

基于支持向量机的 XLPE 电缆接头局部放电的模式识别

刘永生

(南京电力工程设计有限公司, 江苏 南京 210009)

摘 要: 根据 XLPE 电缆接头局部放电的特点, 提出了基于支持向量机 (以偏斜度、陡峭度、放电量因数、相位不对称度、互相关系数、修正的互相关系数这六个统计算子为特征量) 的 XLPE 电缆接头局部放电的模式识别方法, 本文详细分析了基于向量机的基本原理, 阐述了特征量的提取过程和局部放电三维图谱的构造, 并将该方法应用于模拟实验设计的 4 种典型局部放电模型, 效果显著, 能有效地排除电缆接头的故障隐患, 延长电缆的运行寿命。

关键词: XLPE 电缆; 局部放电; 支持向量机; 模式识别

0 引言

交联聚乙烯(XLPE)绝缘电缆在耐热、机械性能方面具有良好优越性, 成为当前应用的主流电缆类型。XLPE绝缘电缆不仅利于城市美化、而且受自然条件(如雷电、风雨、盐雾、污秽等)和周围环境影响较小, 在城市中的应用越来越多。在电缆实际运行中, 除人为破坏和一些偶然因素外, 其绝缘故障一般都是由微小的局部放电发展而来, 电缆接头是电缆线路中绝缘结构相对薄弱的环节^[1], 因此对XLPE电力电缆接头的局部放电检测是及时发现故障隐患, 预测运行寿命, 保障其可靠运行的重要手段^[2]。

目前在电缆局部放电检测中常用到电磁耦合法的Rogowski线圈型宽频带电流传感器^[4], 和超声波传感器。本文将两种传感器的原理融合, 采用基于电磁耦合法的宽频带电流传感器与超声波传感器对电缆局部放电进行联合检测^[6], 并通过对 4 种典型局部放电模型的模拟实验, 检测出不同的局放信号。由于现场环境中存在着各种干扰, 需从噪声中提取真实可靠的局部放电信号, 也就是通常所说的消除噪声, 应用小波消噪法得到模式识别的数据^[7]。

模式识别作为局部放电监测的关键步骤, 其识别方法的选择至关重要^[8], 本文介绍了一种基于向量机的XLPE电缆接头局部放电的模式识别方法^[10], 并通过实验室中 4 种典型局部放电的模拟实验, 检测其识别效果。

1 模拟实验设计

电力电缆产生局部放电的原因有很多, 归纳起

来主要由电缆本身的质量原因、安装施工时的工艺不当、投入运行中由于机械的外力影响或电缆长时间运行造成的老化等原因引起。本文针对 XLPE 电缆局部放电本身的特点, 设计了几种有典型代表意义的放电模型, 分别对这几种放电模型施加电压, 观察其局部放电现象测量并测量放电信号。

电树是电力电缆绝缘故障的主要原因之一。气泡、杂质和金属毛刺等的存在均可导致XLPE绝缘电树的生成和扩大, 最终将导致绝缘击穿。当金属屏蔽层或者导线芯存在毛刺时, 其尖端电场因极不均匀产生放电, 我们在实验室可以利用针板放电模型对其进行模拟。对于绝缘内产生的微小气泡等产生的局部放电, 可以利用内部放电模型进行模拟; 此外, 由于电缆附件制造质量、电缆敷设安装质量等原因, 均易造成在电缆接头内存在悬浮电极, 引起局部放电, 进而造成绝缘劣化, 最终将导致绝缘击穿, 这种情况可以借由悬浮放电模型进行模拟^{[5][11]}。

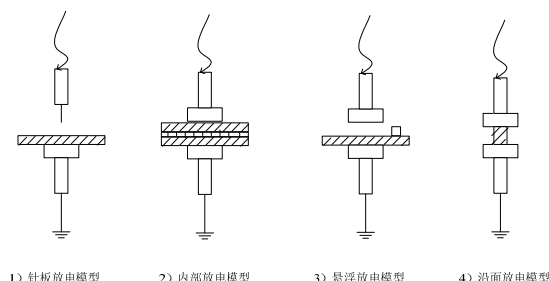


图 1 四种放电模型结构图

图 1 中:

1) 为针-板放电模型, 高压针尖曲率半径为 0.5mm, 锥角 30°, 尖长 15mm, 板电极上放置直径

100mm, 厚 1mm 的环氧树脂绝缘板, 针板距离为 5mm;

2) 为内部放电模型, 放电模型上下 2 层均采用厚 3mm 的环氧树脂板, 中间 1 层采用 1mm 厚的环氧树脂板作绝缘介质, 绝缘板上开有直径为 10mm 的圆孔, 绝缘板之间用环氧树脂胶粘结;

3) 为悬浮放电模型, 两个板电极之间距离为 10mm, 接地电极上放直径 100mm, 厚 5mm 的环氧树脂板, 环氧树脂板上靠近边缘处放置直径 10mm, 高 10mm 的铜柱;

4) 为沿面放电模型, 两个板电极之间纵向放置一个直径 10mm, 长 10mm 的环氧树脂棒。

4 种模型所用的板电极全部都采用直径 50mm, 厚 10mm 的圆形铜板, 为消除电极表面尖角或毛刺的影响, 电极表面和边缘全部打磨光滑。同时为防止电极引线处放电, 电极螺帽专门设计加工为球形螺帽。

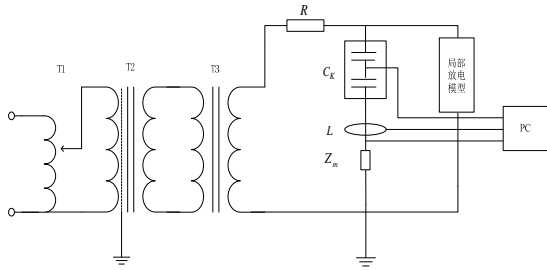


图 2 局部放电实验接线图

局部放电实验接线图见图 2。其中, R 为限流电阻; CK 为耦合电容; Zm 为测量阻抗; L 为宽频带电流传感器; 示波器为 Tektronix 2024 四通道示波器, 分别接收来自电容分压器、测量阻抗、宽频带电流互感器及超声波传感器上的电压信号; 示波器的带宽为 200MHz, 最高实时采样速率为 2GS/s; 超声波传感器选用接收式的探头, 测量时正对局放源, 距离约 200mm。在进行实验时, 所有设备和仪器均与大地进行有效的隔离, 避免了来自大地的外界杂散信号的干扰^[3]。

2 特征量提取与选择

由计算机采集卡获得的数据经消噪后便是局部放电的原始时域波形, 这种信号数据包含了局部放电的全部信息。但是不能将其直接作为模式识别分类器的输入, 所以要對数据进行统计分析, 提取放电工频相位 φ 、放电量 q 和放电次数 n 构成 φ - q - n 三维

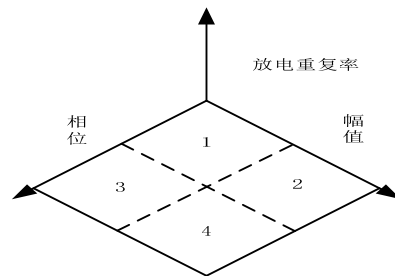
谱图^[10]。

从局部放电原始时域波形中可以得到构造谱图所需要的放电量 q_i 、放电相位 φ_i 和放电重复率 n 三个基本参数。其中放电量 q_i 表示第 i 次放电的视在放电量, 放电相位 φ_i 表示第 i 次放电脉冲所处的工频电压相位; 放电重复率 n 表示单位时间内的放电次数。局部放电的 φ - q - n 三维谱图同时反映一定放电量的放电脉冲发生相位和在该相位的放电重复率, 比较全面地反映了局部放电脉冲的统计特性, 其具体的实现过程是: 将相位 φ 轴等分为 100 份, 放电量 q 轴等分为 360 份, 整个 φ - q 平面被划分为 100×360 个小格, 每个小格的中心为 (φ_i, q_j) , 依据放电脉冲的放电量和发生相位, 统计每个小格中的脉冲个数, 并计算出放电重复率 n 。

$$n(\varphi_i, q_j) = \frac{\sum_k^M n_k(\varphi_i, q_j)}{M} \quad (1)$$

式中: $n(\varphi_i, q_j)$ 是在 (φ_i, q_j) 为中心的小格中单位时间内的放电次数, 简记为 n_{ij} ; M 为对局部放电进行检测的工频周期数; $n_k(\varphi_i, q_j)$ 是第 k 个周期内在小格内的放电脉冲个数; $k=1, 2, \dots, M$ 。在 φ - q - n 坐标系内画出点 (φ_i, q_j, n_{ij}) , 并将这些点连接起来, 就得到 φ - q - n 谱图。

φ - q - n 三维谱图给出了放电幅值、相位、重复率三者之间的关系, 它体现了更丰富更直观的放电特征。借助 φ - q - n 三维谱图可以帮助判断放电类型。在三维谱图中, 垂直坐标为放电重复率, 放电峰值高表示放电重复率高; 反之, 放电重复率低。平面坐标分别为放电幅值和相位, 可以把平面坐标分为 4 个区域, 如图 3 所示。



1-正半周、幅值低; 2-正半周、幅值高;
3-负半周、幅值低; 4-负半周、幅值高。

图 3 三维谱图的平面坐标区域划分

XLPE 电力电缆局部放电 φ - q - n 三维谱图统计算是反应局部放电信号本质特征的一组特征量。其中描述谱图投影形状差异的有偏斜度、陡峭度, 描

述谱图投影正负半周的轮廓差异有放电量因数、相位不对称度、互相关系数以及修正的互相关系数。有了这六个统计算子的定义，就可以将局部放电信号的 φ -q-n三维谱图分为正负两个半波谱图计算其相应的统计算子^{[3][7]}。

1) 偏斜度 S_k

局部放电 φ -q-n三维谱图的偏斜度是 φ -q-n谱图局部放电信号的特征之一，是局部放电模式识别的一个参考依据。偏斜度是用来描述某种形状的分布对比于正态分布形状的偏斜程度。其定义为：

$$S_k = \sum_{i=1}^W (x_i - \mu)^3 \cdot p_i \Delta x / \sigma^3 \quad (2)$$

式中： W 是谱图投影在半周期内的相窗数； x_i 是第 i 个相窗的相位； Δx 是相窗宽度； p_i 、 μ 和 σ 是把谱图投影看成概率密度分布图，以 φ_i 为随机变量时，相窗 i 内的事件出现的概率，均值和标准差。其概率为：

$$p_i = y_i / \sum_{i=1}^W y_i \quad (3)$$

式中： y_i 是谱图投影的纵坐标，它代表视在放电量 q 或放电重复率 n ；

均值和标准差分别为：

$$\mu = \sum_{i=1}^W p_i \varphi_i \quad (4)$$

$$\sigma = \sqrt{\sum_{i=1}^W p_i (\varphi_i - \mu)^2} \quad (5)$$

2) 陡峭度 Ku

陡峭度用于描述某种形状的分布对比于正态分布形状的突起程度。其定义为：

$$Ku = \left[\sum_{i=1}^W (x_i - \mu)^4 \cdot p_i \Delta x / \sigma^4 \right] - 3 \quad (6)$$

其中各量已在偏斜度中定义。

3) 放电量因数 Q

放电量因数 Q 能够反映 φ -q-n 谱图在 φ -q 平面投影的正负半周内平均放电量差异。其计算式为：

$$Q = \frac{\sum_{i=1}^W n_i^- q_i^-}{\sum_{i=1}^W n_i^-} / \frac{\sum_{i=1}^W n_i^+ q_i^+}{\sum_{i=1}^W n_i^+} \quad (7)$$

式中： n_i^+ 、 n_i^- 是相窗 i 内的放电重复率，上标“+”，“-”分别对应于 φ -q-n 谱图在 φ -q 平面投影的正负半周。

4) 相位不对称度 ϕ

相位不对称度 ϕ 能够反映 φ -q-n 谱图在 φ -q 平面投影的正负半周内放电起始相位的差别。其计算式为：

$$\phi = \frac{\varphi_{in}^-}{\varphi_{in}^+} \quad (8)$$

其中： φ_{in}^+ 、 φ_{in}^- 分别是 φ -q-n 谱图在 φ -q 平面投影的正负半周内放电起始相角。

5) 互相关系数 cc

一个局部放电信号的 φ -q-n 三维谱图在 φ -q 平面投影的正负半波相似程度用相关系数 cc 表示。其计算公式为：

$$cc = \frac{\sum_{i=1}^W q_i^+ q_i^- - \left(\sum_{i=1}^W q_i^+ \sum_{i=1}^W q_i^- \right) / W}{\sqrt{\left[\sum_{i=1}^W (q_i^+)^2 - \left(\sum_{i=1}^W q_i^+ \right)^2 / W \right] \left[\sum_{i=1}^W (q_i^-)^2 - \left(\sum_{i=1}^W q_i^- \right)^2 / W \right]}} \quad (9)$$

式中： q_i^+ 、 q_i^- 是相窗 i 内的平均放电量，上标“+”，“-”分别对应着 φ -q-n 谱图在 φ -q 平面投影的正负半周； W 是 φ -q-n 谱图在 φ -q 平面的投影在半周内相窗数。

6) 修正的互相关系数 mcc

修正的互相关系数 mcc 用于评价 φ -q-n 谱图在 φ -q 平面投影的正负半周内放电模式的差异。其计算式为：

$$mcc = Q\phi cc \quad (10)$$

3 支持向量机的模式识别原理

机器学习基本思想是对样本数据进行训练并挖掘出规律，并运用这些规律研究客观对象，来对未知数据或无法预测的数据进行预测和判断。由于其基于统计学习理论，所以机器学习就对解决有限样本情况下的统计学习问题提供了一个有效地途径。支持向量机(Support Vector Machines, SVM)作为一种新的机器学习算法，其主要基于统计学的 VC 维理论和结构风险最小化原理，通过寻求结构风险最小化来实现实际风险的最小化，并以求达到在有限信息的条件下得到最优的结果的目的。

支持向量机的基本思想就是找到最优分类线，

推广到高维就是找到最优分类面^[9]。所谓的最优分类线就是不但要能将两类正确分开，使训练错误率为 0，而且还要使分类间隔最大。这样就既能够保证经验风险最小，也能使真实风险最小。如图 4 所示，白色小圆圈和黑色小圆圈是在二维情况下的两种不同类的数据样本，其中W便是分类线，分类间隔定义为W1 与W2 之间的距离。对非线性分类问题，SVM一般通过核函数将其转化为高维空间的线性问题予以解决。

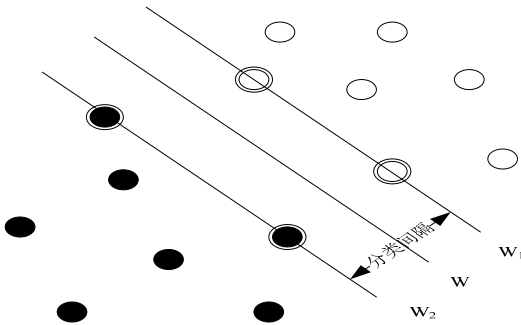


图 4 最优分类面示意图

1) C-SVM 算法

由于本文采用的是基于 C-SVM 算法的 C-支持向量分类机(C-SVC)，所以下面简要介绍一下 C-SVM 算法。

对于非线性分类问题，首先通过引进松弛变量放松约束，再者引进从输入空间 R^n 到 Hilbert 空间 H 的变换 Φ 把数据样本集 (x_i, y_i) , $i=1, \dots, n$, $x_i \in R^n$, $y_i \in \{+1, -1\}$ 映射为 $\phi((x_1), y_1)$, $\phi((x_2), y_2)$, \dots , $\phi((x_i), y_i)$ ，得到如下的原始问题：

$$\min_{\omega \in H, b \in R, \xi \in R^l} \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^l \xi_i \quad (11)$$

$$\text{s.t.} \quad y_i((\omega \cdot x_i) + b) \geq 1 - \xi_i, \quad i=1, L, l \quad (12)$$

$$\xi_i \geq 0, \quad i=1, L, l \quad (13)$$

其中： $C > 0$ 。

若对应于变换 Φ 的核函数为 $K(x_i, x_j)$ ，则原始问题式(11)~式(13)的对偶问题为：

$$\min_{\alpha} \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l y_i y_j \alpha_i \alpha_j K(x_i, x_j) - \sum_{j=1}^l \alpha_j \quad (14)$$

$$\text{s.t.} \quad \sum_{i=1}^l y_i \alpha_i = 0 \quad (15)$$

$$0 \leq \alpha_i \leq C, \quad i=1, L, l \quad (16)$$

若 K 是正定核，则对偶问题式(14)~式(16)必有解，且其解为 $\alpha^* = (\alpha_1^*, L, \alpha_l^*)^T$ ，使得：

$$\omega^* = \sum_{i=1}^l \alpha_i^* y_i x_i \quad (17)$$

$$b^* = y_j \left(1 - \frac{\alpha_j^*}{C}\right) - \sum_{i=1}^l y_i \alpha_i^* K(x_i, x_j) \quad (18)$$

其中：系数 α_i 只有当相应的数据样本点 (x_i, y_i) 刚好使约束式(12)等号成立时才可能非零。

由上述分析可建立 C-支持向量机(C-SVM)，其对应的数学描述如下：

(1) 设已知数据样本集 (x_i, y_i) ，其中 $i=1, \dots, n$, $x_i \in R^d$, $y_i \in \{+1, -1\}$ 。

选取适当的核函数 $K(x_i, x_j)$ 和适当的参数 C ，构造并求解最优化问题：得式(14)~式(16)最优解，

$$\alpha^* = (\alpha_1^*, L, \alpha_l^*)^T \quad (19)$$

(2) 选取 α^* 的一个正分量 $0 < \alpha_j^* < C$ ，并据此计算阈值。

$$b^* = y_i - \sum_{i=1}^l y_i \alpha_i^* K(x_i, y_i) \quad (20)$$

(3) 构造决策函数

$$f(x) = \text{sgn} \left(\sum_{i=1}^l \alpha_i^* y_i K(x, x_i) + b^* \right) \quad (21)$$

2) 核函数

核函数方法就是用非线性变换 Φ 将 n 维矢量空间中的随机矢量 x 映射到高位特征空间，在高维特征空间中设计线性学习算法，若其中各坐标分量间的相互作用仅限于内积，则不需要知道非线性变换 Φ 的具体形式，只要满足 Mercer 条件的核函数替换线性算法中的内积，就能得到原输入空间中对应的非线性算法^{[9] [10]}。常用的满足 Mercer 条件的核函数有多项式函数、径向基函数和神经网络函数等，选取不同的核函数构造的支持向量机也不同。

(1) q 次多项式函数

选取以下核函数

$$K(x, x_i) = [(x, x_i) + 1]^q \quad (22)$$

构造的支持向量机的判别函数为：

$$f(x, \alpha) = \text{sgn} \left\{ \sum_{i=1}^s \alpha_i y_i [(x_i, x) + 1]^q - b \right\} \quad (23)$$

其中: s 为支持向量的个数。可以通过多项式的次数 q 来控制系统的 VC 维数。

(2) 径向基(Radial Basis Function, RBF)函数。
径向基核函数如下:

$$K(\mathbf{x}, \mathbf{x}_i) = \exp(-\gamma \|\mathbf{x} - \mathbf{x}_i\|^2) \quad (24)$$

式中: γ 是一个控制半径的正数。

3) 分类

本文采用决策树 SVM 算法进行分类[8], 该算法是对一个 N 类分类问题, 首先选取其中一类作为正训练样本, 其余 $N-1$ 类作为负训练样本建立 SVM1; 然后选取剩余 $N-1$ 类中一类作为正训练样本, 剩余的 $N-2$ 类作为负训练样本建立 SVM2; 按这样的方法一直建立到 SVM($N-1$); 最终按图 5 进行识别判断得出结果。这种方法的优点是: 1) 训练速度较快; 2) 结果中不会出现同属于多类或任一类的情况。但决策树 SVM 中的一个关键的问题就是决策树的根节点以及其他节点的顺序, 其对分类结果有很大的影响。

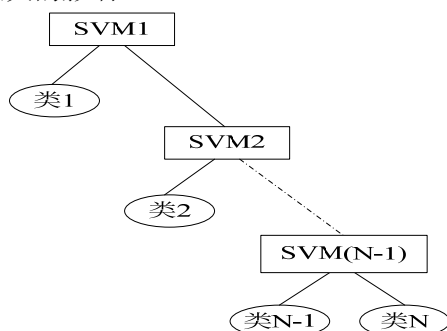


图 5 决策树 SVM

4 实验及结果分析

决策树 SVM 算法的关键在于如何安排决策树的根节点以及其他节点的顺序。为了寻找一个最优的方案, 本文对不同的排序进行了识别判断。其中 A 类缺陷表示内部放电、B 类缺陷表示悬浮放电、C 类缺陷表示沿面放电、D 类缺陷表示针板放电。以表中第一组顺序(ABCD)为例: 当 SVM1 的正训练样本选 A 类缺陷的话, 那从 A、B、C、D 四种类型缺陷中识别出 A 类缺陷的正确率为 0%; 当 SVM2 的正训练样本选 B 类缺陷的话, 那从 B、C、D 三种类型缺陷中识别出 B 类缺陷的正确率为 90%; 当 SVM3 的正训练样本选 C 类缺陷的话, 那从 C、D 两种类型缺陷中识别出 C 类缺陷的正确率为 100%。

针对不同的顺序各分类器正样本的识别正确率并不相同。所以通过对这 24 组顺序的识别正确率研究得出, DCAB 顺序的识别效果最好, 其中 SVM1 选 D 类缺陷为正训练样本的识别正确率为 90%; SVM2 选 C 类缺陷为正训练样本的识别正确率为 85%; SVM3 选 A 类缺陷为正训练样本的识别正确率为 95%。

5 结论

本文通过对支持向量机的基本原理的阐述, 对电缆接头放电模式识别提供了一种新的研究方法。该模式识别方法能够有效地识别出电缆接头的 4 种典型放电类型, 从中对电缆接头局部放电程度进行分析, 实时掌握电缆接头局部放电情况, 能及时有效地排除故障隐患, 确保电缆的可靠运行, 达到延长电缆运行寿命的经济价值。

参考文献:

- [1] 罗俊华, 邱毓昌, 杨黎明. 10kV 及以上电力电缆运行故障统计分析[J]. 高电压技术, 2003, 29(6): 14-16.
- [2] Mashikian M. Preventive maintenance testing of shielded power cable systems[J]. IEEE Trans. on Ind App. l(IA), 2002, 38: 736-743.
- [3] 段乃欣. 用于电力电缆局放检测的宽频带电磁耦合合法的研究[D]. 西安: 西安交通大学, 2003.
- [4] 李华春, 周作春, 张文新, 等. 国内外几种电缆局部放电在线检测方法技术分析[A]. 全国第八次电力电缆运行经验交流会[C]. 2008: 917-926.
- [5] 李燕青, 陈志业, 律方成, 等. 超声波法进行变压器局部放电模式识别的研究[J]. 中国电机工程学报, 2003, 23(2): 108-111.
- [6] T Boczar. Identification of specific type of PD from acoustic emission frequency spectra[J]. IEEE transactions on dielectrics and dielectrical insulation, 2001, 8(4): 598-606.
- [7] 杨孝华, 廖瑞金, 胡建林, 等. 基于 BP 人工神经网络的 XLPE 电力电缆局部放电的模式识别[J]. 高压电器, 2003, 39(4): 35-37.

作者简介:

刘永生 (1984-), 男, 河北保定人, 硕士研究生, 主要从事输电线路设计工作。