

DOI:10.19651/j.cnki.emt.2107230

基于CEEMDAN-HD-云模型特征熵的 油气管道工况识别^{*}

张勇^{1,2} 杨文武^{1,2} 王明吉¹ 孙桐^{2,3} 刘洁^{1,2} 周兴达^{1,2}(1.东北石油大学 物理与电子工程学院 大庆 163318; 2.东北石油大学 人工智能能源研究院 大庆 163318;
3.黑龙江省网络化与智能控制重点实验室 大庆 163318)

摘要: 针对长输油气管道泄漏检测过程中泄漏信号特征信息提取困难,提出一种新的管道负压波信号特征提取方法。采用添加自适应噪声的完备集合经验模态分解算法对采集的负压波信号进行去噪,通过评估CEEMDAN分解后分量与原始信号的概率密度之间的豪斯多夫距离选取有效模态并重构。计算重构信号的云模型特征熵、峭度作为特征参数,用支持向量机进行分类识别。通过实验室数据验证,CEEMDAN、豪斯多夫距离与云模型特征熵结合的方法可以有效提高油气管道泄漏检测的准确性,实现了对流量小于4³ m/h的微小泄漏信号的识别,具有一定的现场应用价值。

关键词: CEEMDAN;豪斯多夫距离;云模型特征熵;支持向量机

中图分类号: TE832 文献标识码: A 国家标准学科分类代码: 440.55

Identification of oil and gas pipeline working condition based on CEEMDAN-HD-cloud model feature entropy

Zhang Yong^{1,2} Yang Wenwu^{1,2} Wang Mingji¹ Sun Tong^{2,3} Liu Jie^{1,2} Zhou Xingda^{1,2}(1. School of Physics and Electronic Engineering, Northeast Petroleum University, Daqing 163318, China;
2. Artificial Intelligence Energy Research Institute, Northeast Petroleum University, Daqing 163318, China;
3. Key Laboratory of Networking and Intelligent Control of Heilongjiang Province, Daqing 163318, China)

Abstract: Aiming at the difficulty in extracting the feature information of the leakage signal in the process of long-distance oil and gas pipeline leakage detection, a new pipeline negative pressure wave signal feature extraction method is proposed. A complete set of empirical mode decomposition algorithm with adaptive noise is used to denoise the collected negative pressure wave signal, and the Hausdorff distance between the probability density of the component after CEEMDAN decomposition and the original signal is evaluated. Select the effective mode and reconstruct. The cloud model feature entropy and kurtosis of the reconstructed signal are calculated as feature parameters, and the support vector machine is used for classification and recognition. Through laboratory data verification, the method of combining CEEMDAN, Hausdorff distance and cloud model feature entropy can effectively improve the accuracy of oil and gas pipeline leak detection, and realize the identification of small leak signals with a flow rate of less than 4³ m/h. Certain field application value.

Keywords: CEEMDAN; Hausdorff distance; cloud model feature entropy; SVM

0 引言

随着我国对石油和天然气等重要战略资源需求量的提升,作为石油、天然气主要运输方式的长输油气管道发展迅猛。相比其他运输方式,长输管道运输具有低污染、高效率

的优点。在管道运输业大力发展的同时,由于管道运行年限增长导致的腐蚀等自然因素以及施工破坏等人为因素,管道事故发生的次数频频增加,对国家和人民造成巨大损失。因此对管道泄漏检测技术的发展、检测准确率的提高等关键性问题的研究具有重要的理论意义和现实价值。

收稿日期:2021-07-10

*基金项目:国家自然科学基金(61873058)、教育部重点实验室开放基金(MECOF2019B02)项目资助

目前各高校及相关单位已研究出许多不同种类的输油管道泄漏检测方法,主要有光纤检测法、声波法、负压波法^[1]等,由于长输油气管道工况复杂,目前并没有一种通用的泄漏检测方法可以适应所有的工况。对不同方法进行分析,发现负压波法具有检测时间短、反应速度快、对系统要求低、容易实施、可定位等优点。在成本有限的情况下能够达到较高的精度,目前被各行业广泛应用。负压波法主要包含信号去噪、特征提取、工况识别 3 个主要步骤。设计合适的去噪算法及特征向量可以有效提高负压波检测准确率。由于长输油气管道工况的复杂性,负压波信号在传输的过程中不可避免的掺杂噪声,噪声的存在会淹没泄漏信号特征,因此需要采用非线性处理方法对负压波信号进行去噪预处理,已有研究人员将小波阈值去噪、经验模态分解(empirical mode decomposition, EMD)、总体经验模态分解(ensemble empirical mode decomposition, EEMD)应用于管道负压波信号的预处理之中^[2],但这些方法存在一些不足。如利用小波阈值去噪,小波基以及小波层数选取复杂,利用 EMD 分解信号容易出现模态混叠,利用 EEMD 分解信号无法抵消添加的噪声。自适应噪声的完备集合经验模态分解(complete ensemble empirical mode decomposition with adaptive noise, CEEMDAN)是在 EEMD 分解时加入特定的白噪声,有效解决了上述问题,本文选用 CEEMDAN 算法处理管道负压波信号,对其进行去噪预处理。目前 CEEMDAN 算法被广泛应用于轴承故障诊断。文献[3]利用信号质量指数算法(signal quality index, SQI)选择 CEEMDAN 分解后的分量,结合奇异值分解(singular value decomposition, SVD)对齿轮裂纹数据进行有效的特征提取。

利用 CEEMDAN 对负压波信号进行处理时,有效模态的选取直接影响去噪效果^[4]。目前模态选择的主要方法有相关系数法、欧氏距离法、能量熵等。文献[4]将 CEEMDAN 算法应用于管道负压波信号,利用相关系数法选取有效模态,并利用支持向量机(support vector machine, SVM)分类,达到良好识别效果。文献[5]将 VMD 算法与蒙斯多夫距离(Hausdorff distance, HD)结合起来用于管道声波信号,目前没有人将 HD 与 CEEMDAN 结合并应用于管道负压波信号。本文通过评估 CEEMDAN 分解后分量与原始信号概率密度函数(probability density function, PDF)的 HD 来筛选有效分量,将分量重构进而达到降噪目的。

特征参数的选择是影响管道泄漏检测准确率主要因素之一。近年来,许多非线性动力学方法如排列熵、样本熵、近似熵、模糊熵等被用于特征提取^[6]。但是排列熵没考虑信号幅值之间的关系、样本熵计算速度慢、近似熵过度依赖数据长度等缺点降低了管道检测的准确率。本文将云模型特征熵应用于管道信号的特征提取,有效地解决了上述问题,并且云模型特征熵不用设置阈值和维数,消除了参数选

择带来的困难,有效地避免了不确定性带来的问题^[7]。

基于上述分析,为更好地选取 CEEMDAN 算法分解后的有效分量及解决信号中的噪声干扰问题,本文提出一种基于蒙斯多夫距离与 CEEMDAN 结合的算法,为解决负压波信号特征提取困难,构造一种 CEEMDAN 与云模型特征熵结合的特征提取方法,并将它们应用到油气管道泄漏检测研究中,通过实验室管道验证了所提方法的可行性。

1 降噪算法理论

1.1 CEEMDAN 分解

CEEMDAN 是在集合经验模态算法的基础上改进的。EEMD 具有无法完全抵消所加入的白噪声、计算时间长等缺点^[8]。2011 年 Torres 等提出 CEEMDAN 方法,既能减少训练次数,又能有效解决模态混叠和端点效应^[9-10]。

CEEMDAN 算法原理具体实现步骤如下。

1) 根据 EEMD 方法分解信号 $X(t) + \epsilon_0 \omega(t)$, 得到第 1 个分量, 其中 $X(t)$ 为原始信号, $\omega(t)$ 为满足 $N(0,1)$ 的白噪声。

$$\overline{IMF_1} = \frac{1}{n} \sum_{i=1}^n IMF_{1i}(t) \overline{IMF_1} \quad (1)$$

2) 当 $k=1$ 时, 计算第 1 个唯一余量:

$$r_1[t] = X[t] - \overline{IMF_1} \quad (2)$$

3) 分解 $r_1[t] + \epsilon_1 E_1(W_1[t])$ ($i=1, 2, \dots, n$) 至第 1 模态分量, 其中 $E_i(\cdot)$ 表示对信号进行分解得到的第 i 个分量, 在此基础上定义第 2 个模态分量为:

$$\overline{IMF_2} = \frac{1}{n} \sum_{i=1}^n E_1(r_1[t] + \epsilon_1 E_1(W_1[t])) \quad (3)$$

4) 对于 $k=2, \dots, K$, 计算第 k 个余量:

$$r_k[t] = r_{k-1}[t] - \overline{IMF_k} \quad (4)$$

5) 分解 $r_k[t] + \epsilon_k E_k(W_k[t])$ ($i=1, 2, \dots, n$) 至第 1 模态分量, 第 $k+1$ 个模态分量可以定义为:

$$\overline{IMF_{k+1}} = \frac{1}{n} \sum_{i=1}^n E_k(r_k[t] + \epsilon_k E_k(W_k[t])) \quad (5)$$

6) 将 k 加 1, 返回步骤 4)。

当余量不适合被分解时(不包含至少两个极点),停止分解最终的余量满足:

$$\overline{R[t]} = X[t] - \sum_{k=1}^K \overline{IMF_k} \quad (6)$$

其中, K 为分解得到的固有模态函数的个数。

1.2 蒙斯多夫距离

HD 可以度量概率密度函数之间的距离,是评价相似程度的指标^[11-12]。为了实现筛选包含原始信号特征信息较多的分量,达到去噪目的。本文利用 HD 测量重构信号与各分量概率密度函数之间的相似性,具有良好的应用效果。分量与原始信号相似程度 $hd(i)$ 定义如下:

$$hd(i) = HD[pdf(x(t)), pdf(BLIMF_i(t))] \quad (7)$$

式中: HD 为所求的概率密度之间的 Hausdorff 距离值; i 为分量个数; $pdf(x(t))$ 为原始管道信号的概率密度;

$pdf(BLIMF_i(t))$ 为分量的概率密度。

在利用 CEEMDAN 分解管道泄漏信号时,通常认为低频模态分量与原始信号的相关性比较大,高频模态分量与原始信号的相关性很小^[13]。本文利用豪斯多夫距离方法选取有效模态的步骤为:

- 1) 计算分量与原始信号的概率密度。
 - 2) 利用式(7)求解每个分量与原信号的 HD。
 - 3) 利用式(8)找到 HD 最大减少量,确定分界模态。
 - 4) 区分噪声模态、有效模态。
- $$\theta_i = L(i+1) - L(i), i = 1, 2, \dots, (N-1) \quad (8)$$
- $$m = \operatorname{argmax} \theta_i$$

式中: θ_i 为相邻模态与原始信号的 HD 减少量, m 为分量分界点。

2 特征提取算法

2.1 云模型理论及逆向云发射器

云模型特征熵参数选择较为简单,能解决不确定性问题,在故障诊断中获得了广泛应用^[14]。已知一组定量数据,可以用 Ex , En , He 3 个参数来表征其定性概念,其中 Ex 为此数据的数学期望, En 为特征熵, He 为特征熵的熵值,将其命名为超熵^[15]。

将定量数据转换为定性语言是利用逆向云发生器计算的。已知管道信号 $y_i (i=1, 2, \dots, n)$, 利用逆向云发生器计算 En 步骤如下:首先求出 4 096 个管道数据点的期望值,将期望代入式(9)可得云模型特征熵。

$$Ex = \frac{1}{n} \sum_{i=1}^n y_i \quad (9)$$

$$En = \sqrt{\frac{\pi}{2}} \times \frac{1}{n} \sum_{i=1}^n |y_i - Ex|$$

2.2 峰度

峰度(Kurtosis)是描述信号波峰尖度的无量纲参数^[16],计算表达式为:

$$K = E(x - \mu)^4 / \sigma^4 \quad (10)$$

截取一段信号 x ,首先算出其标准差 σ 、均值 μ ;将它们代入式(10)即可求得峰度,在软件中利用函数 Kurtosis() 可实现对信号峰度的求解。峰度指标能较好地反映管道泄漏信号中的冲击成分。因此本文将重构信号的峰度作为特征参数,可以较好地区分不同工况信号。

2.3 本文提出的方法

根据对 CEEMDAN、豪斯多夫距离、云模型特征熵理论的深入研究,提出基于 CEEMDAN-HD-云模型特征熵的管道泄漏检测方法,其流程如图 1 所示,其步骤如下。

1) 利用实验室长输管道泄漏检测实验平台采集不同工况下的负压波信号,通过 CEEMDAN 算法将采集的管道信号自适应分解为多个模态分量(无需设置模态个数)。

2) 通过评估 HD 筛选出有效模态重构,得到去噪信号。

3) 提取信号特征,计算重构信号的云模型特征熵、峰度。

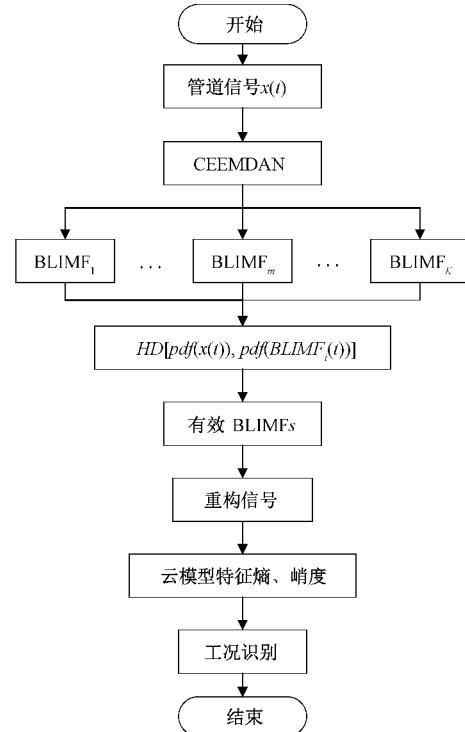


图 1 基于 CEEMDAN-HD-云模型特征熵的特征提取算法流程

4) 将特征参数输入 SVM 中进行工况识别。

3 仿真实验与结果分析

为验证 CEEMDAN-HD 去噪算法对管道负压波泄漏信号处理的有效性,首先利用加入白噪声的 Heavy sine 曲线进行仿真,其信噪比标准方差为 6, 染噪信号如图 2 所示。

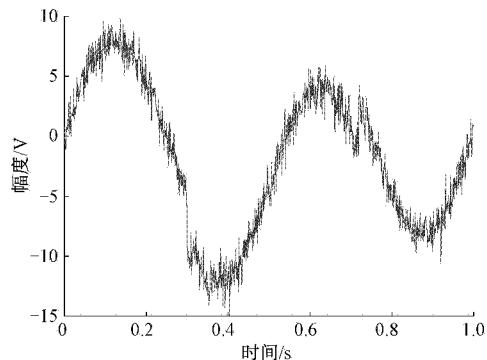


图 2 染噪信号

利用 CEEMDAN 分解含噪信号,设置 CEEMDAN 算法添加的白噪声为强度 0.2,添加次数设为 50 次,最大包络次数设定为 5 000,得到 10 个分量,计算各模态分量与原始含噪信号概率密度函数的豪斯多夫距离如图 3 所示,根据式(8)原则选模态 8、9、10 作为有效分量,其余为噪声分量,将有效分量重构,结果如图 4 所示。

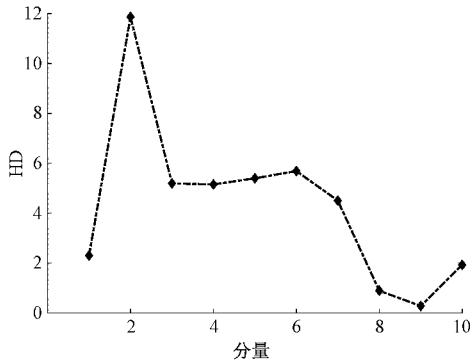


图 3 仿真信号豪斯多夫距离

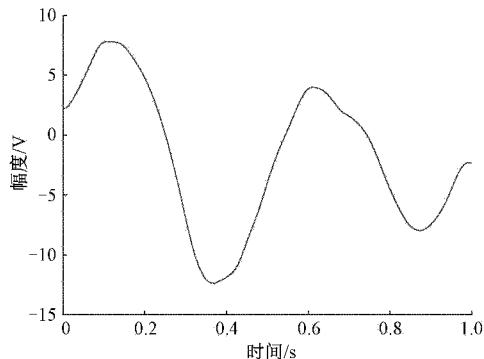


图 4 重构信号

用小波阈值去噪算法、EEMD 与豪斯多夫距离结合算法作对比,选择小波基为 db5,分析去噪效果。结果如表 1 所示,由表 1 可以看出,利用 CEEMDAN-HD 算法处理仿真信号得出重构信号的信噪比为 25.464 4,高于 EEMD-HD 去噪算法和其他对比算法处理仿真信号得到的信噪比,利用其他对比算法得到的均方根误差均高于本文算法。可见本文改进算法得到了较为理想的去噪效果。

表 1 仿真信号去噪后的信噪比与均方根误差

Heavy sine	软阈值函数	硬阈值函数	EEMD-HD	改进算法
SNR	24.853 8	24.202 2	23.616 3	25.464 4
RMSE	0.356 4	0.384 1	0.411 0	0.332 2

4 管道泄漏信号实验分析

4.1 数据采集

本次实验所用的数据采集自东北石油大学人工智能能源研究院院长输管道泄漏检测实验平台。实验平台所用管道型号为 DN80, 材质为不锈钢, 管壁厚度 4 mm, 可以完成气体和液体的实验。管道泄漏检测实验平台如图 5 所示。实验平台软件采用 Visual Studio 开发, 使用 Art 数据采集卡将来自各个传感器的数据输入到计算机中, 采样频率设置为 1 024 Hz。

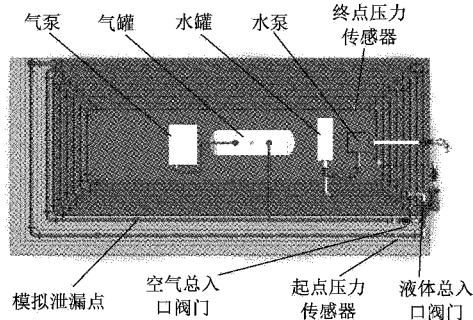


图 5 实验平台原理

4.2 负压波信号的分解与重构

人工选取适用于本实验的有效泄漏信号, 其中包含正常信号、泄漏信号、微小泄漏信号(泄漏量小于 4^3 m/h)。波形如图 6 所示。

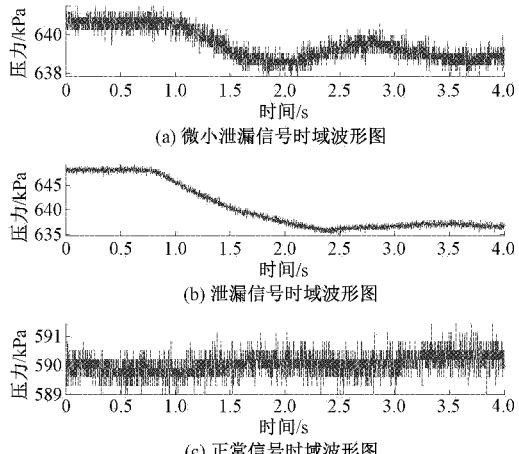


图 6 3 种工况信号的时域波形图

为解决噪声对泄漏信号分析产生的干扰, 更好地提取其特征信息, 本文将采集到的泄漏信号进行 CEEMDAN 分解, CEEMDAN 参数设置同上, 分解得到 13 个分量。利用式(7)计算模态分量与原始管道信号的概率分布的豪斯多夫距离结果如图 7 所示。

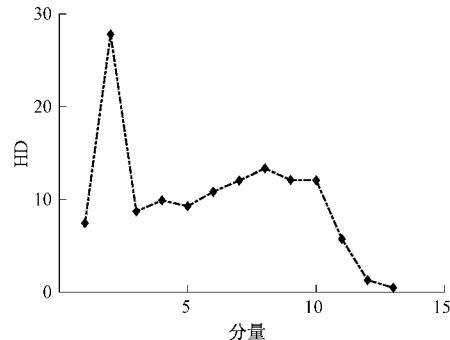


图 7 豪斯多夫距离

如图 7 所示, 由于高频模态分量与原始信号的相关性很小, 并从分解图可以看出, CEEMDAN 前 5 分量都属于

高频分量, 所以从第 6 个模态之后选取有效分量。从豪斯多夫距离曲线可知, 分量 10~11 的 HD 减少量最大, 所以选取模态分量 11 及之后的模态分量作为有效分量进行信号重构。选取模态分量 11、模态分量 12、模态分量 13 作为有效模态进行重构得到的信号如图 8 所示。

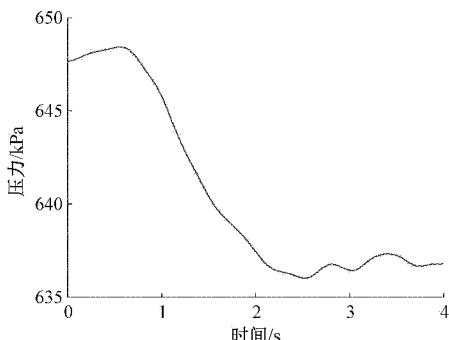


图 8 重构信号

4.3 重构信号的特征提取

为提高系统检测的精度, 本文选择重构信号的云模型特征熵、峭度作为特征向量。首先求解管道泄漏信号、微小泄漏信号、正常信号的云模型特征量, 此过程利用逆向云发生器计算。由表 2 可知, 泄漏信号的特征熵、超熵数值远大于正常信号、微小泄漏信号的特征熵、超熵数值。正常信号的云模型特征熵值不足微小泄漏信号的 1/2。

表 2 3 类信号重构后的云模型特征量

信号类型	特征熵 E_n	超熵 H_e	期望 E_x
泄漏	4.259 4	2.200 4	621.465 5
正常	0.178 2	0.086 8	590.043 2
微小泄漏	0.446 0	0.188 0	619.131 4

为验证本文提出的特征提取算法处理管道负压波信号的优越性, 截取 150 组长度为 4 s、频率为 1 024 Hz 的管道负压波信号, 其中正常、泄漏、微小泄漏信号各 50 组。计算得出的云模型特征熵拟合曲线如图 9 所示。由图可知, 3 种信号的云模型特征熵无交叉, 正常信号的云模型特征熵与泄漏信号的云模型熵差异较大, 与微小泄漏的云模型特征熵差异较小但是区分效果明显。

选取相同的 3 种工况信号(正常、泄漏、微小泄漏)各样本 50 组, 利用式(4)计算重构信号的峭度值, 得到峭度拟合曲线如图 10 所示。正常、泄漏信号经分解重构后计算得到的峭度分布非常稳定且无交叉, 小泄漏信号与泄漏信号的峭度有所交叉, 但与正常重构信号峭度基本无交叉, 区分也较为明显。因此, 云模型特征熵、峭度可以联合作为不同工况信号的特征参数, 可以有效地区分正常和其他状态。

4.4 训练与测试

本文采用 SVM 对提取的管道信号特征参数进行分类, 分类对象为管道运行时不同泄漏程度的 3 种工况信号。

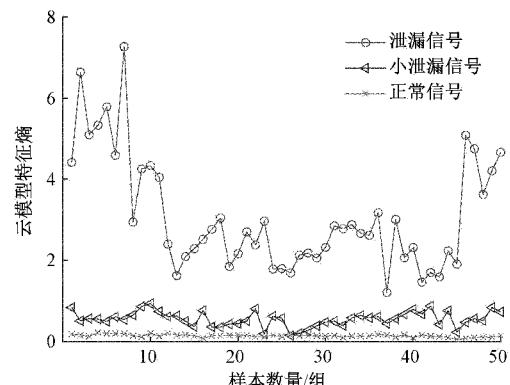


图 9 云模型特征熵拟合曲线

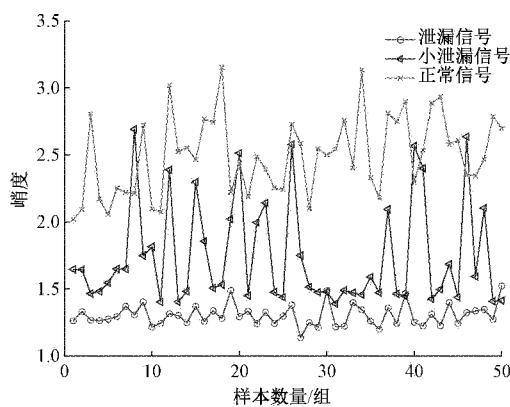


图 10 峭度拟合曲线

首先将云模型特征熵与峭度组成的特征向量进行归一化处理。利用 `randm` 函数将样本数据进行随机分配并设置标签, 其样本数据分配情况及标签设置情况如表 3 所示。

表 3 数据集分配表

管道工况	训练集/个	测试集/个	类别标签
正常	30	20	0
泄漏	30	20	1
微小泄漏	30	20	2

SVM 的核函数选择 RBF 核函数, 惩罚参数“-c”和核函数参数“-g”^[17]采用网格搜索法确定。将训练样本数据输入到 SVM 分类器中训练出分类模型, 把测试样本数据放入模型验证模型性能。SVM 测试集分类结果如图 11 所示。

为验证模型的优越性, 分别采用本文算法、EEMD 算法对管道信号进行分解, 然后求重构信号的云模型特征熵、峭度, 利用 SVM 对不同工况进行分类, 如表 4 所示。

由表 4 可以得出, 本文提出的方法在管道样本数量很有限的情况下, 工况识别分类的正确率达到 98.33%。而集合经验模态分解、豪斯多夫距离、云模型特征熵相结合的算法分类准确率为 91.67%。说明本文提出的算法能较好地处理复杂的管道信号, 有效提高识别准确率。

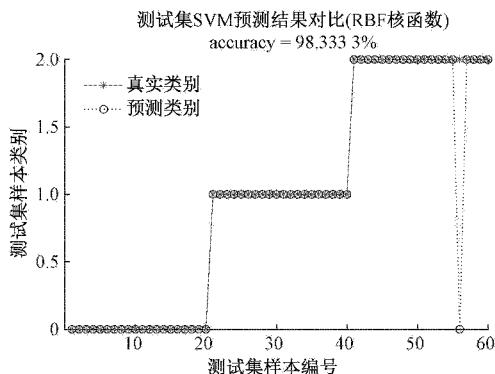


图 11 测试样本分类结果

表 4 CEEMDAN-HD-熵和 EEMD-HD-熵诊断结果

工况种类	分类结果		总体准确率/%
	测试样本/个	正确识别/个	
C-H-熵	60	59	98.33
E-H-熵	60	55	91.67

5 结 论

本文提出一种基于 CEEMDAN-HD-云模型特征熵的管道信号泄漏检测方法。通过仿真信号和实验室管道采集信号进行试验,与其他算法相比,本文提出的算法在信号去噪方面有着较好的优越性。不仅可以得到无失真、更平滑的低频滤波信号,并且能够很好地保留管道信号的特征信息,可以提高管道工况识别准确率。在 CEEMDAN-HD 滤波算法的基础之上,本文设计选取了云模型特征熵、峭度作为特征参数,较好地表征不同管道信号的特征信息,识别准确率高达 98.33%。与 EEMD-HD-云模型特征熵方法比较,准确率提高了 6.66%。说明本文所用方法能够有效区分正常、不同程度泄漏信号,为管道的工况识别提供了新途径。

参考文献

- [1] 王秀芳,王昕. EEMD 近似熵和支持向量机的管道泄漏特征向量提取研究[J]. 化工自动化及仪表,2015,42(11):1226-1229.
- [2] 周怡娜,董宏丽,张勇,等. 基于 VMD 去噪和散布熵的管道信号特征提取方法[J/OL]. 吉林大学学报(工学版),2021:1-13[2021-09-28].
- [3] 古莹奎,曾磊,张敏,等. 基于 CEEMDAN-SQI-SVD 的齿轮箱局部故障特征提取[J]. 仪器仪表学报,2019,40(5):78-88.
- [4] 李传宪,逯斐斐,石亚男,等. 基于改进 CEEMDAN-熵方法的管道泄漏工况识别[J]. 石油化工高等学校学报,2020,33(1):88-96.
- [5] 马雯萍. 基于 VMD 的天然气管道泄漏信号特征提取与检测技术研究[D]. 大庆:东北石油大学,2018.
- [6] 高纯生,周小云,黄祥海. 基于 CEEMDAN-多尺度模糊熵和 ISRNN 的球磨机负荷识别[J]. 矿业研究与开发,2020,40(4):141-146.
- [7] 蔡改贫,宗路,罗小燕,等. 基于 CEEMDAN-云模型特征熵和 LSSVM 的磨机负荷预测研究[J]. 振动与冲击,2019,38(7):128-133.
- [8] 周盛山,汤吉军,王金轩,等. EEMD 和 CNN-XGBoost 在风电功率短期预测的应用研究[J]. 电子测量技术,2020,43(22):55-61.
- [9] 周建民,李家辉,尹文豪,等. 基于 CEEMDAN 和 PSO-OCSVM 的滚动轴承性能退化评估[J]. 电子测量与仪器学报,2021,35(7):194-201.
- [10] LEI Y G, LIU Z Y, OUAZRI J, et al. A fault diagnosis method of rolling element bearings based on CEEMDAN [J]. Proceedings of the Institution of Mechanical Engineers, 2017, 231(10):1804-1815.
- [11] YANG G, LIU Y, WANG Y, et al. EMD interval thresholding denoising based on similarity measure to select relevant modes [J]. Signal Processing, 2015, 109:95-109.
- [12] MA W P, YIN S X, JIANG C L, et al. Variational mode decomposition denoising combined with the Hausdorff distance [J]. The Review of Scientific Instruments, 2017, 88(3):035109.
- [13] SHI M J, ZHAO H H, HUANG Z Q, et al. Signal extraction using complementary ensemble empirical mode in pipeline magnetic flux leakage nondestructive evaluation[J]. The Review of Scientific Instruments, 2019, 90(7):075101.
- [14] 阎玲玲,叶蕾,高内坤,等. 气体管道运行状态特征提取与状态识别[J]. 压力容器,2021,38(2):14-21.
- [15] 赵荣珍,马森财,吴耀春. 云模型和集成极限学习机相结合的滚动轴承故障诊断方法[J]. 兰州理工大学学报,2021,47(4):33-39.
- [16] 王建军,曹新城. 基于主轴振动信号的铣刀磨损监测研究[J]. 国外电子测量技术,2019,38(2):103-108.
- [17] WANG C, HAN F, ZHANG Y, et al. An SAE-based resampling SVM ensemble learning paradigm for pipeline leakage detection[J]. Neurocomputing, 2020, 403:237-246.

作者简介

张勇,工学硕士,副教授,主要研究方向为信号与信息处理。

E-mail:dqpyz@163.com

杨文武,硕士研究生,主要研究方向为信号处理与管道泄漏检测技术。

E-mail:m15776500877@163.com