基于时、频域自相似性的配电变压器环境噪声识别方法

刘 元¹,苏 盛¹,刘正谊²,夏云峰³,刘贯科³,李 彬¹
 (1. 长沙理工大学 电网新能源防灾减灾研究中心,湖南 长沙 410004;
 2. 国网常德供电公司,湖南 常德 415000;3. 广东电网有限责任公司东莞供电局,广东 东莞 523000)

摘要:配电变压器位于开放的嘈杂环境中,环境噪声对基于声音的设备运行状态在线监测具有突出影响。如 何识别和剔除含噪声干扰的录音片段,是推进基于声音的配电变压器状态在线监测亟待解决的瓶颈问题。 分析指出配电设备正常和异常的运行声音与各种环境噪声在时域和频域上的差异性特征,提出基于时、频域 自相似性的配电设备录音数据中含环境噪声片段的识别方法。对分钟级录音数据分帧后提炼每帧的时域和 频域特征指标,通过聚类判断该录音片段是否具有自相似性,识别和剔除含环境噪声的录音数据。实验结果 表明,所提方法能够有效识别和剔除含环境噪声的录音片段,从而得到环境噪声含量较低的录音数据,为后 续开展基于声音信号的配电变压器状态监测提供有力支撑。

关键词:声音信号;状态监测;配电变压器;噪声消除

文献标志码:A

DOI:10.16081/j.epae.202111021

0 引言

中图分类号:TM 41

电力设备运行时发出的声音信号能有效反映设 备运行状态。利用非接触式声音传感器采集电力设 备运行声音信号,利用声音信号分析和诊断运行状 态,具有安装灵活、信息量丰富和可靠性高等优势, 是电力设备状态在线监测的重要发展方向。科研人 员围绕利用声音监测电力设备运行状态开展了大量 研究^[1-3],有力地推动了基于声音的在线监测技术的 工程应用。分散分布的配电变压器由于数量庞大、 巡检周期长,存在的故障隐患问题更突出。基于声 音的无接触式在线监测技术有望发展为提高配电变 压器运维水平的重要手段。需要指出的是,配电变 压器多布设于街头巷尾的嘈杂环境中,环境噪声将 对基于声音的设备运行状态诊断造成较大的影响。 如何消除环境噪声的干扰,是利用声音监测配电变 压器运行状态的关键。

已有的电力设备状态监测研究中,可听声音消噪主要针对变电站内变压器、电抗器等设备,此类应用场景下噪声干扰来源明确,目前主要有2种消噪思路。

1)采用滤波消除较强信号中的微弱干扰,一般 采用小波分析结合软、硬阈值滤波等方式,可在弱噪 声条件下取得良好的消噪效果^[45]。由于电力设备

收稿日期:2021-02-01;修回日期:2021-09-29

基金项目:国家自然科学基金资助项目(51777015);湖南省自 然科学基金资助项目(2020JJ4611);湖南省教育厅重点项目 (19A011)

Project supported by the National Natural Science Foundation of China(51777015), the Natural Science Foundation of Hunan Province(2020JJ4611) and the Key Project of Hunan Provincial Department of Education(19A011) 运行环境的噪声具有强不确定性,该方法主要用于 消除微弱噪声,或信号和噪声的频段相互分离的确 定性噪声,如滤除频率集中在低频段的冷却风扇噪 声,但也存在难以滤除强背景噪声、参数控制困难、 容易削弱有效数据的问题。

2)采用盲源分离技术从观测信号中恢复辐射源 信号,该思路以基于独立分量分析和基于稀疏表示 理论的稀疏分量分析的信号分离方法为主。其中, 基于独立分量分析的信号分离方法在变压器振动、 局部放电检测等领域应用较多[6-7]。但该方法存在 非线性欠定问题,需要满足观测信号数目大于信号 源数目及观测信号具有线性可叠加性的条件。由于 干扰信号源数目难以预知,只能大量增加声音传感 器。且受声音反射、折射等因素影响,在不同位置测 量的电力设备声音并不具有严格的线性可叠加性, 这制约了该方法的推广应用。基于稀疏表示理论的 稀疏分量分析的信号分离方法通过对信号的稀疏表 示解决非线性欠定性问题,从含环境噪声干扰的变 压器运行声音中分离本体振动声音[8-9],再根据变压 器运行声音在频域上具有稀疏性的特点,提前录制 噪声干扰较小的变压器声音作为样本集,训练得到 能标识样本的训练字典,并挑选出与新采集样本最 为匹配的字典后进行重构。但根据含强环境噪声的 声音信号重构设备运行声音,可能会改变声音对应 的设备运行状态,难以真实还原声音构成。

此外,经验模态分解和变模态分解相结合的算法近年来也被用于信号消噪^[10]。但前者存在模态混 叠问题,后者存在需提前确定模态数量等问题,使得 该类算法还需得到进一步优化。

配电变压器运行在开放、嘈杂的环境中,环境噪 声具有强不确定性,难以预判一般性规律,传统的噪 声消除方法难以适用,容易将含噪声的声音监测样 本判断为设备异常状态。因此,有必要研究嘈杂环 境下环境噪声干扰的识别和消除方法,推动基于声 音的配电设备状态监测技术的发展。

因为电力设备故障多为慢变过程,状态异常检 测的时效性要求不高,加之基于声音的在线监测系 统可持续不断地记录设备声音数据,只要能准确识 别和剔除含环境噪声干扰的录音样本,仍可利用含 环境噪声较小的录音样本聚类识别设备运行状态。 本文利用正常和故障异常状态下电力设备运行声音 与环境噪声的特性差异,提出了一种基于时、频域自 相似性的环境噪声识别与剔除方法,将每分钟的录 音数据分帧后提取每帧的时、频域特征,利用基于 MeanShift聚类的相似度分析方法对特征进行相似 度分析,识别其中是否包含不具有时、频域自相似性 的环境噪声,从而筛选出有效样本,为后续基于声音 信号识别电力设备运行状态监测提供支撑。

1 配电变压器与环境噪声特性分析

1.1 配电变压器运行声音的时、频域特征

变压器振动发出的声音主要由铁芯和绕组振动 产生^[12]。其中,铁芯振动是由硅钢片的磁致伸缩效 应、硅钢片接缝处的电磁力引起的,绕组振动是由通 电绕组在漏磁场中受力引起的。在分钟级时间尺度 上,电力设备的电气量参数波动有限,振动声音信号 的时、频域特征随时间变化较小,具有稳定性和自相 似性。配电变压器在轻载、重载以及底座螺栓松动故 障下发出声音的时域和时频域图像见附录A图A1。

1)从图A1左边的时域波形图能直观看出,声音 强度在轻载时波动幅度很小,重载时声音强度略有 加大且频繁出现小幅毛刺;底座螺栓松动时发出声 音的强度明显高于正常运行场景,也频繁出现小幅 毛刺。

2)图 A1 右边的时频域特征图展示了频率成分 随时间变化的情况,颜色对应能量大小,颜色越亮的 区域对应频率成分的能量越强。轻载时频率成分在 60 s内基本平稳;重载时频率成分和频率波动与轻 载时大致相同,但低频能量明显更强;底座螺栓松动 时,出现了较多的高频分量,包含低频能量在内的各 部分能量进一步加强。此外,在3种故障下频率分 量在60 s内均基本平稳。

1.2 典型环境噪声的时、频域特征

配电设备运行环境中的可听声音可分为设备运 行声音和环境噪声两大类,如表1所示。第1类是配 电房或环网柜内目标监测配电设备运行时的可听声 音,主要包括正常运行时的设备振动声音和异常时 的振动、放电声音;第2类是周围环境噪声,如语音、 车辆经过、蝉虫鸣叫、蛙鸣、风声、雨声、雷电声等。 其中,电力设备在正常和异常状态下发出的声音均 具有稳定持续的特性,而环境噪声均为短期内不稳 定变化的声音,无论在时域上还是在频域上都不具 有持续稳定的特性。含典型环境噪声的录音数据的 时域和时频域图像见附录A图A2。

表1 配电设备运行环境中的可听声音分类

Table 1 Classification of audible sound in operation environment of distribution equipment

声音类型	包括内容	特征(分钟级别)
设备运行 声音	正常运行时的设备振动 声音以及异常时的 振动、放电声音	稳定持续,时、频域 特征变化小
环境噪声	语音、车辆经过、 蝉虫鸣叫、蛙鸣等	不稳定,时、频域 特征变化大
	时间较长的下雨、 刮风等发出的声音、 雷电声等	不稳定,呈无规律的 起伏状态,时、频域 特征变化大

1)图 A2(a)所示的录音包含语音和车辆经过的 噪声。由图可见:语音大小和频率无规律,不具有平 稳性,与配电设备运行声音存在较大差异;车辆经过 时发出的声音随车辆与配电设备的距离的变化而持 续波动,车辆从远处驶来,经过配电设备且车身侧面 正对配电设备时,发出的声音先增大后减小。

2)图 A2(b)所示的录音包含蝉鸣和鸟叫,由图 可见,两者在时域上部分区域存在规律的脉冲,在短 时间内具有一定的周期性。但动物鸣叫是靠发声器 官的肌肉收缩,声音随时间变化较大,不会长期持续 稳定存在。

3)图 A2(c)为下雨时的录音数据,雨点落在声 音传感器附近造成类似脉冲的噪声干扰,使得声音 的时域波形存在较多毛刺,在时频域图像中表现为 颜色较亮的条纹。

4)图 A2(d)为大风天气下的录音。刮风时风速 的起伏将在经过声音传感器时留下呼啸声,随风速 变化产生较明显的不稳定噪声。

综上所述,配电设备运行环境噪声在时域和时 频域上具有和设备运行声音显著不同的差异性特征,可考虑利用这一差异识别和剔除含环境噪声的 录音数据。

2 基于时、频域自相似性的噪声识别方法

2.1 噪声识别原理

电力设备的振动类和放电类故障为慢变过程, 异常声音将持续存在。在分钟级时间尺度上,电力 设备的电压、电流等参数波动幅度有限,发出声音的 时域和频域特征随时间变化较小,在时域和频域上 具有较高的自相似性。而环境噪声在时域和频域上 随时间有较明显的变化,不会长期持续,并不具有自 相似性。将分钟级的声音样本分帧分割为若干片段 后,可对各帧进行相似性分析,发现样本声音在时、频 域上不具有相似性时,可判断该样本为含噪声样本。

利用环境噪声不具有自相似性的特点,按图1所 示流程,采用无需预设类簇数的MeanShift聚类算法 判断录音数据的自相似性。将录制的分钟级声音分 帧分割为若干秒级片段,然后提取每帧的时域能量、 小波包频域能量特征,作为特征向量进行聚类。不含 噪声的声音片段在时、频域特征上具有较高的相似 性,会紧密聚成一个类簇,而含噪声的声音片段和不 含噪声的片段具有差异性,会被识别为不同的类簇。



图1 含环境噪声样本的识别、剔除流程

Fig.1 Identification and elimination flowchart of samples with environmental noise

2.2 预处理及特征提取

2.2.1 预处理

识别电力设备监测录音中的环境噪声时,对声音的预处理主要包括分帧和加窗2个步骤。进行信号分帧时,为了使相邻帧具有连续性,可参照图2在两帧之间设置部分重叠区域,重叠区域长度一般为帧长的1/3~1/2。图2中,分帧帧长为T,重叠区域为T/2,后一帧相对前一帧的位移量称为帧移。







2.2.2 特征提取

对分帧后的信号进行分析,提取每帧的时、频域特征。时域上,本文结合文献[13-14]的研究以及多次实验选取了均方根X_{RMS}、均方差X_{MES}、峭度K、偏度特征S这4个指标。其中,均方根反映信号的能量大小;均方差反映信号的离散程度;峭度反映信号

的冲击成分;偏度反映信号的数据分布对称测度。 各指标定义式见附录A表A1,表中 \bar{X} 为离散振动信 号 $\{x_i\}(i=1,2,\dots,N)$ 的平均值, $\bar{X}=\frac{1}{N}\sum_{i=1}^{N}|x_i|,N$ 为采 样点数。

本文选择小波包分解后的能量特征作为频域特征,用于表征不同频段上的频率成分的能量分布^[15]。 小波包分解是一种频段划分方法,能将信号按任意 时频分辨率分解到不同频段,具有精确细分的特点 和较强的时频局部化能力,能很好地满足信号特征 提取的要求。

对于已知声音信号*s*(*t*),经过*i*层小波包分解后将在第*i*层获得2ⁱ个信号的子频段,此时*s*(*t*)可表示为:

$$s(t) = \sum_{i=0}^{2^{i}-1} f_{i,j}(t_j) = f_{i,0}(t_0) + f_{i,1}(t_1) + \dots + f_{i,2^{i}-1}(t_{2^{i}-1}) \quad (1)$$

式中: $i=0, 1, 2, ..., 2^{i}-1; f_{i,j}(t_{j})$ 为进行i层小波包分 解后节点j在第i层的重构信号。若信号的最小频 率为 f_{\min} ,最大频率为 f_{\max} ,则经过i层分解后第i层的 每个子频段频率宽度为 $(f_{\max} - f_{\min})/2^{i}$ 。

由于含噪声与不含噪声的声音片段在频率分布 上具有差异性,可提取不同频段上的特征如能量进 行区分。上述信号*s*(*t*)经过小波包分解后,每个子 频段能量*E*_{ii}可表示为:

$$E_{i,j} = \int_{T} \left| f_{i,j}(t_j) \right|^2 dt = \sum_{k=1}^{N} \left| x_{(j,k)} \right|^2$$
(2)

式中: $x_{(j,k)}$ 为重构信号 $f_{i,j}(t_j)$ 的离散点幅值,j=0,1,2,…, 2^i-1 。s(t)的总频段能量E和各节点所占频段 能量与总频段能量之比 $P_{i,j}$ 分别如式(3)、(4)所示。

$$E = \sum_{j=0}^{2^{-1}} E_{i,j}$$
(3)

$$P_{i,i} = E_{i,i} / E \times 100 \%$$
 (4)

2.3 基于 MeanShift 聚类的相似性分析方法

聚类算法是以相似性为基础,将样本集合按照 特征属性划分为不同类簇,它认为属于同一类簇的 样本具有较大的相似性,属于不同类簇的样本间具 有较大的差异性。

MeanShift 算法又称为均值漂移算法,最早由 Fukunage 提出,经Yizong Cheng扩充完善后,被广泛 应用于图像分割^[16]、目标追踪^[17]、聚类等领域。与 传统的K均值聚类算法相比,其优势在于无需确定 分布的形状和类簇数量^[18]。MeanShift 算法的思想 是假设数据集 $\{x_i | i=1, 2, ..., n\}$ 服从概率密度函数 f(x),并将数据集稠密区域认为是局部极大值或一 种模式,对空间中的每个点进行峰值搜索直至收敛, 将收敛到同一个点的作为一类。设 $x_i(i=1, 2, ..., n)$ 为d维空间 \mathbf{R}^d 中的样本点,其概率密度函数为f(x), 利用核函数 $K(\mathbf{x})$ 进行多元核密度估计,得到 $f(\mathbf{x})$ 的 核函数估计 $\hat{f}(\mathbf{x})$ 为:

$$\hat{f}(\boldsymbol{x}) = \frac{1}{nh^{d}} \sum_{i=1}^{n} K\left(\frac{\boldsymbol{x} - \boldsymbol{x}_{i}}{h}\right)$$
(5)

式中:h为带宽参数;核函数K(x)满足式(6)。

$$K(\mathbf{x}) = c_{k,d} k(\|\mathbf{x}\|^2)$$
(6)

式中: $c_{k,d}$ 为保证 $K(\mathbf{x})$ 积分为1的标准化常数。

另外定义向量函数k(x)的负导函数g(x),即 g(x)=-k'(x),其对应的核函数G(x)为:

$$G(\mathbf{x}) = c_{g,d} g(\|\mathbf{x}\|^2)$$
(7)

则 $f(\mathbf{x})$ 的梯度 $\nabla f(\mathbf{x})$ 的估计 $\hat{\nabla} f(\mathbf{x})$ 为:

$$\hat{\nabla}f(\mathbf{x}) = \nabla \hat{f}(\mathbf{x}) = \frac{2c_{k,d}}{nh^{d+2}} \sum_{i=1}^{n} (\mathbf{x} - \mathbf{x}_i) g\left(\left\|\frac{\mathbf{x} - \mathbf{x}_i}{h}\right\|^2\right) = \frac{2c_{k,d}}{nh^{d+2}} \left[\sum_{i=1}^{n} g\left(\left\|\frac{\mathbf{x} - \mathbf{x}_i}{h}\right\|^2\right)\right] (\mathbf{m}_h(\mathbf{x}) - \mathbf{x}) \quad (8)$$

$$\boldsymbol{m}_{h}(\boldsymbol{x}) = \sum_{i=1}^{n} \boldsymbol{x}_{i} g\left(\left\| \frac{\boldsymbol{x} - \boldsymbol{x}_{i}}{h} \right\|^{2} \right) / \sum_{i=1}^{n} g\left(\left\| \frac{\boldsymbol{x} - \boldsymbol{x}_{i}}{h} \right\|^{2} \right) \quad (9)$$

式中: $m_h(x)$ 为均值漂移向量。

均值漂移向量 m_h 始终指向密度增加最大的方向,如图3所示。图中, x_i^0 为迭代中的质心; x_i^1 为矢量移动后的质心; x_i^n 为最终质心。MeanShift算法具体步骤如下:

1)给定初始点 $x \cdot g(x)$ 的核函数G(x)、容许误差 ε ,计算均值漂移向量 $m_k(x)$;

2)将 $m_h(x)$ 赋值给x;

3)重复步骤1)、2),直至**||m**_h(**x**)−**x** ||≤ε 时结束 循环。



图 3 MeanShift聚类示意图



将一段声音信号经过分帧后得到的n帧信号 记作{ y_i |i=1, 2, ..., n},将其时域特征向量记作 $S_{i,1} - S_{i,4}$,将频域特征向量记作 $P_{i,1} - P_{i,128}$,则第i帧声音信号的特征向量 M_i 、单个声音样本总的特征 矩阵M分别如式(10)、(11)所示。

$$\boldsymbol{M}_{i} = [\boldsymbol{S}_{i}]_{1} \cdots \boldsymbol{S}_{i} \boldsymbol{A} \boldsymbol{P}_{i}]_{1} \cdots \boldsymbol{P}_{i}]_{128}$$
(10)

$$\boldsymbol{M} = \begin{bmatrix} \boldsymbol{M}_{1} \\ \vdots \\ \boldsymbol{M}_{n} \end{bmatrix} = \begin{bmatrix} \boldsymbol{S}_{1,1} & \cdots & \boldsymbol{S}_{1,4} & \boldsymbol{P}_{1,1} & \cdots & \boldsymbol{P}_{1,128} \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{S}_{n,1} & \cdots & \boldsymbol{S}_{n,4} & \boldsymbol{P}_{n,1} & \cdots & \boldsymbol{P}_{n,128} \end{bmatrix}$$
(11)

对样本的特征矩阵进行聚类分析,若能聚成一 类,则说明各帧信号具有较高的相似性,判断为可用 于诊断设备运行状态的有效样本;若形成多个类簇, 则说明样本不具有自相似性,判断为含噪声样本。

3 现场实验

3.1 数据采集及处理

采用 TEAC 公司的 TASCAMDR-05 录音机在部 署有一台配电变压器的箱式变压器内持续采集2个 月的可听声音,现场布置见附录B图B1。声音传感 器的测量误差在3dB以内,采样频率为48kHz。录 制了变压器正常运行状态和底座螺栓松动状态下, 大量含不同类型环境噪声的声音样本,录制时的环 境噪声以车辆路过、语音、雨声、风声等为主。由于 录音是现场采集,将人工挑选的高信噪比声音样本 作为不含环境噪声的正常运行声音样本。

将录音数据分割成长度为60s的子音频样本, 以其中一个子样本为例进行说明。首先,按分帧帧 长10s对子样本进行分帧,如图4所示;然后提取每 帧的时、频域特征。将小波包分解层数设置为7, 小波基选择db10,利用小波包分解提取不同频段的 能量特征,得到128个频段能量,每频段带宽约为 156 Hz。取前3帧的前10个频段的能量分布进行分 析,如图5所示,图中为含噪声帧(第1帧)与不含噪



图4 含环境噪声样本的分帧







Fig.5 Frequency energy ratio distribution of frames with and without noise

声帧(第2、3帧)的前10个频段的能量分布。由图可见,含噪声与不含噪声帧在不同频段上的能量分布 有明显差异,不含噪声的正常运行帧之间的频段能 量差异较小,因此小波包频段能量可作为区分噪声 的特征向量。

3.2 噪声剔除

利用 MeanShift 算法对每个样本的特征矩阵进 行聚类分析,当聚类结果中只有1个类簇时,为不含 噪声的有效样本,否则为含噪声样本。该算法需设 置带宽参数*h*,基于大量历史样本实验测试本文设 置*h*=0.02。

由于高维数据无法直接显示,本文采用多维标 度法将高维数据降维至二维^[19],部分样本聚类结果 如附录B图B2所示。图B2(a)、(b)中,样本各帧聚 类形成1个类簇,各帧录音具有较高的相似性,可见 图 B2(a)、(b)所示的样本为有效样本。图 B2(d)-(f)为含语音、雨声、风声噪声的变压器正常运行声 音样本,由图可见,样本各帧聚类后形成多个类簇, 可判断为含噪声样本。图B2(c)为含鸟叫噪声的正 常运行样本,由图可见,样本各帧同样也被聚为一 类,出现了含噪声样本各帧聚类形成单个类簇的情 况,相比其他含噪声样本,该样本中的噪声具有持续 时间较短、能量小的特点。经分析比较,笔者认为分 帧的帧长设置是影响判断的关键要素。该算例中, 分帧长度为10s,长度偏大,使得每帧信号的时、频 域统计特征相对平滑,减小了含噪声帧与不含噪声 帧的特征差异,使得具有较高相似性的样本被聚为 一类。但帧长设置过小也存在放大局部细节、削弱 样本之间相似性的问题。

3.3 分帧长度对聚类结果的影响

为分析分帧长度对样本聚类结果的影响,在不同的分帧帧长下对附录B图B3所示的测试样本进行聚类,分析聚类类簇数目对分帧长度变化的影响,结果如图6所示。图中,样本1-4分别对应图B3中含鸟叫的正常运行声音样本、不含噪声的正常运行声音样本、底座螺栓松动声音样本、绕组松动声音样本。由图可见:分帧长度T≥6.4 s时,含噪声样本



Fig.6 Change of number of clusters with different framing lengths

聚类为1个类簇,而在 $T \leq 3.2$ s时无法聚类为1个类 簇,因此,为识别该含噪声的样本,分帧长度不宜超 过6.4s;不含噪声的声音样本在T > 0.1s时均可聚类 为1个类簇,而当T < 0.1s时出现多个类簇;底座螺 栓松动声音样本在T > 0.4s时能聚类为1个类簇,当 T < 0.4s时开始出现多个类簇;绕组松动声音样本在 $T \geq 0.8$ s时只能聚类为1个类簇。为了保证后3种有 效样本不被聚类为多个类簇,T应大于0.8s。

3.4 分帧长度范围的确定

3.3节的分析表明分帧长度会影响聚类结果,并 给出了能同时识别含噪声样本和有效样本的大致分 帧长度的取值范围,但该取值范围仅利用少量样本 得到,不能代表实际情况。为保证有效样本(不含噪 声的正常、故障声音样本)只能聚类为1个类簇,同 时含噪声样本不能聚类为1个类簇,本节利用大量 不同类型的声音样本进行训练,得到各种不同类型 样本的优选分帧长度取值范围,并取交集得到总体 的优选取值范围。

在设置有效样本的优选分帧长度时,可定义式 (12)所示的有效样本识别率*P*_{T1},*P*_{T1}越大,类簇数目 为1的有效样本数量越多,则有效样本识别率越高。

$$P_{\rm T1} = \frac{N_{\rm 1}}{N_{\rm e}} \times 100 \,\% \tag{12}$$

式中:N₁为类簇数目为1的有效样本数量;N_e为有效 样本数量。

在设置含噪声样本的优选分帧长度时,可定义 式(13)所示的噪声样本识别率 P_{T2} , P_{T2} 越大,表明越 多的噪声样本难以形成1个类簇,噪声样本识别率 越高。

$$P_{\rm T2} = \frac{N_0}{N_{\rm n}} \times 100 \,\% \tag{13}$$

式中: N_0 为类簇数目不为1的样本数量; N_n 为含噪声的样本数量。

当 P_{T1} =100%时,优选的有效样本分帧长度 T_1 存在一取值范围,当 P_{T2} =100%时,优选的含噪声样本分帧长度 T_2 存在一取值范围,理论上这2个取值范围之间存在交集,该交集即为总体的最优分帧长度取值范围。

设置初始分帧长度为 0.1 s,以步长为 0.1 s进行 迭代, T最大取值为 15 s,训练样本的信息及其优选 分帧长度范围如表 2 所示。由表可见,为有效识别 有效样本,需要 $T_1 \ge 0.9$ s,为有效识别含噪声样本, 需要 $T_2 \le 3.5$ s,取交集得到总体的最优分帧长度取 值范围为[0.9,3.5] s。

为测试优选的分帧长度取值范围能否适应其他样本,利用人工挑选的测试集(包括40组有效样本、40组含各种噪声的样本),得到指标P_{TI}、P_{T2}的测试结果如图7所示。由图可见:当分帧长度设

Table 2 Ti	aining result	S		
样本类型	训练样本数	优选帧长范围		
正常运行,不含噪声	20	$T_1 \! \ge \! 0.2 \text{ s}$		
故障	20	$T_1 \ge 0.9 \text{ s}$		
正常运行,含语音噪声	20	$T_2 \! \leqslant \! 3.8 \ { m s}$		
正常运行,含车辆路过噪声	20	$T_2 \! \leqslant \! 5.2 \; \mathrm{s}$		
正常运行,含鸟叫、蝉鸣噪声	20	$T_2 \leq 3.5 \text{ s}$		
正常运行,含雨声噪声	20	$T_2 \leq 10 \text{ s}$		
正常运行,含刮风噪声	20	$T_2\!\leqslant\!15~{\rm s}$		
⁸ 100 − − − − − − − − − − − − − − − − − −	• • • • • • • •	100 %		
2 + 95 = 90 = 95 = 95 = 90 = 1.0 = 1.5 = 2	% .0 2.5	3.0 3.5		
$T \neq s$				

表2 训练结果



Fig.7 Test results of P_{T1} and P_{T2}

 $--P_{T1}, --P_{T2}$

置在[0.9, 3.5]s范围内时, P_{T1} 、 P_{T2} 始终大于95%;在 [1.5, 3]s范围内, P_{T1} 、 P_{T2} 同时达到100%。由此可见,所确定的优选分帧长度范围能够满足现场需求, 可保留大部分有效样本,同时识别含噪声样本。

4 结论

本文利用电力设备在正常和故障状态下运行 声音具有自相似性而环境噪声不平稳、无自相似性 的差异性特征,提出了一种基于时、频域自相似性分 析的电力设备环境噪声识别与剔除方法,主要结论 如下:

1)对比设备正常、故障运行声音和典型环境噪 声的时、频域特征,分析发现两者在稳定性和自相似 性上具有明显差异;

2)提出基于时、频域自相似性的配电设备录音 监测数据中环境噪声的识别与剔除方法,将录音数 据分割为分钟级片段后再进行分帧,对各帧的时、频 域特征指标进行聚类分析识别含非平稳环境噪声的 录音数据;

3)基于现场实测数据的分析表明,所提方法的 环境噪声剔除准确度主要受分帧长度影响,将分帧 长度设置在[1.5,3]s范围内时可准确识别并剔除所 有含环境噪声的录音数据。

本文所提方法为后续利用声音信号识别配电设备运行状态奠定了重要基础。需要指出的是,所提 方法是以环境噪声不具有稳定性和自相似性为前提的,在本文的测试环境下能有效识别和剔除环境噪声。未来一方面需要进行更多运行场景下的测试分析,探寻可能出现的具有自相似性的环境噪声,另一 方面也可以从设备运行状态诊断分析的算法设计上 入手,利用夹杂稳定的具有自相似性的环境噪声的 设备运行录音识别配电设备运行状态。

附录见本刊网络版(http://www.epae.cn)。

参考文献:

- [1]潘亮亮,赵书涛,李宝树.基于声波信号分析的电气设备故障 诊断新方法[J].电力自动化设备,2009,29(8):87-90.
 PAN Liangliang, ZHAO Shutao, LI Baoshu. Electrical equipment fault diagnosis based on acoustic wave signal analysis
 [J]. Electric Power Automation Equipment,2009,29(8):87-90.
- [2] 李建鹏,赵书涛,夏燕青.基于双谱和希尔伯特-黄变换的断路器故障诊断方法[J].电力自动化设备,2013,33(2):115-119,125.

LI Jianpeng,ZHAO Shutao,XIA Yanqing. Fault diagnosis based on bispectrum and Hilbert-Huang transform for circuit breaker [J]. Electric Power Automation Equipment, 2013, 33(2):115-119,125.

- [3] NAKANO S, TSUBAKI T, HIRONAKA S. Applying a voice recognition system for SF₆ gas insulated switchgear's inspection/maintenance services[J]. IEEE Transactions on Power Delivery, 2001, 16(4):534-538.
- [4] 马春雷,谢荣斌,赵莉华,等.变压器可听声信号特征分析[J]. 电力大数据,2018,21(2):18-26.
 MA Chunlei, XIE Rongbin, ZHAO Lihua, et al. Characteristic analysis of transformer audible acoustic signals[J]. Power Systems and Big Data,2018,21(2):18-26.
- [5]范竞敏,曹建,丁家峰.电力变压器绕组状态实时监测算法
 [J].电力自动化设备,2010,30(3):81-85.
 FAN Jingmin, CAO Jian, DING Jiafeng. Real-time monitoring algorithm of power transformer windings status[J]. Electric Power Automation Equipment,2010,30(3):81-85.
- [6] 郭俊, 汲胜昌, 沈琪, 等. 盲源分离技术在振动法检测变压器故障中的应用[J]. 电工技术学报, 2012, 27(10):68-78.
 GUO Jun, JI Shengchang, SHEN Qi, et al. Blind source separation technology for the detection of transformer fault based on vibration method[J]. Transactions of China Electrotechnical Society, 2012, 27(10):68-78.
- [7] 刘字舜,程登峰,夏令志,等. 基于单通道盲源分离算法的局部 放电特高频信号去噪方法[J]. 电工技术学报,2018,33(23): 5625-5636.
 LIU Yushun, CHENG Dengfeng, XIA Lingzhi, et al. Partial discharge ultra-high frequency signal de-noising method based on single-channel blind source separation algorithm[J]. Transactions of China Electrotechnical Society,2018,33(23):5625-5636.
- [8] ZOU L, GUO Y K, LIU H, et al. A method of abnormal states detection based on adaptive extraction of transformer vibroacoustic signals[J]. Energies, 2017, 10(12): 2076.
- [9]周东旭,王丰华,党晓婧,等.基于稀疏表示理论的特高压交流 变压器声信号盲分离研究[J]. 电网技术,2020,44(8):3139-3148.
 ZHOU Dongxu, WANG Fenghua, DANG Xiaojing, et al. Blind separation of UHV power transformer acoustic signal preprocessing based on sparse representation theory[J]. Power System Technology,2020,44(8):3139-3148.
- [10] 焦彦军,胡春.基于改进EEMD方法的数字滤波器[J].电力自动化设备,2011,31(11):64-68.
 JIAO Yanjun,HU Chun. Digital filter based on improved EEMD method[J]. Electric Power Automation Equipment,2011, 31(11):64-68.
- [11] 姜万录,王浩楠,朱勇,等.变分模态分解消噪与核模糊C均值 聚类相结合的滚动轴承故障识别方法[J].中国机械工程,



2017,28(10):1215-1220,1226.

JIANG Wanlu, WANG Haonan, ZHU Yong, et al. Integrated VMD denoising and KFCM clustering fault identification method of rolling bearings[J]. China Mechanical Engineering, 2017,28(10):1215-1220,1226.

- [12] 汲胜昌,张凡,师愉航,等. 基于振动信号的电力变压器机械状态诊断方法研究综述[J]. 高电压技术,2020,46(1):257-272.
 JI Shengchang,ZHANG Fan,SHI Yuhang,et al. Review on vibration-based mechanical condition monitoring in power transformers[J]. High Voltage Engineering,2020,46(1):257-272.
- [13] 张恒,赵荣珍.故障特征选择与特征信息融合的加权 KPCA方法研究[J].振动与冲击,2014,33(9):89-93,121.
 ZHANG Heng,ZHAO Rongzhen. Weighted KPCA based on fault feature selection and feature information fusion[J]. Journal of Vibration and Shock,2014,33(9):89-93,121.
- [14] 杨毅,刘石,张楚,等. 基于振动分布特征的电力变压器绕组故 障诊断[J]. 振动与冲击,2020,39(1):199-208.
 YANG Yi,LIU Shi,ZHANG Chu, et al. Winding fault diagnosis of power transformer based on vibration distribution features[J]. Journal of Vibration and Shock,2020,39(1):199-208.
- [15] 白辉,许志红. 基于小波包变换和高阶累积量的电弧故障识别 方法[J]. 电力自动化设备,2020,40(11):195-202,224.
 BAI Hui, XU Zhihong. Arc fault identification method based on wavelet packet transform and high-order cumulant[J]. Electric Power Automation Equipment,2020,40(11):195-202,224.
- [16] GUO Y H, ŞENGÜR A, AKBULUT Y, et al. An effective color image segmentation approach using neutrosophic adaptive mean

shift clustering[J]. Measurement, 2018, 119:28-40.

- [17] 曹帆之,朱述龙,朱宝山,等.均值漂移与卡尔曼滤波相结合的 遥感影像道路中心线追踪算法[J]. 测绘学报,2016,45(2): 205-212,223.
 CAO Fanzhi,ZHU Shulong,ZHU Baoshan, et al. Tracking road centerlines from remotely sensed imagery using mean shift and Kalman filtering[J]. Acta Geodaetica et Cartographica Sinica,2016,45(2):205-212,223.
- [18] YAMASAKI R, TANAKA T. Properties of mean shift[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020,42(9):2273-2286.
- [19] 郝智勇,贺明科,谭文堂,等. 基于多维标度法的专利文本可视 化聚类研究[J]. 计算机应用研究,2010,27(12):4608-4611.
 HAO Zhiyong, HE Mingke, TAN Wentang, et al. Research on MDS-based visual clustering in patent texts[J]. Application Research of Computers,2010,27(12):4608-4611.

作者简介:



刘 元(1995—),男,安徽宣城人,硕 士研究生,主要研究方向为电力设备故障诊 断(**E-mail**:302620970@qq.com);

苏 盛(1975—),男,湖南长沙人,教 授,博士,通信作者,主要研究方向为电力 网络安全防护和大数据技术应用(E-mail: eessheng@163.com)。

刘元

(编辑 任思思)

Environmental noise recognition method for distribution transformer based on time domain and frequency domain self-similarity

LIU Yuan¹, SU Sheng¹, LIU Zhengyi², XIA Yunfeng³, LIU Guanke³, LI Bin¹

(1. Disaster Prevention and Reduction Center of Renewable Energy System, Changsha University of

Science and Technology, Changsha 410004, China;

2. State Grid Changde Power Supply Company, Changde 415000, China;

3. Dongguan Power Supply Bureau of Guangdong Power Grid Co., Ltd., Dongguan 523000, China)

Abstract: The distribution transformer is located in an open noisy environment, and environmental noise has a prominent impact on the online monitoring of sound-based equipment operating status. How to identify and eliminate recordings containing noise is a bottleneck problem to be solved urgently in the promotion of sound-based online monitoring of the status of distribution transformer. The analysis points out the difference between the normal and abnormal operating sound of power distribution equipment and various kinds of environmental noises in time and frequency domains. The method for identifying environmental noise fragments in the recording data of power distribution equipment based on self-similarity in time and frequency domains is proposed. After the minute-level recording data is divided into frames, the time domain and frequency domain characteristic indexes of each frame are extracted. Whether the recording fragments have self-similarity is judged by clustering, and the recording data containing environmental noise are identified and eliminated. The experimental results show that the proposed method can effectively identify and eliminate the recording segments containing environmental noise, thereby recording data with lower environmental noise content can be obtained, and providing strong support for the subsequent development of sound signalbased status monitoring of distribution transformer.

Key words: sound signal; status monitoring; distribution transformer; noise rejection





fault operation of transformer

附录 A



图 A2 环境噪音的时域、时频域图像 Fig.A2 Time domain and time-frequency domain images of environmental noise

Table A1 Time domain characteristics index				
名称	表达式	作用		
均方根值(RMS)	$X_{\rm RMS} = \sqrt{\sum_{i=1}^{N} x_i^2 / N}$	反映信号的能量 大小		
均方差(MES)	$X_{\text{MES}} = \frac{1}{N} \sum_{i=1}^{N} (x_i - \frac{1}{N} \sum_{i=1}^{N} x_i)^2$	反映信号的离散 程度		
峭度(K)	$K = \sum_{i=1}^{N} x_i^4 / (NX_{\rm rms}^4)$	反映信号的冲击 成分		
偏度(S)	$S = \frac{\frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{X})^3}{\left[\frac{1}{N-1} \sum_{i=1}^{N} (x_i - \bar{X})^2\right]^{3/2}}$	反映信号的数据 分布对称测度		

表 A1 时域特征指标



图 B1 现场传感器布置 Fig.B1 On-site sensor placement 注:箱式变压器包括低压室、高压室以及变压器室,为了保证人员安全,录音设备放置在箱式变压器的 低压室柜门上,录音时柜门处于关闭状态。

附录 B



图 B2 部分样本聚类结果 Fig.B2 Partial sample clustering results



