

DOI:10.19651/j.cnki.emt.2210507

基于精细复合多尺度散布熵的墙体管道敲击探测方法^{*}

李 瑾 行鸿彦 王海峰 吴叶丽 陈 梦

(南京信息工程大学江苏省气象灾害预报预警与评估协同创新中心 南京 210044)

摘要: 为了提高墙体管道敲击探测的准确率,本文采用精细复合多尺度散布熵检测敲击声音信号的频率和幅值的变化,提取信号中的多尺度管道特征;将构建的多维度管道特征矩阵输入到支持向量机中,使用麻雀搜索算法确定支持向量机参数最优值,通过模型训练,完成墙内埋设管道有无的分类,提出了基于精细复合多尺度散布熵的墙体管道敲击探测方法。将此方法与其它信号处理方法进行对比分析,结果证明,本文所提方法探测准确率高达97%,远远高于其他两种方法。

关键词: 管道探测;SSA-SVM;敲击声音;精细复合多尺度散布熵

中图分类号: TP391 **文献标识码:** A **国家标准学科分类代码:** 510.4040

Knock detection method for wall pipes based on refined composite multiscale dispersion entropy

Li Jin Xing Hongyan Wang Haifeng Wu Yeli Chen Meng

(Jiangsu Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters, Nanjing University of Information Science & Technology, Nanjing 210044, China)

Abstract: In order to improve the accuracy of pipe percussion detection in the wall, this paper uses the fine composite multi-scale dispersion entropy to detect the frequency and amplitude changes of the percussion sound signal, and extract the multi-scale pipe features in the signal. The multi-dimensional pipeline feature matrix was input into the support vector machine, and the sparrow search algorithm was used to determine the optimal value of the parameters of the support vector machine. The classification of buried pipelines in the wall was completed through model training, and a knocking detection method of pipelines in the wall based on fine composite multi-scale dispersion entropy was proposed. Comparing this method with other signal processing methods, the results show that the detection accuracy of the proposed method is up to 97%, which is much higher than the other two methods.

Keywords: pipeline detection;SSA-SVM;percussive sound;refined composite multiscale dispersion entropy

0 引 言

管道是城市建设中必不可少的重要部分,当今在对建筑物进行设计时,管道设计大多采用隐线设计,将管道埋藏于墙体内部。在后期不清楚墙体管道走向的情况下对房屋进行二次施工时,为了避免对墙体管道造成损坏,需要对墙体管道进行准确的探测^[1]。而如今常用的管道探测方法主要针对地下管道探测,例如电磁感应法、高密度电法、雷达探测法等,应用于墙体管道探测时存在成本高、操作繁琐、无法探测非金属管道

等问题^[2-3]。声振检测法是激励墙体产生机械振动,通过测量墙体振动声音的特征来判断其中是否含有管道的无损检测方法,操作简单、成本低^[4]。目前已在复合材料检测、农作物检测、木建筑检测等多个领域得到了广泛应用。但到目前为止,声振检测法在管道探测方面的应用都还很少。

早期声振检测法是通过人耳根据敲击声音判断被测物是否有缺陷,十分依赖操作者经验,准确率也不高^[5]。为了减少对操作者经验的依赖,引入传感器采集敲击声音信号并将其转变为电信号进行处理分析。

收稿日期:2022-06-28

^{*} 基金项目:国家重点研发计划资助项目(2021YFE0105500)、国家自然科学基金(62171228)项目资助

管道探测所采集敲击声音信号是非平稳、非线性信号,采用有效的信息分析方法对其进行处理提取管道特征对提升探测准确率至关重要。针对非平稳、非线性信号的分析方法,例如小波分析、经验模式分解、排列熵、样本熵等,都已在敲击探测领域得到了广泛应用。但是,这些方法都属于单一尺度分析方法,普遍存在计算耗时长,应用场景受限,特征提取不明显等缺点。为了克服样本熵与排列熵的部分缺陷,2016年,Rostaghi等^[6]提出了散布熵(dispersion entropy,DE),能够检测噪声带宽以及同时频率和幅值的变化,计算速度快,不易受突变信号影响。但单一尺度DE不能完整提取信号复杂性特征,因此,Hamed等^[7]和Zhou等^[8]在散布熵、多尺度熵和复合多尺度基础上提出了精细复合多尺度散布熵(refined composite multiscale dispersion entropy,RCMDE),RCMDE稳定性更强,在计算误差、特征提取效果等方面都更有优势。

在此基础上,将采集信号的噪声滤除并提取出第一次敲击的声音,采用RCMDE方法提取一次敲击声音信号中的多尺度管道特征;将构建的多维度管道特征矩阵输入到支持向量机(support vector machine,SVM)中,使用麻雀搜索算法(sparrow search algorithm,SSA)确定SVM参数最优值后进行模型训练并完成管道有无的分类,提出一种RCMDE和SSA优化的支持向量机(sparrow search algorithm-support vector machine,SSA-SVM)相结合的墙体管道敲击探测方法。最后将此方法与基于功率谱估计的墙体管道敲击探测方法和基于小波包分析的墙体管道敲击探测方法进行对比。

1 基于RCMDE的墙体管道敲击探测方法

1.1 特征提取

根据敲击探测理论,敲击墙体不同区域,存在管道的区域比正常墙体区域频率更低^[9]。因此可以通过对墙体敲击声音信号的频率分析来判断墙体内部是否有管道。

1) 散布熵算法

DE方法能够检测频率和幅值的变化,更有效的提取信息,且其计算时间也大大缩短,抗干扰性强,对噪声不敏感,可以有效反应墙体管道特征^[10]。

对采集到的敲击声音信号 $x = \{x_j, j = 1, 2, \dots, N\}$, 利用正态分布函数映射到 $y = \{y_1, y_2, \dots, y_N\}, y \in (0, 1)$, 即:

$$y_j = \frac{1}{\delta \sqrt{2\pi}} \int_{-\infty}^{x_j} \exp\left(-\frac{(t-\mu)^2}{2\delta^2}\right) dt \quad (1)$$

式中: μ, δ 分别是期望和标准差。

将 y 映射到 $z_j^{(c)}$, 即:

$$z_j^{(c)} = \text{int}(cy_j + 0.5) \quad (2)$$

式中: int 为取整函数, c 为类别个数。计算嵌入向量 $z_j^{(m,c)}$:

$$z_i^{(m,c)} = (z_i^{(c)}, z_{i+d}^{(c)}, \dots, z_{i+(m-1)d}^{(c)}), i = 1, 2, \dots, N - (m-1)d \quad (3)$$

式中: m 为嵌入维数, d 是时延。

若 $z_j^{(c)} = \nu_0, z_{i+d}^{(c)} = \nu_1, \dots, z_{i+(m-1)d}^{(c)} = \nu_{m-1}$, 则 $z_i^{(m,c)}$ 的散步模式为 $\pi_{\nu_0 \nu_1 \dots \nu_{m-1}}$ 。

每种散步模式的概率为:

$$P(\pi_{\nu_0 \nu_1 \dots \nu_{m-1}}) = \frac{\text{Num}(\pi_{\nu_0 \nu_1 \dots \nu_{m-1}})}{N - (m-1)d} \quad (4)$$

其中, $\text{Num}(\pi_{\nu_0 \nu_1 \dots \nu_{m-1}})$ 表示 $z_i^{(m,c)}$ 映射到 $\pi_{\nu_0 \nu_1 \dots \nu_{m-1}}$ 的个数 c^m 。

根据香农熵的定义, 敲击声音信号的DE值为:

$$E_{DE}(x, m, c, d) = - \sum_{\pi=1}^{c^m} P(\pi_{\nu_0 \nu_1 \dots \nu_{m-1}}) \ln(P(\pi_{\nu_0 \nu_1 \dots \nu_{m-1}})) \quad (5)$$

2) 精细复合多尺度散布熵算法

RCMDE是基于多尺度散布熵(multiscale dispersion entropy, MDE)提出的。MDE是先对敲击声音信号进行不同尺度的粗粒化,再计算其散布熵值。在计算过程中没有考虑数据间的关系,当采集到的敲击声音信号时间过短时,很可能会出现未定义熵,使结果出现误差^[11]。RCMDE是先将敲击声音信号分割、粗粒化,再计算每个粗粒化序列散布模式的概率,求得均值再计算RCMDE^[12]。这样处理能够减少粗粒化过程中部分统计信息的丢失,减少计算误差,克服了MDE的缺点。

当尺度因子为 τ 时,计算粗粒化序列。将采集到的敲击声音信号 $\{u(i), i = 1, 2, \dots, L\}$ 分成每段长度为 τ 的小段,并算出每小段的均值按大小顺序排列。第 k 个粗粒化序列 $x_k^\tau = \{x_{k,1}^{(\tau)}, x_{k,2}^{(\tau)}, \dots\}$ 为:

$$x_{k,j}^\tau = \frac{1}{\tau} \sum_{i=k+\tau(j-1)}^{k+j\tau-1} u_i, 1 \leq j \leq L/\tau, 1 \leq k \leq \tau \quad (6)$$

算出每个粗粒化序列的散布模式 π 的概率,求出平均值:

$$\bar{P}(\pi_{\nu_0 \nu_1 \dots \nu_{m-1}}) = \frac{1}{\tau} \sum_{k=1}^{\tau} P_k^{(\tau)} \quad (7)$$

RCMDE定义为:

$$E_{RCMDE}(X, m, c, d, \tau) = - \sum_{\pi=1}^{c^m} \bar{P}(\pi_{\nu_0 \nu_1 \dots \nu_{m-1}}) \ln(\bar{P}(\pi_{\nu_0 \nu_1 \dots \nu_{m-1}})) \quad (8)$$

根据Rostaghi等的研究表示,在DE的计算过程中,初始信号数据长度应大于2000, m 一般取2或3, d 取1, c 取4.8。而RCMDE的计算过程中,参数与DE计算的参数保持一致。

1.2 麻雀搜索算法优化的支持向量机算法

1) 支持向量机算法

SVM就是实现支持向量的分类,通过在输入数据(支持向量)的子集中构建生成一个最优超平面将圆形数据和三角形数据分隔开来执行二元分类问题,如图1所示^[13]。由于本文数据集是二分类、小规模、非线性的,而SVM对这类数据集进行分类时更有优势,因此采用SVM算法进

行墙体内有无管道的识别分类^[14]。

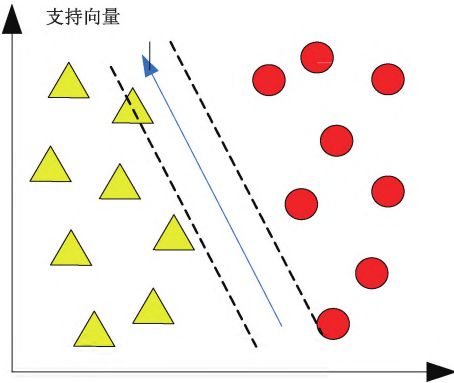


图 1 SVM 示意图

其算法运算过程如下所示:

对于训练集:

$$\{(x_i, y_i) \mid i = 1, 2, \dots, l; x_i \in R\} \quad (9)$$

回归函数表达式为:

$$f(x) = \omega^T \varphi(x) + b \quad (10)$$

式中: ω 为超平面权值, b 为偏差量。

求解最优方程:

$$\min = \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^l (\xi_i + \xi_i^*) \quad (11)$$

$$s. t. \quad y_i \omega^T \varphi(x) + b - y_i = \xi_i$$

式中: $C > 0$ 为惩罚因子, ξ 和 ξ^* 为引入的松弛系数。采用拉格朗日函数将式(10)转化为:

$$\omega(\alpha, \alpha^*) = -\frac{1}{2} \sum_{i,j=1}^l (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)(x_i \cdot x_j) + \sum_{i=1}^l (\alpha_i - \alpha_i^*) y_i - \sum_{i=1}^l (\alpha_i + \alpha_i^*) \epsilon \quad (12)$$

最后得到 SVM 回归模型为:

$$\hat{y}(x) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) K(x_i, x) + b \quad (13)$$

式中: $K(x_i, x) = \varphi(x_i) \cdot \varphi(x)$ 为核函数, 这里采用专门用于非线性样本而且参数较少的高斯径向基核函数 $K(x_i, x) = \exp(-\|x_i - x\|^2 / \sigma^2)$ 。

2) 麻雀搜索算法优化支持向量机模型

本文采用的高斯径向基核函数参数为惩罚因子 C 和核函数参数 σ , C 会影响 SVM 的泛化能力, 影响训练误差和分类模型的复杂度; σ 表明训练集数据的范围特性, 影响 SVM 分类模型的学习能力^[15]。SVM 分类模型的优劣与其参数有直接的联系^[16]。

为了提升分类的准确率, 通常会采用智能优化算法进行参数寻优, 如遗传算法, 粒子群算法, 麻雀搜索算法等。SSA 是 Xue 等^[17]于 2020 年根据麻雀的觅食和反捕食行为提出的新型群智能算法。因其具有高性能的搜索能力, 寻优能力强、收敛速度快、并且能够避免陷入局部最优的问题^[18]。因此本文选用 SSA 进行寻优。

SSA 优化 SVM 模型的流程为:

将输入的数据进行归一化处理, 构建 SVM 预测模型 (support vector regression, SVR); 初始化 SSA 参数, 定义 C 和 g 的取值范围; 定义适应度函数并计算个体适应度值; 计算最佳适应度并更新麻雀最佳个体位置, 得到 C 和 g 的最优值; 以最优参数 C 和 g 进行 SVM 分类模型训练; 将测试集输入预测模型, 得到分类结果。流程如图 2 所示^[19]。

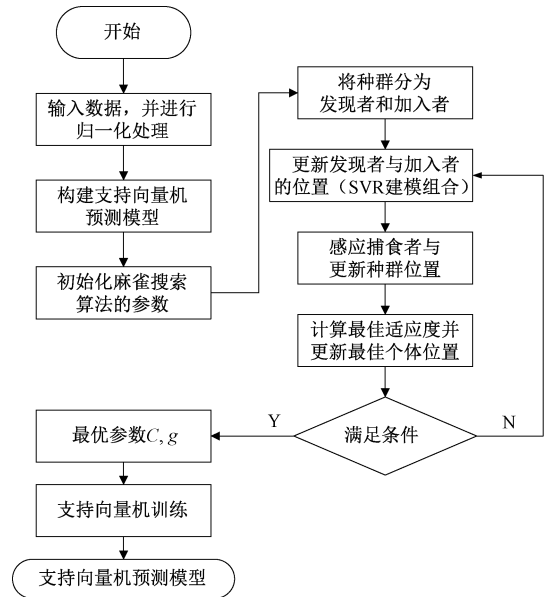


图 2 SSA-SVM 流程

2 基于 RCMDE 的墙体管道探测实验

2.1 实验装置及步骤

将一截 PVE 水管封入混凝土墙左边模拟墙体和墙体内管道, 如图 3(a) 所示。用两条长度相等的鱼线将一铁球悬挂在铁架上, 以保证能准确地敲击在墙上指定的位置, 且可以通过控制抬起小球的角度使其下落来控制对墙体的敲击力度, 如图 3(b) 所示。将铁架放置于墙体 10 cm 处, 如图 4 所示, 声音采集装置 VS1053 模块放在墙体模型左侧 5 cm 处。

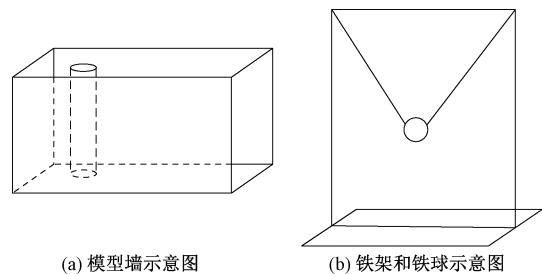


图 3 实验装置示意图

实验在模型墙上的不同部分进行敲击, 采集时将小球拉至与墙面垂直 90° 时放下, 以保证每次敲击力度相同; 将采集到的敲击声音信号进行预处理; 计算处理后的敲击声



图 4 墙体内管道敲击探测实验模型图

音信号的 RCMDE, 选择合适尺度下的 RCMDE 值作为特征值; 对 SVM 参数进行初始化, 从所有墙体敲击声音信号的特征值样本中随机选取一些样本作为训练集, 利用 SSA 算法优化 SVM 参数进行学习训练; 将特征向量输入得到的分类模型进行匹配分析得到探测结果。探测流程如图 5 所示。

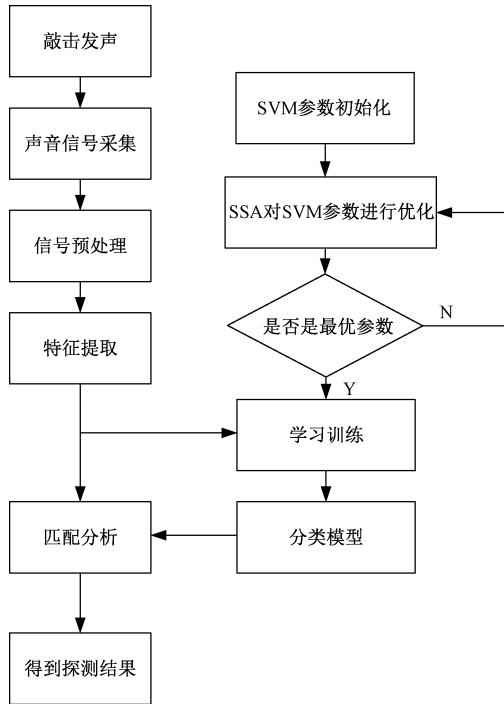
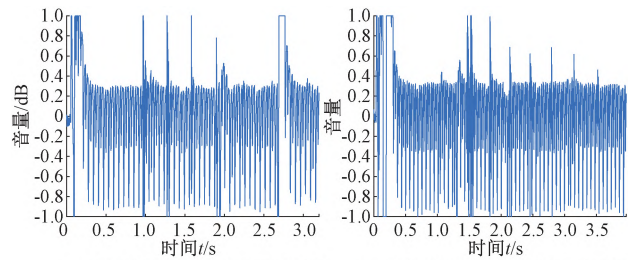


图 5 基于 RCMDE 的墙体内管道探测流程

2.2 实验数据分析

现场采集到的信号是由多次敲击声音、大量现场噪声和异常信号等组成且敲击声音被淹没在噪声中, 无法直接拿来作管道存在分析, 如图 6 所示。为了获得干净的敲击声音信号, 使其便于分析, 需要对其进行预处理——滤除噪声以及提取端点。预处理后得到的一次敲击声音信号如图 7 所示。

计算墙体一次敲击声音信号的 RCMDE 值, 不同尺度 RCMDE 均值方差图如图 8 所示, 尺度因子 τ 最大为 15。在图 8 中可以发现, 随着尺度因子 τ 增大, 有管道墙体的敲击声音信号的 RCMDE 值和无管道墙体的敲击声音信号的



(a) 原始墙体内有管道信号时域图 (b) 原始墙体内无管道信号时域图

图 6 信号时域图

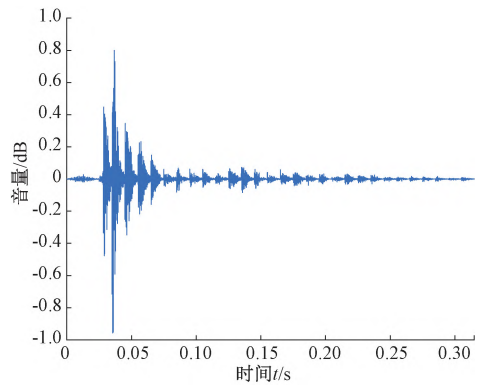


图 7 一次敲击声音信号时域图形

RCMDE 值都在减小, 且下降速度也逐渐降低。有管道和无管道两部分墙体的敲击声音信号对应的 RCMDE 曲线可以区分, 且有管道墙体的敲击声音信号的 RCMDE 值比无管道墙体的敲击声音信号的 RCMDE 值大, 也就是有管道墙体的敲击声音信号的频率比无管道墙体的敲击声音信号的频率高, 符合敲击探测理论, 证明采用 RCMDE 方法提取墙体内管道特征进行分类是可行的。

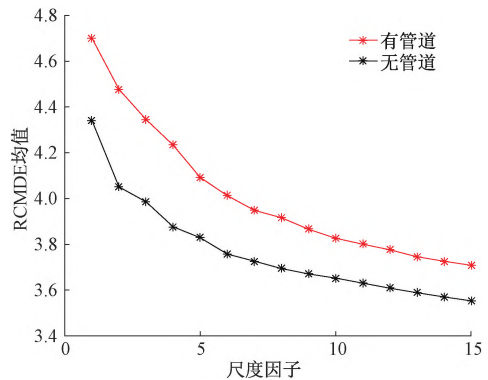


图 8 敲击声音信号的 RCMDE 均值

从图 8 中还可以看出, 墙体敲击声音信号的 RCMDE 值在尺度因子较大时差别较小, 不利于分类识别; 而在尺度因子较小时, 特征值数量较少, 可能不能完整的反映出管道特征, 导致分类准确率下降。综合考虑, 选择前 6 个尺度的 RCMDE 值作为敲击声音信号的管道特征向量, 用于表达信号包含的管道信息, 输入 SSA 优化的 SVM 分类模型进行墙体内有无管道的分类。

经过计算后,每个敲击声音信号可以得到 6 个特征值。采集到的有管道部分墙体敲击声音信号和无管道部分墙体敲击声音信号各 157 个。将这些敲击声音信号进行分析,提取特征值,得到有 314 个样本的特征值数据集。将数据集分成管道和无管道两类,有管道样本标记为 1,无管道样本标记为 0。其中从有管道样本中选择 136 个,无管道样本中选择 103 个作为训练集,余下 75 个样本为测试集。将训练集中的特征向量输入 SSA-SVM 中训练出分类模型,对测试样本分类,将分类结果比较,得出预测的准确率。定义麻雀种群为 100,最大迭代次数 1 000,目标函数的维度 $D=2$,初始值的上界为 100,下界为 0.01,得到最佳准确率作为适应度函数。最后得到最佳的惩罚因子 c 为 58.982 3,核函数 g 为 0.658 125,探测准确率达到 97%。实际所得出的结果与预测所得出的结果如图 9 所示。因此,基于 RCMDE 的墙体管道探测方法能准确的探测出墙体内是否有管道。

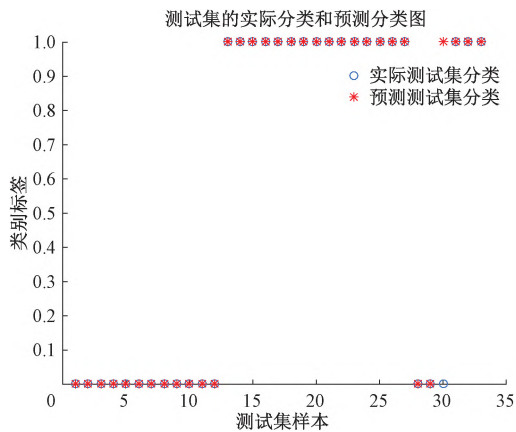


图 9 基于 RCMDE 方法的实际分类和预测分类图

为了验证此方法的优越性,将此方法与基于功率谱估计的墙体管道敲击探测方法和基于小波包分析的墙体管道敲击探测方法进行对比。采用 welch 方法估计功率谱,将功率谱中 7 个最明显的峰值作为特征向量输入 SVM 分类模型进行墙体内是否存在管道的判断,探测准确率为 89%,实际所得出的结果与预测所得出的结果如图 10 所示。

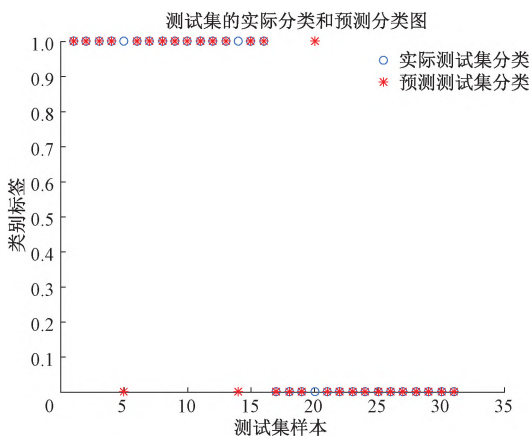


图 10 基于功率谱估计方法的实际分类和预测分类图

示;采用 db7 小波基对敲击信号做 3 层小波分解,将 8 个节点的能量值作为特征参数输入 SVM 分类模型进行管道有无的分类,探测准确率为 64%,实际所得出的结果与预测所得出的结果如图 11 所示。3 种方法的对比结果如表 1 所示。

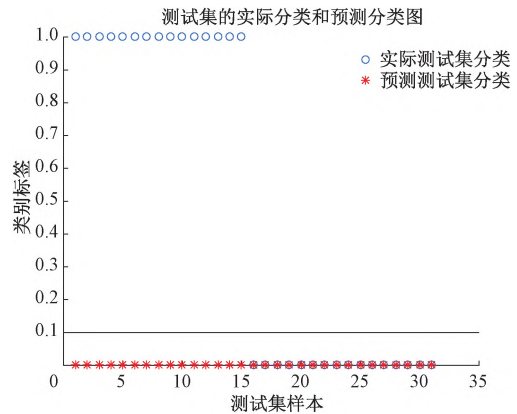


图 11 基于小波包分析方法的实际分类和预测分类图

表 1 3 种信号分析方法分类准确率对比

方法	惩罚因子 c	核参数 g	准确率/%
功率谱+SSA-SVM	100	0.5	89.8
小波包+SSA-SVM	56	0.37	64
RCMDE+SSA-SVM	58.982 3	0.658 125	96.97

由表 1 可知,相比于基于功率谱估计的墙体管道敲击探测方法和基于小波包分析的墙体管道敲击探测方法,RCMDE 方法对于墙体管道探测准确率更高。

3 结 论

本文提出了基于 RCMDE 和 SSA-SVM 的墙体管道探测方法,将采集到的墙体敲击声音信号的噪声滤除并通过双门限算法进行端点提取得到第一次敲击的声音,采用 RCMDE 方法提取敲击声音信号中的多尺度管道特征;将构建的多维度管道特征矩阵输入到支持向量机中,利用 SSA 算法得到 SVM 最优参数后进行分类模型的学习训练,实现对有管道墙体部分敲击声音信号样本的自动识别。经实验验证,相对于功率谱估计和小波包分析提取的管道特征,RCMDE 提取的管道特征误差更小,稳定性更高;而且所需特征向量数量更少,识别率更高达到了 97%,能够准确的区分有管道样本和无管道样本。

参考文献

- [1] 李瑾,行鸿彦,吴叶丽,等.基于敲击声音信号的墙体管道探测方法[J].国外电子测量技术,2022,41(2):1-6,DOI:10.19652/j.cnki.femt.2103310.
- [2] 李帅永,韩明秀,文井辉.基于 VMD-SVD 自优化的管道微泄漏信号增强方法[J/OL].电子测量与仪器学报:1-13 [2022-01-05].http://kns.cnki.net/kcms/

- detail/11.2488.TN.20211208.1214.004.html.
- [3] 王晓娟,赵锴,郑毅.管道纵向超声导波的对称模态转换特性分析[J].仪器仪表学报,2020,41(9):151-160, DOI:10.19650/j.cnki.cjsi.J2006453.
- [4] 曹丽,路加,刘杰,等.基于振动敲击的软质物体硬度检测[J].仪器仪表学报,2002(S1):74-76, DOI:10.19650/j.cnki.cjsi.2002.s1.035.
- [5] 王铮,李硕宁,郭广平.敲击检测技术在某雷达天线罩在役检测中的应用[J].无损检测,2012,34(6):29-32.
- [6] ROSTAGHI M, AZAMI H. Dispersion entropy: A measure for time-series analysis [J]. IEEE Signal Processing Letters, 2016, 23(5): 610-614.
- [7] HAMED, AZAMI, MOSTAFA, et al. Refined composite multiscale dispersion entropy and its application to biomedical signals [J]. IEEE Transactions on Bio Medical Engineering, 2017, 64(12): 2872-2879.
- [8] ZHOU F M, HAN J, YANG X Q. Multivariate hierarchical multiscale fluctuation dispersion entropy: Applications to fault diagnosis of rotating machinery[J]. Applied Acoustics, 2021, 182: 108271.
- [9] CAWLEY P, ADAMS R D. The mechanics of the coin-tap method of non-destructive testing [J]. Academic Press, 1988, 122(2): 299-316.
- [10] 郑近德,李从志,潘海洋.复合多尺度散布熵在滚动轴承故障诊断中的应用[J].噪声与振动控制,2018, 38(S2):653-656.
- [11] 李从志,郑近德,潘海洋,等.基于精细复合多尺度散布熵与支持向量机的滚动轴承故障诊断方法[J].中国机械工程,2019,30(14):1713-1719,1726.
- [12] 王金东,欧凌非,赵海洋,等.基于 CEEMDAN 和 RCMDE 的往复压缩机轴承故障诊断方法[J].机床与液压,2021,49(5):168-172,161.
- [13] 王清杰,全海燕.基于单形进化算法优化支持向量机的运动想象脑电分类研究[J].电子测量与仪器学报, 2021,35(9):157-163, DOI:10.13382/j.jemi.B2103989.
- [14] 梁钊.基于敲击信号的刹车片内部缺陷检测方法研究[D].淄博:山东理工大学,2019.
- [15] 张熙,许浩,李沁程.基于 SVM 的飞机驾驶舱头顶板字符识别方法研究[J].电子测量技术,2020,43(22):82-87, DOI:10.19651/j.cnki.emt.2005171.
- [16] 张作霖,王利恒.基于小波分析和 GA-SVM 的小电流接地故障选线方法[J].自动化与仪表,2021,36(8):8-12,23, DOI:10.19557/j.cnki.1001-9944.2021.08.002.
- [17] XUE J, SHEN B. A novel swarm intelligence optimization approach: Sparrow search algorithm[J]. Systems Science & Control Engineering, 2020, 8(1): 22-34.
- [18] 魏鹏飞,樊小朝,史瑞静,等.基于改进麻雀搜索算法优化支持向量机的短期光伏发电功率预测[J].热力发电, 2021,50(12):74-79, DOI:10.19666/j.rld.202104127.
- [19] 汤安迪,韩统,徐登武,等.基于混沌麻雀搜索算法的无人机航迹规划方法[J].计算机应用,2021,41(7): 2128-2136.

作者简介

李瑾, 硕士研究生, 主要研究方向为仪器仪表技术, 信号检测与处理。

E-mail: 346909631@qq.com

行鸿彦(通信作者), 教授、博士生导师, 主要研究方向为气象仪器设计与计量, 信号检测与处理等。

E-mail: xinghy@nuist.edu.cn