

DOI:10.19651/j.cnki.emt.2313449

## VMD 和 SO 优化 SVM 的光纤复合海缆故障诊断研究\*

李俊卿<sup>1</sup> 刘若尧<sup>1</sup> 何玉灵<sup>2</sup> 张承志<sup>1</sup> 耿继亚<sup>1</sup>

(1. 华北电力大学电力工程系 保定 071003; 2. 华北电力大学机械工程系 保定 071003)

**摘要:** 为了进一步提高光纤复合海底电缆的故障诊断准确率,提出了基于 VMD 及 SO 优化 SVM 的故障诊断方法。首先,使用 VMD 对故障数据进行分解,得到若干条 IMF 分量并利用皮尔逊相关系数做进一步筛选。其次,对筛选得到的 IMF 分量进行特征提取,分别提取各分量的峭度、近似熵及模糊熵。最后,将上述特征值构成的特征向量输入经 SO 优化的 SVM 中进行训练及分类,得到故障诊断结果。实验结果表明,采用本文提出的基于 VMD 和 SO 优化 SVM 的故障识别方法,光纤复合海底电缆的故障识别准确率达到 100%,分别比 SVM、GA-SVM、GWO-SVM、CNN 方法的识别准确度高 7.5%、5%、5%、7.5%。

**关键词:** 变分模态分解;蛇优化算法;支持向量机;IMF 分量;光纤复合海缆故障

**中图分类号:** TP277 **文献标识码:** A **国家标准学科分类代码:** 470.4

## Fault diagnosis of fiber-optic composite submarine cable based on VMD and SO optimized SVM

Li Junqing<sup>1</sup> Liu Ruoyao<sup>1</sup> He Yuling<sup>2</sup> Zhang Chengzhi<sup>1</sup> Geng Jiya<sup>1</sup>

(1. Department of Electric Power Engineering, North China Electric Power University, Baoding 071003, China;

2. Department of Mechanical Engineering, North China Electric Power University, Baoding 071003, China)

**Abstract:** In order to further improve the fault diagnosis accuracy of fiber optic composite submarine cable, a fault diagnosis method based on VMD and SO optimization SVM is proposed. Firstly, VMD was used to decompose the fault data, several IMF components were obtained, and Pearson correlation coefficient was used for further screening. Secondly, feature extraction is carried out on the selected IMF components to extract the kurtosis, approximate entropy and fuzzy entropy of each component respectively. Finally, the eigenvectors composed of the above eigenvalues are input into the SVM optimized by SO for training and classification, and the fault diagnosis results are obtained. The experimental results show that the fault recognition accuracy of fiber-optic composite submarine cable can reach 100% by using the optimized SVM method based on VMD and SO, which is 7.5%, 5%, 5% and 7.5% higher than that of SVM, GA-SVM, GGO-SVM and CNN respectively.

**Keywords:** variational mode decomposition; snake optimization algorithm; support vector machine; IMF component; the faulty of fiber-optic composite submarine cable

## 0 引言

在大力推进经济全球化和建设“海洋强国”的背景下,海上风电进入了蓬勃发展的时期<sup>[1]</sup>。光纤复合海底电缆(以下简称“海缆”)集跨海通信和电能传输于一身,已成为经济全球化和海上风电场并网的重要媒介。海缆常年工作于恶劣的海底环境,易受到船舶抛锚、渔船捕鱼、地质活动<sup>[2]</sup>及海水冲刷的影响,从而造成海缆发生短路、漏电等电气故障及锚砸、钩挂等机械故障。海缆由于其每公里高达

百万元以上的造价,一旦发生故障,不仅给国民经济造成巨大损失,还会严重影响海边居民的正常生产生活<sup>[3]</sup>。因此,对海缆故障进行快速和准确识别,具有重要的研究意义。

海缆的故障诊断问题可归纳为多分类问题,其故障数据主要通过分布式光纤传感技术获得。分布式光纤传感技术以海缆中的冗余光纤作为传感器,通过相位敏感光时域反射、布里渊光时域反射等方法提取光纤中的温度和应变信号。根据海缆发生短路及漏电故障时温度变化较大,发生锚砸及钩挂故障时应变较大的特点<sup>[4]</sup>,对温度及应变信

收稿日期:2023-04-23

\* 基金项目:国家自然科学基金(52177042)项目资助

号的特征提取是进行故障诊断的关键一环。

施鑫麒等<sup>[5]</sup>利用db小波包对海缆数据进行特征提取,经过重构后可还原故障特征,但该方法存在分解层数过高时重构精度不足的缺陷。郑国强采用极点对称模态分解(extreme-point symmetric mode decomposition, ESMD)算法将海缆故障信号分解为有限个本征模态函数(intrinsic mode function, IMF)分量,利用支持向量机(support vector machine, SVM)进行海缆故障识别<sup>[6]</sup>,但ESMD的信号分解效果易受到噪声干扰且不适合处理非线性海缆故障信号,另外,未经优化的SVM存在惩罚参数和核参数难以确定最优值的缺点。为此,尚秋峰等<sup>[7]</sup>利用鸟群优化算法(bird swarm algorithm, BSA)优化后的SVM对海缆故障进行分类,虽然在一定程度上解决了SVM的参数寻优问题,但该算法的收敛速度仍有一定的提升空间。

相比以上方法,变分模态分解(variational mode decomposition, VMD)通过迭代搜索模态最优解,具有更好的抗噪性能,能够对含噪的海缆故障信号进行有效的特征提取,并且避免了过包络、欠包络及模态频率混叠问题<sup>[8]</sup>。蛇优化算法(snake optimization, SO)<sup>[9]</sup>是于2022年提出的一种新型智能优化算法,具有收敛速度快、性能稳定、鲁棒性好、结果准确的优点。

为了解决上述问题,本文采用VMD和SO优化SVM相结合的方法,对海缆短路、漏电、锚砸及钩挂4种类型的故障进行识别。本文通过修改海缆正常监测数据得到各故障状态下的温度或应变信号,信号经VMD分解得到有限个IMF分量。提取各IMF分量的峭度、近似熵和模糊熵组合为特征向量,输入经SO优化的SVM中,输出分类结果。将本文方法与遗传算法优化支持向量机(genetic algorithm-SVM, GA-SVM)、灰狼算法优化支持向量机(grey wolf optimization-SVM, GWO-SVM)、SVM、卷积神经网络CNN方法进行了对比分析,结果表明VMD和SO优化SVM相结合的方法具有更好的识别效果。

## 1 基于VMD和SO优化SVM的故障诊断方法

### 1.1 变分模态分解

VMD算法将经典的维纳滤波器推广到多个自适应波段,可根据各模态的中心频率将一组信号分解为多个IMF分量,并且保证每个模态在解调到基带后都是平滑的,同时它还克服了传统的经验模态分解算法存在的模态混叠的缺陷。VMD算法将原始故障数据分解为IMF分量的步骤如下:

1) 建立约束变分问题。

$$\begin{cases} \min_{\{p_k\}, \{\omega_k\}} \sum_{k=1}^K \|\partial(t) \left[ \delta(t) + \frac{j}{\pi t} \times p_k(t) \right] \times e^{-j\omega_k t}\|^2 \\ s. t. \sum_{k=1}^K p_k = f(t) \end{cases} \quad (1)$$

式中: $p_k$ 为原始信号经VMD分解后得到的各IMF分量; $K$ 为IMF分量的个数; $f(t)$ 为待分解的原始信号,即短路、漏电、锚砸、钩挂4种类型的故障信号; $\delta(t)$ 为狄拉克分布; $t$ 为采样时刻; $\omega_k$ 为各IMF分量对应的频率中心。

2) 求解约束变分问题的最优解。求解的具体步骤如下:

步骤(1)引入拉格朗日乘法因子,进行拉格朗日变换,将约束变分问题转换为非约束变分问题。

$$L(\{p_k\}, \{\omega_k\}, \lambda) = \alpha \sum_{k=1}^K \|\partial_i \left[ \delta(t) + \frac{j}{\pi t} \times p_k(t) \right] e^{-j\omega_k t}\|^2 + \|f(t) - \sum p_k(t)\|^2 + \langle \lambda(t), f(t) - \sum p_k(t) \rangle \quad (2)$$

式中: $\lambda$ 为拉格朗日乘法因子; $\alpha$ 为二阶惩罚因子,用于保证高斯噪声存在时信号重构的准确性。

步骤(2)采用交替方向乘法更新拉格朗日乘法因子、各IMF分量及其中心频率,当满足步骤(3)中的迭代终止条件时就得到了无约束模型的鞍点,即原问题的最优解 $p_k$ 。

$$\hat{p}_k^{n+1}(\omega) = \frac{\hat{f}(\omega) - \sum_{i \neq k} \hat{p}_i(\omega) + \frac{\hat{\lambda}(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k)^2} \quad (3)$$

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |\hat{p}_k(\omega)|^2 d\omega}{\int_0^\infty |\hat{p}_k(\omega)|^2 d\omega} \quad (4)$$

$$\hat{\lambda}^{n+1}(\omega) = \hat{\lambda}^n(\omega) + \gamma(\hat{f}(\omega) - \sum_k \hat{p}_k + 1_k(\omega)) \quad (5)$$

式中: $n$ 为迭代次数; $|\hat{p}_k(\omega)|$ 为各IMF分量经傅里叶变换后的维纳滤波; $\gamma$ 为噪声容限。

步骤(3)当满足下式时,迭代终止,经傅里叶变换后,输出得到的 $K$ 个IMF分量的值。

$$\frac{\sum_{i=1}^k \|\hat{p}_k^{n+1}(t) - \hat{p}_k^n(t)\|^2}{\|\hat{p}_k^n(t)\|^2} < \epsilon \quad (6)$$

式中: $\epsilon$ 为设定的阈值。

### 1.2 蛇优化算法

现实中,蛇只会在温度较低且有食物时进行交配,否则蛇只会寻找食物或吃掉现有的食物。蛇优化算法是一种新型的仿生智能优化算法,该算法通过设置一定的温度范围和食物数量来激发蛇的一系列行为,包括搜索食物、战斗模式和交配模式。根据蛇的上述行为不断更新蛇群中个体的位置,最终得到的最优个体的位置即为最优解,其优化过程如下:

1) 搜索食物,具体流程为:

将蛇的种群划分为雄性和雌性两个子种群,且两个子种群的个体数量保持相同。当食物数量小于阈值,即食物

匮乏时,蛇会通过选择任意随机位置来搜索食物,并根据以下公式更新它们的位置:

$$X_{i,m}(t+1) = X_{\text{rand},m}(t) \pm c_1 \times A_m \times ((X_{\text{max}} - X_{\text{min}}) \times \text{rand} + X_{\text{min}}) \quad (7)$$

$$X_{i,f}(t+1) = X_{\text{rand},f}(t) \pm c_1 \times A_f \times ((X_{\text{max}} - X_{\text{min}}) \times \text{rand} + X_{\text{min}}) \quad (8)$$

式中: $X_{i,m}$  和  $X_{i,f}$  分别为第  $i$  只雄性蛇和雌性蛇的位置; $X_{\text{rand},m}$  和  $X_{\text{rand},f}$  分别为随机选择的雄性蛇和雌性蛇的位置; $X_{\text{max}}$  和  $X_{\text{min}}$  分别为优化问题取值范围的上下边界; $A_m$  和  $A_f$  分别为雄性蛇和雌性蛇寻找食物的能力; $c_1$  为常数,经反复试验,取值 0.05 时效果最优;rand 为  $0 \sim 1$  的随机数。

当食物数量高于阈值,温度高于阈值时,蛇同样只会搜索食物,其位置通过下式更新:

$$X_{i,j}(t+1) = X_{\text{food}} \pm c_2 \times \text{Temp} \times \text{rand} \times (X_{\text{food}} - X_{i,j}(t)) \quad (9)$$

式中: $X_{i,j}$  为雄性或雌性蛇的位置; $X_{\text{food}}$  为最佳个体的位置; $c_2$  为常数,经反复试验,取值为 2 时效果最优;Temp 为环境的温度值。

2) 进入战斗模式,具体流程为:

当食物数量高于阈值,而温度低于阈值时,雄性蛇之间会因争夺最优秀的雌性而发生战斗,同样的,雌性蛇也会战斗以获得最好的雄性。战斗模式中,蛇的位置利用下式进行更新:蛇会进入战斗模式或交配模式。

$$X_{i,m}(t+1) = X_{i,m}(t) + c_2 \times FM \times \text{rand} \times (Q \times X_{\text{best},f} - X_{i,m}(t)) \quad (10)$$

$$X_{i,f}(t+1) = X_{i,f}(t) + c_2 \times FF \times \text{rand} \times (Q \times X_{\text{best},m} - X_{i,f}(t)) \quad (11)$$

式中: $X_{\text{best},f}$  和  $X_{\text{best},m}$  分别为雌性和雄性个体的最佳位置; $FM$ 、 $FF$  分别为雄性和雌性蛇的战斗能力; $Q$  为存在的食物数量。

3) 进入交配模式,具体流程为:

当食物充足,环境寒冷时,雄性蛇与雌性蛇之间还会发生交配行为。交配模式下,两者位置采用下式更新:

$$X_{i,m}(t+1) = X_{i,m}(t) + c_2 \times M_m \times \text{rand} \times (Q \times X_{i,f}(t) - X_{i,m}(t)) \quad (12)$$

$$X_{i,f}(t+1) = X_{i,f}(t) + c_2 \times M_f \times \text{rand} \times (Q \times X_{i,m}(t) - X_{i,f}(t)) \quad (13)$$

式中: $M_m$  和  $M_f$  分别为雄性和雌性蛇的交配能力。

交配模式下会发生产卵现象,卵孵化后产生的小蛇会替换种群中最差的雄性和雌性个体。蛇群中最差的雄性和雌性个体的计算按下式进行:

$$X_{\text{worst},m} = X_{\text{min},m} + \text{rand} \times (X_{\text{max},m} - X_{\text{min},m}) \quad (14)$$

$$X_{\text{worst},f} = X_{\text{min},f} + \text{rand} \times (X_{\text{max},f} - X_{\text{min},f}) \quad (15)$$

式中: $X_{\text{worst},m}$ 、 $X_{\text{worst},f}$  分别为雄性和雌性中的最差个体; $X_{\text{min},m}$ 、 $X_{\text{max},m}$  分别为雄性个体活动范围的下边界和上边界; $X_{\text{min},f}$ 、 $X_{\text{max},f}$  分别为雌性个体活动范围的下边界和上

边界。

### 1.3 SO 优化 SVM

SVM 是一种用于智能分类的深度学习算法,主要应用于样本数量较少情况下的模式识别问题<sup>[10-11]</sup>。SVM 通过寻找最优分类超平面来达到分类目的。寻找最优分类超平面可转化为求解以下带约束条件的目标函数:

$$\begin{cases} \min \frac{1}{2} \|\omega\|^2 + c \sum_{i=1}^n \xi_i \\ \text{s. t. } \begin{cases} y_i(\omega x_i + b) \geq 1 \\ c \geq 0 \end{cases} \end{cases} \quad (16)$$

式中: $\omega$  为权重; $\xi_i$  为松弛因子; $b$  为偏置; $c$  为惩罚参数,用于平衡最大间隔和分类误差, $c$  值越大 SVM 的分类准确率越高,但  $c$  值过大会导致模型的训练时间延长与过拟合等问题。

SVM 采用核函数(kernel function, KF)进行原始数据的非线性映射,以在高维空间解决多分类问题。本文采用径向基函数(radial basis function, RBF)作为核函数,其定义为:

$$K(\mathbf{x}_i, \mathbf{x}) = \exp(-\|\mathbf{x}_i - \mathbf{x}\|^2 / 2g^2) \quad (17)$$

式中: $\mathbf{x}_i$  和  $\mathbf{x}$  为输入所在空间的任意两个特征向量; $g$  为核参数,直接影响空间维度的大小。

本文采用 SO 算法对 SVM 的惩罚参数  $c$  和核参数  $g$  进行迭代寻优,具体步骤如下:

1) 初始化 SO 参数。设置种群数量  $N$ 、迭代次数  $D$ 、常数  $c_1$ 、 $c_2$  的值。之后,设定蛇种群个体活动的区域,即参数  $c$  和  $g$  的取值范围。

2) 数据归一化。为避免原始数据量纲不同带来的影响,采用 Z-score 归一化方法,使数据满足标准正态分布,其公式如下:

$$x^* = \frac{x(t) - \mu}{\sigma} \quad (18)$$

式中: $x(t)$  为原始故障数据; $\mu$  为原始故障数据的均值; $\sigma$  为原始故障数据的标准差。

3) 将 SVM 分类精度作为适应度函数,计算初始种群的适应度值。适应度函数可表示为:

$$v = \frac{x_T}{x_T + x_F} \quad (19)$$

式中: $x_T$  为分类正确的个数; $x_F$  为分类错误的个数; $v$  为适应度值。

4) 根据搜索食物、战斗模式、交配模式相应的公式,更新蛇种群最优个体的位置。最优个体的位置表示为一个包含两参数的向量,分别对应核参数  $g$  和惩罚参数  $c$ 。

5) 将更新后的  $g$  和  $c$  输入 SVM,计算适应度值;

6) 判断是否满足迭代终止条件。当适应度函数的值达到最大或不再增加时,迭代终止,此时的  $g$  和  $c$  即为 SVM 的最优参数。否则,返回步骤 4) 继续循环。

本文提出的 SO 优化 SVM 的算法流程如图 1 所示。

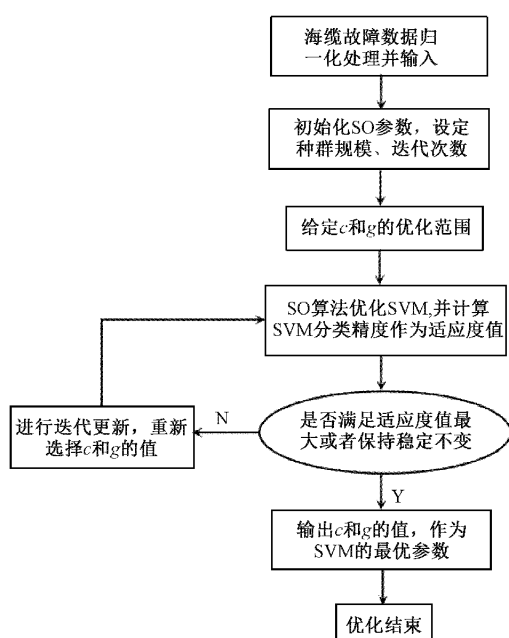


图1 SO优化SVM的流程

#### 1.4 故障诊断方法及流程

将VMD与SO优化SVM相结合,针对海缆故障数据的诊断具体步骤如下。

##### 1) 数据获取

依据经验公式及长期运行经验,修改海缆正常状态下的监测数据,得到短路、漏电、锚砸及钩挂4种故障类型的数据。利用ANSYS Workbench软件进行仿真建模,验证故障数据的合理性。

##### 2) 数据处理

对故障数据进行VMD分解,得到有限个IMF分量并利用皮尔逊相关系数法进行筛选。

##### 3) 特征提取

分别提取筛选后的分量的峭度、近似熵、模糊熵作为特征值,进而组合成特征向量。

##### 4) 故障诊断

将上述特征向量的80%作为训练集,20%作为测试集,输入经SO优化后的SVM中进行训练及分类,输出诊断结果。

本文所述故障诊断方法的流程如图2所示。

## 2 实验验证

### 2.1 海缆故障数据获取

海缆发生的故障可分为机械故障和电气故障。机械故障主要是指以应变为主的锚砸、钩挂,电气故障主要包括以温度变化为主的短路、漏电。分布式光纤传感技术以冗余光纤作为传感器,可以探测光纤中光的偏振态变化,经光干涉仪解调后可进一步转化为光强度的变化<sup>[21]</sup>。布里渊光时域反射技术作为分布式光纤传感技术的重要一员,可精

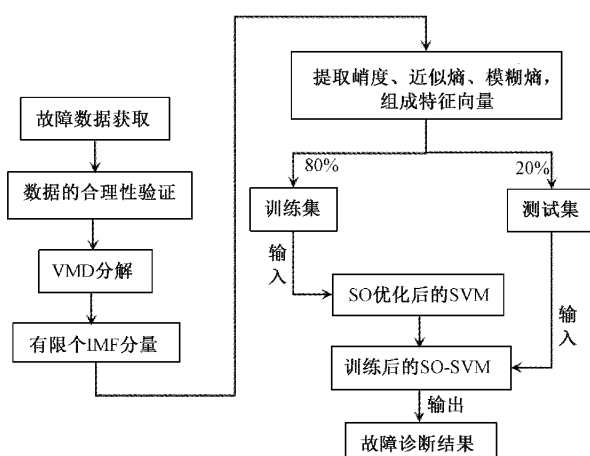


图2 海缆故障诊断流程图

确测量海缆的布里渊频移信号,再利用式(20)将其还原为温度及应变信号。

$$\Delta f_B = C_T \Delta T + C_\epsilon \Delta \epsilon \quad (20)$$

式中: $\Delta T$ 和 $\Delta \epsilon$ 分别为光纤的温度变化量和应变量; $C_T$ 和 $C_\epsilon$ 分别为布里渊频移的温度系数和应变系数; $\Delta f_B$ 为布里渊频移变化量。

海缆由于其复杂的工作环境和高昂的制作成本,致使故障实验难以开展。故障数据的贫乏是阻碍深度学习应用于海缆故障诊断领域的一大难题。本文选用位于福建省某海岸、长度为3.5 km的110 kV光纤复合海底电缆,依据式(21)~(22)及长期运行经验修改其正常监测时获得的温度及应变信号得到故障数据,用于海缆故障诊断研究。

其中,海缆发生锚砸、钩挂时,应变幅值与船锚重量、船锚速度、海水流速等有关,依据运行经验修改即可;当海缆的敷设深度和土壤的导热系数不变时,海缆发生短路或漏电后的温升只与海缆的载流量和环境温度有关<sup>[13]</sup>,其温升满足下式,总的温升为两部分温升之和。

$$\Delta T_1 = \rho I^2 \quad (21)$$

式中: $\Delta T$ 为温度的变化量; $I$ 为海缆的载流量,单位为A; $\rho$ 为随温度而变化的系数。

$$\Delta T_2 = \rho_1 T + \rho_2 \quad (22)$$

式中: $T$ 为环境温度,单位为 $^{\circ}\text{C}$ ; $\rho_1$ 、 $\rho_2$ 是随载流量变化的系数。

文献[5]指出,海缆发生短路故障时,光纤温度发生改变的长度范围为10~30 m,温度上升区域位于故障点到电源之间;发生漏电故障时,光纤温度上升的区域仅为故障点周围几米内;发生锚砸故障时,光纤发生应变的长度范围为数十米,且应变呈现由故障点向两端逐渐减小的趋势;发生钩挂故障时,光纤的应变长度范围可达到数百米,变化趋势与锚砸相同。

据此,通过设置不同的故障长度范围,不同的故障起始位置,不同的温度及应变变化幅度,不同的温度及应变初始值,依据长期运行经验及上述公式修改海缆正常监测数据,

得到合理的海缆故障数据。另外,为模拟海缆实际工况,向故障数据中添加不同程度的高斯噪声。最终形成包含短路、漏电、锚砸及钩挂 4 种故障类型的 200 组数据的数据集,其中,每种故障类型对应 50 组数据。每种故障类型任意选取一组数据,如图 3~6 所示。

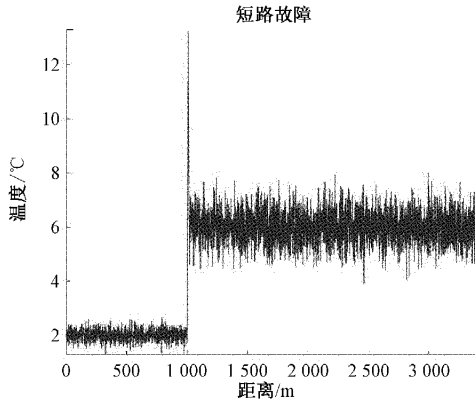


图 3 短路故障数据

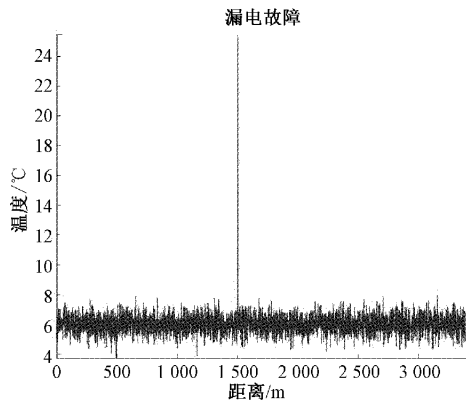


图 4 漏电故障数据

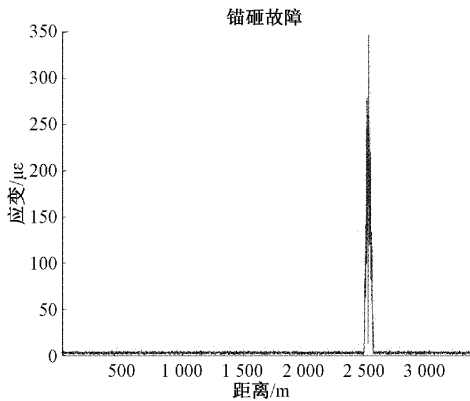


图 5 锚砸故障数据

其中,图 3 为海缆发生短路故障,故障范围 20 m,故障起始位置为 1 km 处,温度上升 10 °C,高斯噪声标准差为 0.1;图 4 为海缆发生漏电故障,故障范围 5 m,故障起始位置为 1.5 km 处,温度上升 18 °C,高斯噪声标准差为 0.1;图 5 为海缆发生锚砸故障,故障范围 80 m,故障起始位置

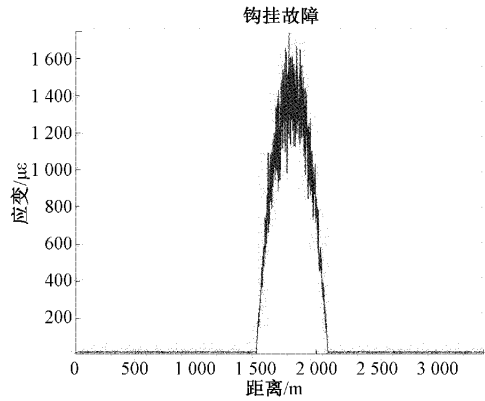


图 6 钩挂故障数据

为 2.5 km 处,应变为 240  $\mu\epsilon$ ,高斯噪声标准差为 0.3;图 6 为海缆发生钩挂故障,故障范围 600 m,故障起始位置为 1.5 km 处,应变为 1 400  $\mu\epsilon$ ,高斯噪声标准差为 0.2。

为检验所用锚砸及钩挂故障数据的合理性,选用 YJQ41 型单芯交联聚乙烯 (crosslinked polyethylene, XLPE) 光纤复合海底电缆,利用 ANSYS Workbench 仿真软件对该海缆受到锚砸和钩挂作用时的工况进行仿真。考虑到计算资源的有限性,在保证求解精度的前提下对海缆的结构进行一定的简化,简化后的海缆结构如图 7 所示,其中,光纤存在于光单元结构中。图 8~9 分别为海缆受到锚砸和钩挂作用时的网格划分图。

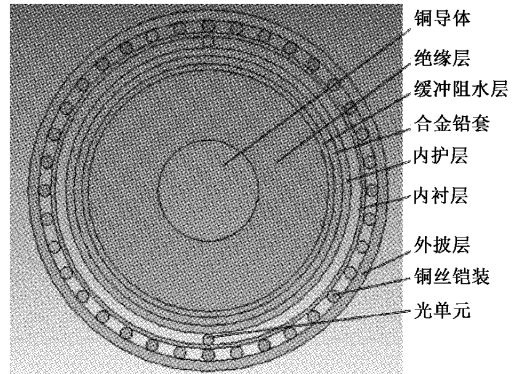


图 7 简化后的海缆截面图

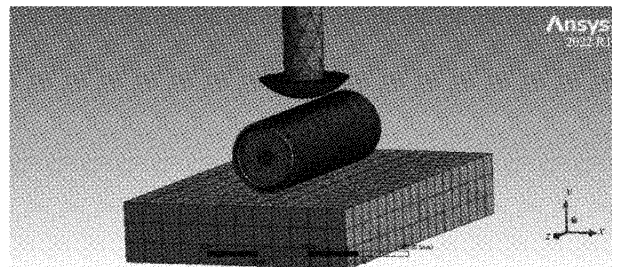


图 8 锚砸作用下的网格划分效果图

设置海缆长度为 1.5 m,船锚质量为 200 kg,船锚砸中海缆时的速度为 5 m/s,船锚钩挂海缆时的速度为 2 m/s。海缆各层结构、船锚、地面选择相应的材料,篇幅所限,此处

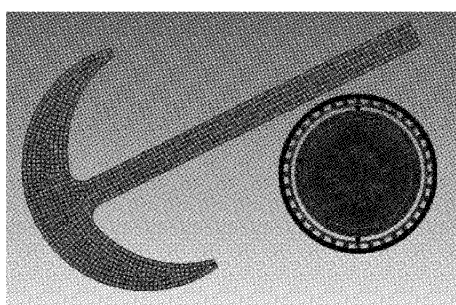


图9 钩挂作用下的网格划分效果图

不再赘述。提取锚砸及钩挂作用下,海缆光单元的应变仿真结果,如图10~11所示。

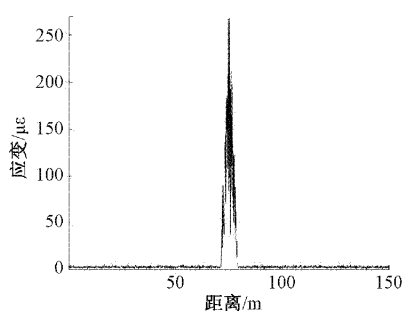


图10 锚砸时光单元应变仿真结果

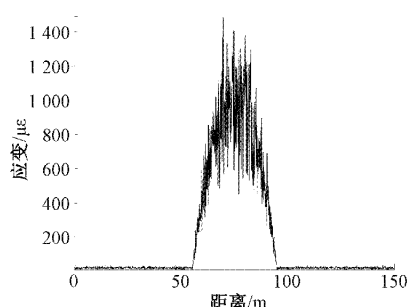


图11 钩挂时光单元应变仿真结果

类比上述光单元应变的仿真结果与修改得来的故障数据,可见两者具有一致性,说明本文采用的修改海缆正常运行数据得到故障数据的做法具有合理可行性。

## 2.2 故障数据VMD分解

设置VMD的各个参数:二阶惩罚因子 $\alpha$ 设为6400,模态分量分解个数 $K$ 设为6,迭代次数 $n$ 设为100。利用VMD对含有不同程度高斯噪声的故障数据进行分解,得到6个IMF分量。原始信号的采样频率为12kHz,采样点数为3400。以短路故障为例,其中一组故障数据经VMD分解后的IMF分量如图12所示。图13为图12中各IMF分量对应的频谱图,可以看出各IMF分量按低频到高频的顺序排列。

## 2.3 筛选有效分量

利用皮尔逊相关系数法对所得的6个IMF分量进行筛选,相关系数的计算公式为式(23),各IMF分量对应的

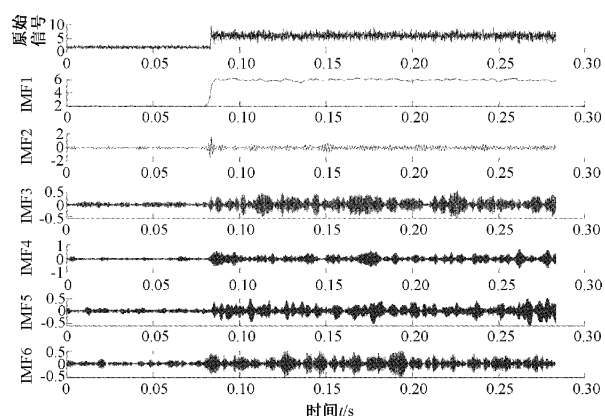


图12 VMD分解

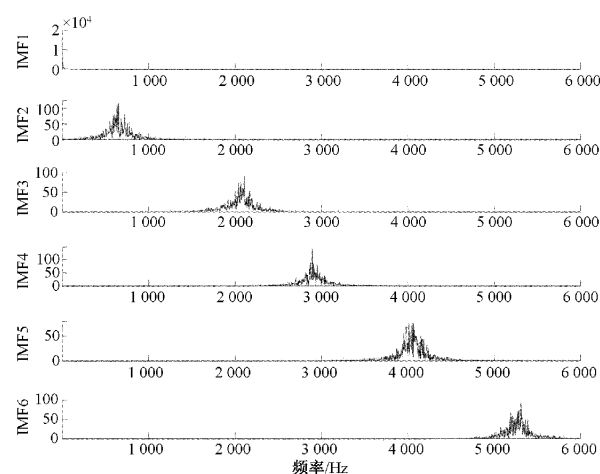


图13 频谱图

相关系数如表1所示。

$$R = \frac{E[(p_k - \bar{p}_k)(f - \bar{f})]}{E[(\mu_k - \bar{\mu}_k)^2] E[(f - \bar{f})^2]} \quad (23)$$

式中: $f$ 和 $p_k$ 分别为原始故障数据和VMD分解得到的IMF分量; $\bar{f}$ 和 $\bar{p}_k$ 分别为其均值; $E$ 为数学期望。

表1 IMF分量与原始数据的相关系数

故障类型	IMF1	IMF2	IMF4	IMF5	IMF6
短路	0.854 6	0.448 4	0.084 5	0.584 0	0.078 5
漏电	0.426 1	0.578 1	0.484 3	0.107 3	0.103 0
锚砸	0.553 0	0.580 6	0.325 5	0.205 9	0.304 1
钩挂	0.957 6	0.430 7	0.080 0	0.316 2	0.323 6

当相关系数大于0.3时,分量与原始信号之间才具有相关性<sup>[14]</sup>,以短路故障为例,IMF1、IMF2、IMF3、IMF5为经筛选得到的有效分量。

## 2.4 IMF分量特征提取

对4种故障类型的各有效分量进行特征提取,分别提取其峭度(kurtosis,  $K$ )、近似熵(approximate entropy, AE)<sup>[15]</sup>和模糊熵(fuzzy entropy, FE)<sup>[16]</sup>作为特征值,并组

合成为特征向量,计算 K 的公式如式(24)所示,计算 AE 和 FE 的过程参看文献[15]及文献[16]。峭度可反映故障数据中冲击信号的影响并随着冲击信号的增大而增大<sup>[17]</sup>,近似熵和模糊熵用来衡量信号的复杂程度,熵值越小,说明信号中存在的频率个数越少,噪声含量越低,故障特征越明显。

$$K = \frac{E[x(t) - \mu]^4}{\sigma^4} \quad (24)$$

上述特征向量的作用是在利用支持向量机进行故障诊断时作为其输入<sup>[18]</sup>。经计算,海缆各类型故障数据的特征向量如表 2 所示,表 2 中的 IMFA~IMFD 为经筛选得到的各故障类型有效分量,每种故障类型由 4 个有效分量表示,每个有效分量经计算得到 3 个特征值,故每个特征向量由 12 个特征值组成。

表 2 故障数据的特征向量

	特征值	短路	漏电	锚砸	钩挂
IMFA	K	1.828 4	22.875 3	54.705 0	22.113 3
	AE	0.002 9	0.213 8	0.010 2	0.013 8
	FE	0.000 3	0.039 3	0.009 2	0.002 3
IMFB	K	11.018 3	39.978 2	89.417 5	52.013 0
	AE	0.602 3	0.575 1	0.016 6	0.069 7
	FE	0.400 0	0.538 1	0.015 9	0.038 3
IMFC	K	3.413 5	28.835 5	132.150 0	59.581 9
	AE	0.416 9	0.395 1	0.105 4	0.123 2
	FE	1.181 5	1.157 1	0.487 9	0.165 9
IMFD	K	3.801 3	20.416 8	148.388 8	50.793 6
	AE	0.408 8	0.348 8	0.300 7	0.045 6
	FE	1.489 3	1.495 8	0.247 5	0.061 2

### 2.5 故障诊断

设置 SO 算法的种群数量  $N$  为 30,迭代次数  $D$  为 100,核参数  $g$  和惩罚参数  $c$  的取值范围均为 0.01~1 000。以 SVM 的分类精度作为适应度值,SO 优化 SVM 的优化曲线如图 14 所示,可以看出,当迭代次数仅为 15 次时适应度值就已达到了最大而后保持稳定,说明 SO 算法的优化过程迅速,收敛速度快。优化后,得到  $g$  和  $c$  的最优值分别为 0.114 6 和 0.506 2,将该值作为 SVM 的最优参数即完成了 SVM 的优化过程。

本文共有 4 种故障类型,通过 2.2~2.4 节得到数据样本,每种故障类型包含 50 组特征向量,将其 80% 作为训练集,20% 作为测试集,最终得到 160 个样本的训练集和 40 个样本的测试集。给定各故障类型对应的标签:短路故障对应标签为 1;漏电故障对应标签为 2;锚砸故障对应标签为 3;钩挂故障对应标签为 4。利用训练集训练经 SO 优化后的 SVM,再将测试集输入得到识别结果。图 15 为经 SO 优化后的 SVM 在测试集上的识别结果,分类准确率达到

了 100%,识别精度良好。

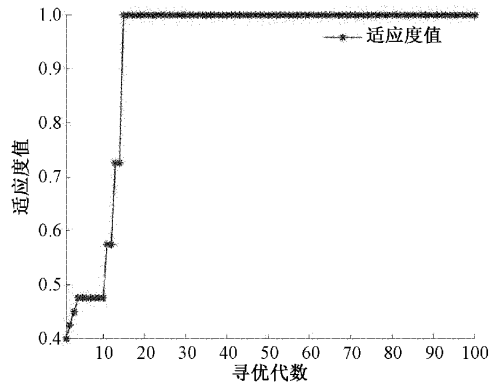


图 14 SO 优化 SVM 的适应度曲线

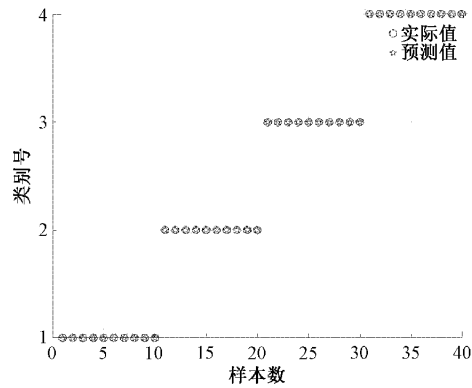


图 15 SO 优化 SVM 的识别结果

为进一步验证 VMD 结合 SO 优化 SVM 的故障诊断方法在真实海缆故障数据上的分类效果,选取分布于我国沿海的某条光纤复合海缆,收集其故障数据进行实验验证。其中,海缆锚砸和钩挂故障数据各有 20 组,短路和漏电故障数据各有 5 组,原因在于锚害引起的机械类事故远远多于海缆自身结构损坏引起的电气类事故。分别给定锚砸、钩挂、短路、漏电数据的标签为 1、2、3、4,利用本文提出的方法对上述数据进行分类,分类结果如图 16 所示。可知:VMD 结合 SO 优化 SVM 的故障诊断方法在真实海缆故障数据上的整体识别准确率为 98% (49/50),其中,锚砸、短

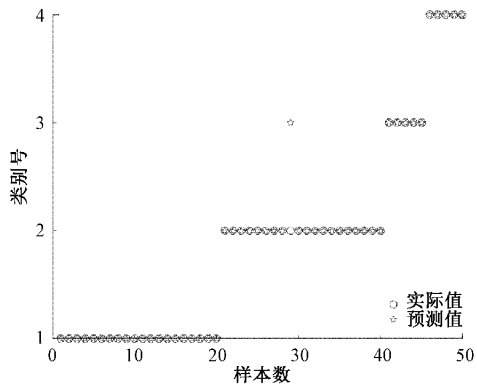


图 16 海缆真实故障数据的识别结果

路、漏电的识别准确率均达到了100%，只有一组钩挂类型的故障数据错判为了短路故障，说明了本文所提方法在现实中应用的可行性。

### 3 对比实验

为了验证本文提出的SO优化SVM算法在提高海缆故障诊断准确率方面的优越性，将经VMD分解后得到的特征向量分别输入SVM、GA-SVM、GWO-SVM、CNN中进行训练和测试，在测试集上的识别结果如图17~图20所示。由图17~20的结果得知：未经优化的SVM及CNN的故障诊断准确率最低，为92.5% (37/40)；经GA或

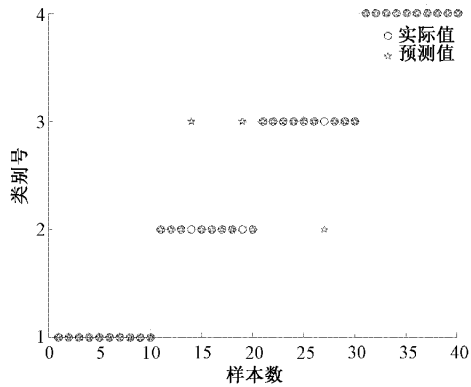


图17 SVM的识别结果

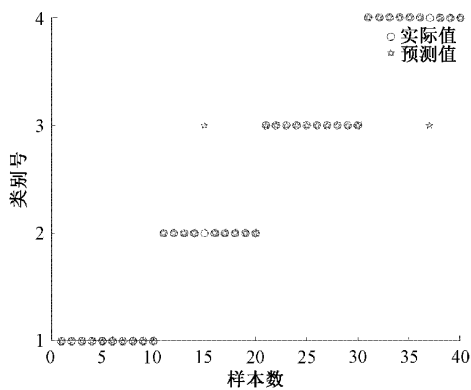


图18 GA优化SVM的识别结果

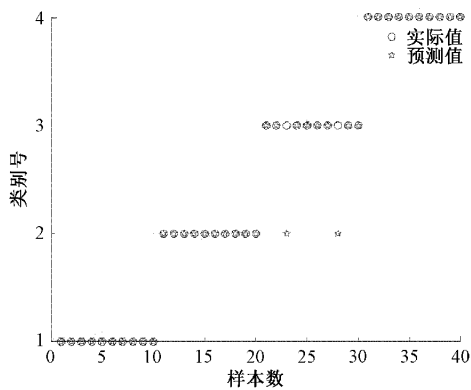


图19 GWO优化SVM的识别结果

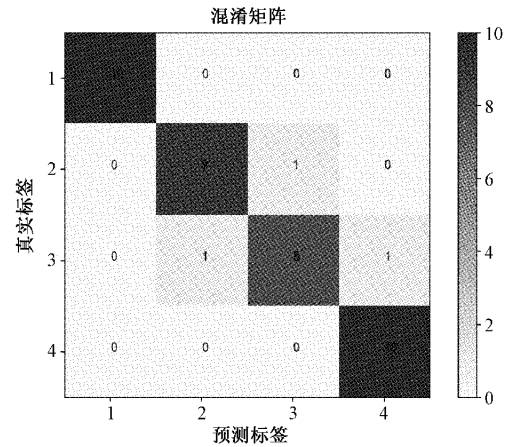


图20 CNN的识别结果

GWO优化后的SVM的故障诊断准确率有所提高，均为95% (38/40)。上述诊断准确率均低于经SO优化的SVM，说明了本文所提方法具有较高的精度。

### 4 结论

本文研究了基于VMD及SO优化SVM的故障诊断方法，并将其应用于光纤复合海底电缆的故障诊断中，取得了良好的识别效果。考虑到海缆故障数据的贫乏，依据经验修正正常监测数据得到合理的故障数据，并利用有限元模型进行故障数据的准确性验证。通过VMD算法进行信号分解，提取经筛选后的分量的峭度、近似熵、模糊熵组合成为特征向量，输入到经SO优化后的SVM中进行训练及分类，得到了在测试集上100%的准确率。对比实验结果表明，SO优化SVM算法的故障诊断准确率高于SVM、GA优化SVM、GWO优化SVM及CNN，能对海缆的四种故障类型进行准确识别。另外，本文方法在识别真实海缆故障数据中识别精度达到了98%，具有应用价值。

### 参考文献

- [1] 唐文虎,梁启恒,赵柏宁,等.基于改进稀疏表示的海上风电场交流海底电缆短路故障分类方法[J].中国电机工程学报,2023,43(06):2212-2222.
- [2] DIMITRIOS E, RAUL V. Sensors to increase the security of underwater communication cables: A review of underwater monitoring sensors[J]. Sensors, 2020,20(3):173-179.
- [3] 冀大雄,周佳龙,钱建华,等.海底电缆检测方法发展现状综述[J].南方电网技术,2021,15(5):36-49.
- [4] 陈永,尹成群,吕安强,等.光纤复合海底电缆故障检测与诊断方法[J].光通信研究,2014,18(3):56-59.
- [5] 施鑫麒,陈元林,安博文.基于MapWinGIS的海缆故障数据全方位可视化仿真[J].计算机仿真,2021,38(8):264-268.
- [6] 郑国强.基于ESMD和SVM的光纤复合海缆振动信



- 号分析方法研究[D]. 北京:华北电力大学,2021.
- [7] 尚秋峰,郭家兴. 基于 VMD-BSA-SVM 的海缆振动信号识别方法[J]. 激光与光电子学进展,2022,59(17): 65-74.
- [8] 郭梓良,郝如江,杨文哲,等. 优化 VMD 和改进残差网络的齿轮箱故障诊断研究[J]. 机械科学与技术,2022, 41(1):1-7.
- [9] HASHIM F A, HUSSIEN A G. Snake optimizer: A novel meta-heuristic optimization algorithm [J]. Knowledge-Based Systems,2022,17(3):242-248.
- [10] ZHANG A H, YU D L, ZHANG Z Q. TLSCA-SVM fault diagnosis optimization method based on transfer learning[J]. Processes, 2022,10(2):362-362.
- [11] 李红月,高英杰,朱文昌. IAO 优化 SVM 的电机滚动轴承故障诊断[J]. 电子测量技术,2022,45(10): 126-132.
- [12] 周路遥,刘黎,蒋愉宽,等. 光纤分布式传感技术在海底电缆状态监测中的应用[J]. 浙江电力,2018,37(2): 6-10.
- [13] 胡志伟. 海缆电气和机械故障对光纤温度/应变影响的建模分析和特征提取[D]. 北京:华北电力大学,2019.
- [14] 杨东,宋汉江,霍柏琦,等. 基于特征相关分析的柴油机故障诊断方法研究[J]. 现代制造工程,2018,23(12): 147-152,5.
- [15] RICHMAN J S, MOORMAN J R. Physiological time-series analysis using approximate entropy and sample entropy[J]. American Journal of Physiology-Heart and Circulatory Physiology, 2012,278(6):2039-2049.
- [16] 李生鹏,韦朋余,丁峰,等. 基于多尺度模糊熵和主成分分析的轴承故障特征提取[J]. 船舶力学,2018, 22(10):1277-1285.
- [17] 贺志军,李军霞,刘少伟,等. CEEMD-VMD 与参数优化 SVM 结合的托辊轴承故障诊断[J]. 机械科学与技术,2022,42(2):1-7.
- [18] 刘秀丽,王鸽,吴国新,等. VMD 及 PSO 优化 SVM 的行星齿轮箱故障诊断[J]. 电子测量与仪器学报,2022, 36(1):54-61.

### 作者简介

**李俊卿**,教授,博士,主要研究方向为电机及其系统分析,电气设备状态监测与故障诊断。

E-mail:junqing03@163.com

**刘若尧**,硕士,主要研究方向为深度学习及故障诊断。

E-mail:694986539@qq.com

**何玉灵**,教授,博士,主要研究方向为电机特性分析、检测与失效预防。

E-mail:heyuling1@163.com