 王序文, 王小捷, 等. 基于 OPTICS 的变电设备状态监测异常数据过滤算法. 电力信息与通信技术, 2015, 13(6): 8-14.