

基于声纹的高泛化性风机叶片异常检测方法研究



邹宜金, 连应华, 黄新宇, 刘培*, 王颖蕊

(福建华电可门发电有限公司 福州 350512)

【摘要】对风力发电机组叶片异常检测进行研究,发现依靠单台风力发电机 3 个叶片声纹之间的参考和对比即可诊断该台风力发电机叶片是否故障。为此,该文提出基于聚类和中值收敛的周期性音频切割方法,对声纹进行有效的切割,减少了计算量,为后续异常检测提供了基础。采用风力发电机组 3 个叶片之间的稳态差异法对异常进行检测,绕开了待检物,信道等变化所带来的算法迁移失效问题,具有良好的泛化性。为风力发电机组叶片检测提供了一种有效的技术手段。

关键词 异常检测; 故障诊断; 声纹; 风机叶片

中图分类号 TM315 文献标志码 A doi:10.12178/1001-0548.2021080

High Generalization in Anomaly Detection of Wind Turbine Generator Based on Voiceprint

ZOU Yijin, LIAN Yinghua, HUANG Xinyu, LIU Pei*, and WANG Yingrui

(Fujian Huadian Kemen Power Generation Co., Ltd. Fuzhou 350512)

Abstract The major research work of the paper is the anomaly diagnosis of wind turbine generator based on voiceprint. The research finds that the reference and comparison between the three blades of a single wind turbine can diagnose whether the wind turbine is faulty. On this base, the paper proposes a periodic audio cutting method based on clustering and median convergence, which effectively cuts the voiceprint, reduces the amount of calculation, and provides a basis for subsequent anomaly detection. The steady-state difference method between three blades of wind turbine is used to detect anomalies, which avoids the migration failure of algorithm caused by the changes of channels and objects to be checked. The paper provides an effective technical means for fan blade inspection.

Key words anomaly detection; fault diagnosis; voice print; wind turbine blades

当前,风力发电正以极快的速度发展,但由于风电机组大多安装在高山、荒野、海滩、海岛等相对恶劣的环境中,容易出现故障,维修起来耗费大量人力物力。

在风力发电机组中,叶片作为整个风电机组最前端的捕风机构,造价约占整机的 20% 左右。从叶片本身看其复合材料结构决定了其运行期间自身状态的复杂性;从运行环境看,风机叶片是整个机组暴露在外部环境中长期运行的最大部件,风沙、盐雾、紫外线、雷电等都会对叶片造成一定的损伤;从功能发挥上看,作为捕风机构的叶片在运行过程中要承受各种风况施加的多变震、弯曲、扭转的交替应力变化^[1]。在以上因素的影响下,随着运行时间的推移将导致叶片发生破损、开裂乃至断裂

等质量问题,这些问题无论从影响程度还是维修成本上都居高不下。由于叶片的维修和更换,相对于其他设备更为复杂,耗时更长,所以叶片故障引起的停机时长,在机组总故障中的比例高达 34%^[2],因此对风电机组叶片工作状况进行异常检测非常必要。有效的风机叶片异常检测能提高生产率,减少运维成本支出,增强作业安全性。

1 相关工作

在风机故障检测领域,已发展基于声发射^[3-4]、振动分析^[5-7]、光纤光栅^[8]、红外热成像^[9]等多种原理的检测技术。考虑到实际情况,能符合检测效率实时、非接触式且便于安装要求的有 SCADA 数据分析和声学检测两类方法。

收稿日期: 2021-03-22; 修回日期: 2021-05-26

作者简介: 邹宜金(1977-),男,高级工程师,主要从事电力电子与电力传动方面的研究。

*通信作者: 刘培, E-mail: afrapei@163.com

SCADA 数据分析^[10-11]是对 SCADA 传感器数据进行分析,采用聚类或者趋势预测对风机叶片可能发生的异常进行分类判定。此类方法传感器数据汇总属性数据维度过高,采样频率较低,相关属性分析模糊,工况对数据影响大,模型准确度受限。

声学检测方法主要实现手段有:

1) 传统数值统计分析,对波形进行时域频域分析,依赖其均方根值、峰值水平、峭度、斜度、波峰因数和其他数值特点进行风机叶片是否异常的判定。该方法是比较早期的研究方案,在实际运用中难以得到满意的结果。

2) 非线性和非稳态的分析^[12],采用小波变换或者希尔伯特-黄变换得到故障特征频率变化的特征。不过,输入信号复杂度高,采用小波变换也没有从根本上解决任意变频的问题,没有切合风机是一个线性系统的非定常响应问题。

3) 普通机器学习采用分类方法、回归方法、神经网络、支持向量机、决策树、随机森林等^[13]。这需要大量故障风机叶片的信号样本,但实际中风机叶片的故障类型、受损程度与部位不同都将导致叶片故障数据收集不全面,从而影响所建模型的准确度。此类方法,如果换个场景和风机,模型的准确度会存在问题。

实际应用中风场位置不同、风机类型各异、故障类型众多、背景环境多变等因素复杂交错。

上述基于规则和特征的手段都面临非线性表征能力弱,泛化能力不强,识别率不高的问题,所以无法满足实际工程需要。基于深度学习的方法,受限于数据采集的规模难以实现,和实现过程中易过拟合及调参难的通病,难以泛化,故障诊断的可迁移性存在问题。同时,信号复杂度高、数据不均衡、数据收集困难也是解决风机故障监测必须直面的问题。

2 方法的设计

2.1 模型思考和设计

通过项目研究,发现风机实际运行中,由于叶片旋转具有周期性,采集叶片声音传感器位置在叶轮迎风侧正下方,故叶片扫风的声音信号也具有周期性。所以,可以利用受损叶片和正常叶片的周期性和差异性特点进行高效准确的异常检测。

区别于目前基于大量数据样本的智能诊断算法,不需要采集大量的数据信息,不需要建立诊断识别的模型,仅依靠单台风机 3 个叶片之间的参考和对比就可以诊断该台风机是否故障,操作流程简单,诊断精度也较高;从算法角度,减少了计算量,提出基于聚类和 中值收敛的周期性音频切割方法,对声纹进行有效的切割,为增强声纹变化的鲁棒性问题提供基础。

本文具体方法的设计如图 1 所示。

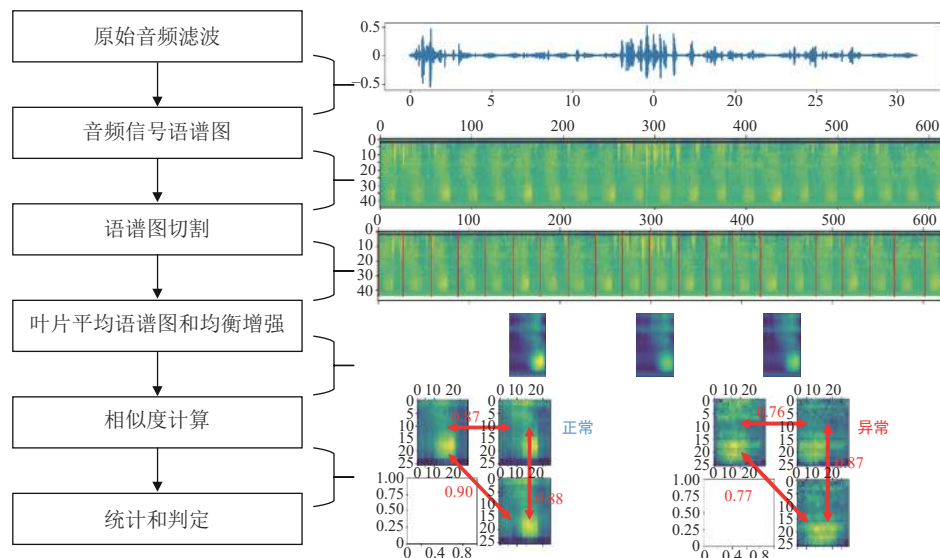


图 1 本方法实现示意图

2.1.1 原始音频预处理

声音采集器工作在复杂的室外,采集到的音频

信号一般包含大量噪声,如鸟叫、风声、人声和其他风机造成的噪声等。目前,风机的启动条件一般

为平均风速不小于 3.5 m/s, 声音传感器采集的信号必然包含风噪, 且相比其他背景噪声, 风噪影响较大。风噪的频谱能量集中在 350 Hz 以下的低频区域, 需要滤波器将低频的风噪滤除。

2.1.2 计算音频信号语谱图

声音信号的特征提取是核心步骤之一, 提取到有效、可靠的特征能提高结果的准确性和有效性, 降低处理的复杂度。经过实验语谱图能很好的表征风机叶片的音频特征。

2.1.3 语谱图切割

对语谱图峰图和峰图之间的间隔进行识别, 实现精准的音频分割, 得到分割的坐标。值得一提的是, 虽然采用了声纹分割聚类^[14]这个名字, 但是其含义不是传统语音中的识别说话人数量和归属, 是识别不同风机叶片扫风音频和扫风间隔音频。

不采用实时转速情况下, 叶片周期的声纹切割算法有以下两种:

1) 波峰波谷峰值计算。正常纯净扫风音频, 可以使用该办法, 但是在实际风场中, 经常有无规律间隔叠加能量的干扰, 且风机 3 个叶片的功率强度轴也可能不同, 导致相邻叶片的波峰波谷有较大的震荡区间, 对峰值检测的精度和准度带来挑战。少识别的波峰, 会导致错周期切割, 直接影响后续的叠加算法。

2) 深度学习。采用一维卷积神经网络或者其他模型, 标注峰值区域进行特征学习, 需要大量采集数据, 还需要手工标注不少数据, 模型也需要提前训练, 工程实现较为冗杂。

本文采用聚类方法对中高维语谱图矩阵进行聚类, 得到扫风声和安静声的标签序列, 接着用中值收敛调整修正该标签序列, 能精确得到安静音频到扫风音频阶跃的坐标点。这种设计不仅能无参数动态分割, 还有较强的鲁棒性和泛化性。

具体实现如下:

1) 采用 K_means 算法将语谱图矩阵聚成两类, 分割出扫风音频和扫风间隔期音频。

2) 对语谱图的中高维度进行截取, 将降维后的矩阵输入 K_means 两类聚类算法。

3) 得到聚类的 0,1 标签序列。K_means 聚类后的效果不错, 得到扫风声和安静声的标签序列, 能大体划分出扫风周期, 不过由于信号复杂度高, K_means 聚类对异常值敏感, 还是存在毛刺和缺失。需要对标签序列的毛刺和尖峰进行过滤。

4) 接着进入收敛修正聚类标签操作, 至此, 可以得到分割点序列坐标。

2.1.4 叶片平均语谱图和均衡增强

透明度叠加, 生成风机每个叶片平均语谱图。用分割的坐标将长矩阵切割成 t 个矩阵 $T_i(0 < i \leq t-1)$, 然后按照透明度 p 叠加矩阵值。

$$p = \frac{3}{t} \quad (1)$$

$$T_{col1} = p \sum_{i=1}^{t/3} T_{3i-2} \quad (2)$$

$$T_{col2} = p \sum_{i=1}^{t/3} T_{3i-1} \quad (3)$$

$$T_{col3} = p \sum_{i=1}^{t/3} T_{3i} \quad (4)$$

T_{col1} , T_{col2} , T_{col3} 为叶片音频的平均语谱图画像的矩阵表示, 其叠加效果的热图如图 2 所示。

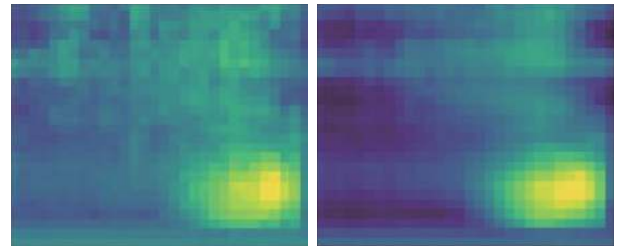


图 2 透明度叠加前后音频语谱图热图对比

风力微弱情形下, 语谱图中有用数据的对比度会比较接近, 采用直方图均衡算法^[15]进行处理, 通过这种方法, 亮度可以更好地在直方图上分布。文献 [16] 提出了灰度直方图均衡化算法。由于彩色 RGB 图片, 有 3 个通道, 不能简单为每个通道应用灰度直方图均衡化算法, 然后组合起来。非线性直方图均衡是非线性过程, 均衡每个通道的做法会得到失真度非常大的结果。均衡涉及图像的强度值, 而不是颜色分量, 因此, 需要将 RGB 格式转为不干扰图像颜色平衡, 均衡亮度值的 YCbCr 格式。在强度平面上执行灰度直方图均衡化操作, 然后再把 YCbCr 格式转回 RGB 格式。

2.1.5 相似度计算

采用图片化的 SSIM (structural similarity) 值来计算 3 个特征矩阵热图图片相似度。结构相似性作为结构相似性理论的实现, 结构相似度指数从图像组成的角度将结构信息定义为独立于亮度、对比度

的,反映场景中物体结构的属性,并将失真建模为亮度、对比度和结构3个不同因素的组合。采用均值作为亮度的估计,标准差作为对比度的估计,协方差作为结构相似程度的度量。

文献[17]提出了SSIM值计算方法。3个特征矩阵热图图片两者之间比较得到3个SSIM值,接着采用这3个结果值进行统计和判断异常。

2.1.6 统计和告警

SSIM值在连续6分钟内,如果相似度都一致,而且大于阈值 k ($0 < k < 1$),那么判断该叶片正常。如果3个SSIM值在连续一段时间内,有一组值稳定偏低而且小于阈值 k ,那么判定该叶片检测为异常,发出检测异常告警。以实现风机叶片的异常检测。

3 实验与结果分析

3.1 输入样本构造集

以山风场和海风场采集的故障和正常风机叶片声信号作为原始数据集,配置为单通道,44100采样率,采样频率区间为0~12 kHz,选择不同特性和背景风机的声信号作为原始数据集,如表1所示。

表1 原始数据集来源

风场	风机编号	数据编号	类型	环境特征
海边风场	5	Data0	Normal	轻微叠加周边其他风机扫风声
	3	Data1	Normal	叠加持续蛙叫虫鸣
	9	Data2	Normal	叠加海风海浪声音
	2	Data3	Normal	叠加雨声
	7	Data4	Normal	叠加围墙造成的轻微回音
	13	Data5	Normal	无明显干扰
	5	Data6	Normal	偶尔塔筒门敲击声
山地风场	16	Data7	Normal	叠加强劲的风噪
	14	Data8	Normal	声音微弱
	11	Data9	Normal	快转速,扫风急促
	19	Data10	Fault	一个叶片扫风异常
	8	Data11	Fault	两个叶片扫风异常

图3是不同音频频谱图表征。

第一组是正常叶片纯净扫风音频,第二组是叠加鸟鸣正常扫风音频,第三组是扫风微弱音频,第四组是叠加敲击声音频,第五组是风噪比较大的音频。

对原始音频进行样本重构得到数据集。重构方法主要有以下3种:

1) 由于风机是包含转速(3~17 r/min)所以,采用拉伸和收缩处理得到不同转速下音频。

2) 采用语音合成,对周期进行中途变化或者叠加背景噪声。

3) 对音频进行软件数据增强处理。

通过样本重构,共得到200组2000 min的音频样本。其中,异常组占40组,异常组占比20%。

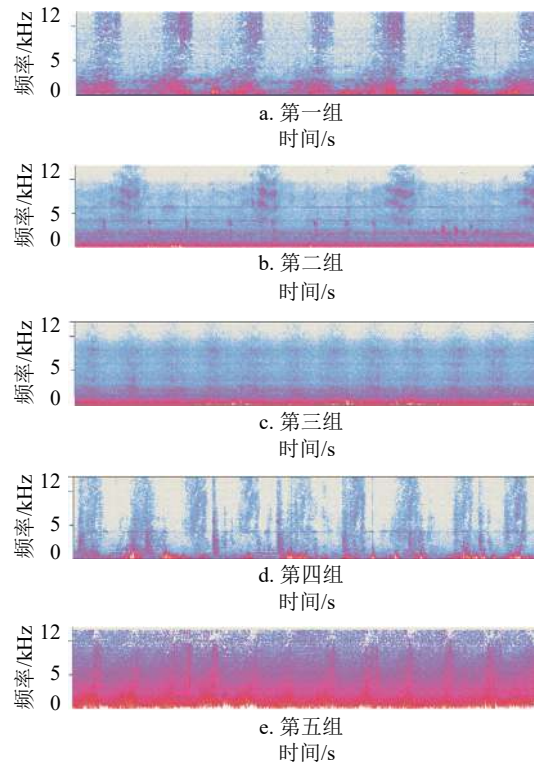


图3 不同风机叶片扫风音频频谱图

3.2 模型结果分析

样本输入模型后,输出结果如下:

异常识别的Recall为0.975, Precision为0.929,具体值如图4所示,该结果与参数选择密切相关。

	实际异常	实际正常
模型识别为异常	39	3
模型识别为正常	157	1

图4 模型识别结果

3.3 参数选择

3.3.1 声纹切割算法的表现

声纹切割算法的前提是聚类成功,由于聚类算法对异常值敏感,在音频受到严重干扰,转速过慢(伴随聚类中心点贴近问题)和转速过快(异常值敏感度加大)的音频切割准确度下降。如图5所示,但在正常转速情况下,切割算法的精度足够。

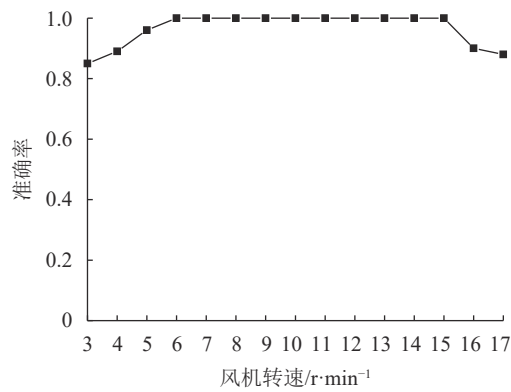


图5 切割算法在叶片不同转速下的精度

3.3.2 相似度方法选择

相似度方法有余弦相似度、均值哈希相似度、差值哈希相似度、感知哈希算法相似度、三直方图相似度、单通道直方图相似度及 SSIM 结构相似度。设一组标签为相似度高的语谱图, 与一组标签为相似度低的语谱图, 输入不同相似度计算方法, 其结果如图 6 所示。差值明显的方法为比较适合的相似度方法。

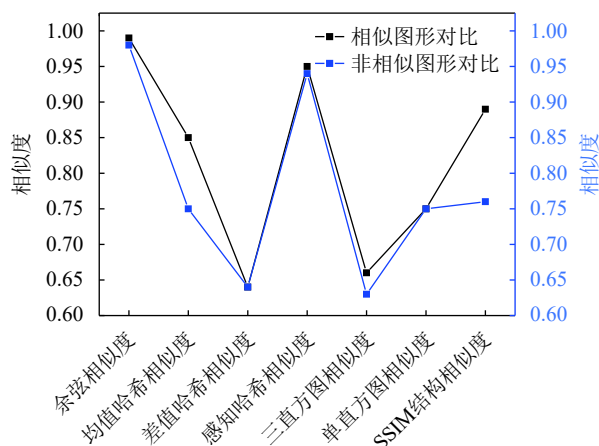


图6 不同相似度算法效果

SSIM 相似度方法在相似度高和相似度低情况下, 有比较大的差值, 能比较好地表征叶片频谱图相似度。

4 结束语

本方法的独特性在于解决了声音异常检测的可行性和可泛化性。

1) 异常检测的可行性: 抛弃了主流的寻找异常声音特性和正常声音特性的方法。重新归纳定义了异常的概念: 某个叶片只要和其他两个叶片声音特性不一致就标记异常。不依赖输入的大数据样本, 不需要提前训练, 易于部署和工程实践, 从声音到图像的转变也使得异常检测有可靠和成熟的量化标准。

2) 异常检测的可泛化性: 能迁移, 具有时间上和空间上的可泛化性。

研究工作的不足之处总结为以下两点:

1) 只对风机叶片的异常检测进行检测和提示, 没有进行具体异常的分类, 比如哨音明显、叶片前缘异常、叶片后缘异常、叶片开裂等。

2) 未进行合适的风机运行状态识别, 未识别导致变桨、偏航、关机过程、停机过程等风机操作控制期间, 识别失效的问题。

参考文献

- [1] 李奎, 印厚飞, 杨智, 等. 风机叶片智能化及应用[J]. 科技创新与应用, 2020, 325(33): 180-181.
LI K, YIN H F, YANG Z, et al. Intelligentization and application of fan blades[J]. Technology Innovation and Application, 2020, 325(33): 180-181.
- [2] 刘昊, 代海涛, 黄宜森, 等. 大型风电机组叶片故障声学诊断技术研究[J]. 风能产业, 2019, 4: 272-276.
LIU H, DAI H T, HUANG Y S, et al. Research on large-scale wind turbine blade failure diagnosis of acoustic technology[J]. Wind Energy Industry, 2019, 4: 272-276.
- [3] TANG J, SOUA S, MARES C, et al. An experimental study of acoustic emission methodology for in service condition monitoring of wind turbine blades[J]. Renewable Energy, 2016, 99: 170-179.
- [4] JOOSSE P A, BLANCH M J, DUTTON A G, et al. Acoustic emission monitoring of small wind turbine blades[J]. Journal of Solar Energy, 2013, 124(4): 401-411.
- [5] JOSHUVA A, SIVAKUMAR S, SATHISHKUMAR R, et al. Fault diagnosis of wind turbine blades using histogram features through nested dichotomy classifiers[J]. International Journal of Recent Technology and Engineering (IJRTE), 2019, 8: 193-201.
- [6] 张彬. 风机叶片损伤特征分析[D]. 上海: 上海电机学院, 2016.
ZHANG B. Damage characteristics analysis of wind turbine blade[D]. Shanghai: Shanghai Dianji University, 2016.
- [7] 张霆. 风力发电机叶片状态监测与故障诊断系统设计和实现[D]. 上海: 上海电机学院, 2016.
ZHANG T. Design and implementation of condition monitoring and bearing fault diagnosis system for wind turbine blades[D]. Shanghai: Shanghai Dianji University, 2016.
- [8] KIM S W, KIM E H, RIM M S, et al. Structural performance tests of down scaled composite wind turbine blade using embedded fiber bragg grating sensors[J]. International Journal of Aeronautical and Space Sciences, 2011, 12: 346-353.
- [9] 黄涛, 顾桂梅. 含裂纹缺陷的红外热图像处理算法研究[J]. 红外技术, 2014, 9: 732-736.
HUANG T, GU G M. Research on processing algorithms of infrared thermal image with crack defect[J]. Infrared Technology, 2014, 9: 732-736.
- [10] 李宇. 基于数据驱动的风机核心部件健康监测方法研究

- 及应用[D]. 成都: 电子科技大学, 2019.
- LI Y. A research and application for wind turbine health condition monitoring based on data-driven[D]. Chengdu: University of Electronic Science and Technology of China, 2019.
- [11] WEI X, VERHAEGEN M, ENGELEN V T. Sensor fault detection and isolation for wind turbines based on subspace identification and KALMAN filter techniques[J]. *International Journal of Adaptive Control & Signal Processing*, 2010, 24(8): 687-707.
- [12] 饶金根, 顾桂梅. 基于谐波小波包和支持向量机的风机叶片损伤识别研究[J]. *玻璃钢/复合材料*, 2014, 4: 37-41.
- RAO J G, GU G M. Wind turbine blade damage identification based on harmonic wavelet packet and support vector machine[J]. *Fiber Reinforced Plastics/Composites*, 2014, 4: 37-41.
- [13] 赵娟. 风机叶片健康监测声学特征提取方法研究[D]. 北京: 北京邮电大学, 2018.
- ZHAO J. Research on acoustical feature extraction in turbine blades health monitoring[D]. Beijing: Beijing University of Posts and Telecommunications, 2018.
- [14] 王泉. 声纹技术: 从核心算法到工程实践[M]. 北京: 电子工业出版社, 2020, 9: 167-168.
- WANG Q. Voice identity techniques: From core algorithms to engineering practice[M]. Beijing: Publishing House of Electronics Industry, 2020, 9: 167-168.
- [15] 余玲飞, 刘强. 基于深度循环网络的声纹识别方法研究及应用[J]. *计算机应用研究*, 2019, 36(1): 153-158.
- YU L F, LIU Q. Research and application of deep recurrent neural networks based voiceprint recognition[J]. *Application Research of Computers*, 2019, 36(1): 153-158.
- [16] HORN B K. *Robot Vision*[M]. Massachusetts: MIT Press, 1986.
- [17] WANG Zhou, BOVIK A C, SHEIKH H R, et al. Image quality assessment: From error visibility to structural similarity[J]. *IEEE Transactions on Image Processing*, 2004, 13(4): 600-612.

编辑 刘飞阳