

基于样本加权 FCM 聚类的未知类别局部放电信号识别

贾亚飞^{1,2}, 朱永利¹, 高佳程¹, 袁 博³

(1. 华北电力大学 新能源电力系统国家重点实验室, 河北 保定 071003;

2. 国网河北省电力有限公司雄安新区供电公司, 河北 雄安新区 071800;

3. 国网河北省电力公司经济技术研究院, 河北 石家庄 050000)

摘要:针对电力变压器待识别局部放电信号中可能存在不属于已知类别的未知样本的问题,提出了一种基于样本加权模糊 C 均值(FCM)聚类的未知类别局部放电信号识别方法。对已知类别的局部放电信号进行 FCM 聚类,确定各已知类的聚类中心;分别计算已知类别和待识别局部放电信号的样本权值,并根据 Otsu 准则确定样本权值的自适应阈值;将各待识别局部放电信号的权值与确定的阈值进行比较,判断其是否属于已知类别;采用支持向量机(SVM)对得到的属于已知类别的待识别局部放电信号进行分类,对未知类别样本进行人为分析判断。采用所提方法对实验室条件下的放电信号进行分析,实验结果表明,所提方法可以有效地区分待识别局部放电信号中的未知类别样本。

关键词:电力变压器;局部放电;模式识别;未知样本;样本加权;FCM 聚类;Otsu 准则;支持向量机

中图分类号:TM 41;TM 835

文献标识码:A

DOI:10.16081/j.issn.1006-6047.2018.12.016

0 引言

电力变压器是电力系统输变电的关键性设备,其运行的可靠性直接关系到整个电力系统的安全与稳定。变压器故障统计分析表明^[1],绝缘故障是影响变压器可靠运行的重要原因之一,而局部放电是导致变压器绝缘破坏的主要原因。不同类型缺陷产生的局部放电具有不同的特征^[2-3],因此对局部放电信号进行有效的模式识别,可以准确地了解和掌握变压器内部缺陷类型的性质和特征,对指导变压器的检修工作意义重大。

分类器的设计是实现局部放电模式识别的关键环节之一,主要分为无监督模式识别和有监督模式识别两大类。聚类分析是无监督模式识别中的一个重要分支,文献[4]利用模糊 C 均值 FCM(Fuzzy C-Means)算法针对提取的特征向量对放电源脉冲进行聚类,取得了较好的聚类效果;文献[5]提出了基于 K-means 聚类分析的局部放电谱图自动识别方法,经现场实际测试数据和在线监测数据测试证明,该方法在相位信息缺失的情况下能够实现对局部放电监测数据的自动诊断分析,得到局部放电谱图的识别结果;文献[6]利用 Mahalanobis 距离算法系统地阐述了各类放电在三角坐标系中的聚类分布情况,实现了对典型局部放电信号模式的准确识别。常用的有监督模式识别方法有支持向量机 SVM(Support

Vector Machine)、神经网络、相关向量机等。文献[7]提出了一种基于多分组特征的组合核多分类 SVM 的局部放电信号识别方法,寻找最优核函数组合分类模型,对多个特征空间数据具有普适性,且融合效果理想。文献[8]提出了一种基于自适应小波神经网络的局部放电信号识别方法,构建了一个 4 层自适应小波神经网络模型,并且采用粒子群优化算法进行一次优化,采用 BP 算法进行二次优化,研究结果表明该方法的训练效果明显优于单独使用 BP 算法的训练效果。文献[9]采用多分类相关向量机 M-RVM(Multiclass Relevance Vector Machine)实现局部放电信号的多分类问题,该方法的基函数权值少数非零,诊断速度快,可以有效解决小样本、高维、非线性分类问题。无监督的模式识别方法虽然可以有效、快速地进行局部放电信号的划分,但是这一诊断过程只是实现了局部放电信号的初步分析,并没有进一步给出局部放电信号的类型等信息。SVM^[10-11]和神经网络^[12]等有监督分类方法需要对具有明确放电类别的样本进行训练和学习,然后用训练好的分类器对待识别样本进行分类。

目前,虽然多种有监督模式识别方法已经成功应用于局部放电信号的模式识别,但是还没有关于待识别局部放电信号中存在不属于已知类别的局部放电信号,即未知类别局部放电信号识别的研究。对于变压器而言,经过长期的监测工作和经验积累,可整理得到一部分可用的具有明确类型信息的局部放电信号样本,但变压器本身的结构复杂,要想获得全部局部放电信号类型的训练样本十分困难。而有监督模式识别方法只能识别已知放电类别,如果不对不属于已知类别(未知类别)的局部放电信号进行分类,该类信号会被分类到已知类别中,导致错误的

收稿日期:2018-05-17;修回日期:2018-09-17

基金项目:国家自然科学基金资助项目(51677072);中央高校基本科研业务费专项资金资助项目(2016XS101)

Project supported by the National Natural Science Foundation of China(51677072) and the Fundamental Research Funds for the Central Universities(2016XS101)

诊断。因此,在对待识别局部放电信号进行分类前,有必要对待识别局部放电信号中不属于已知类别的局部放电信号进行识别。

局部放电是一个渐变的过程,提取的特征通常具有模糊性,若直接采用提取的特征值来辨识放电类型会存在一定的难度。以模糊数学为基础的聚类分析方法为解决这类问题提供了一种可行的途径。FCM 是模糊聚类中应用最为广泛的一种算法,它依据数据样本间的相似度,通过迭代优化目标函数,将相似度高的样本对象划分为同一类来实现数据的分类。因此,FCM 在局部放电信号模式识别中得到了广泛的应用。

基于上述问题,为了充分利用收集到的已知类别的局部放电信号,并且能有效识别不属于已知类别的局部放电信号,本文提出了一种基于样本加权 FCM 聚类的未知类别局部放电信号识别方法,对待识别局部放电信号进行分析。该方法根据各待识别局部放电信号与已知类别聚类中心的距离确定权值,利用 Otsu 准则,即采用最大类间方差法自适应地选取样本权值阈值,通过比较确定各局部放电信号是否属于已知类别,挑选得到未知类别的局部放电信号,然后对属于已知类别的局部放电信号进行分类。可以对挑选得到的未知类别的局部放电信号进行人工分析判断,确定其放电类型后将其加入已知类别的训练样本中。为了验证本文所提方法的有效性,采用 SVM 方法分别对所有待识别局部放电信号样本(包含未知类别的放电信号样本)和经样本加权 FCM 聚类方法处理后的样本进行分类。实验结果表明,经基于样本加权 FCM 聚类方法处理后的样本的正确识别率要远远高于直接对所有待识别样本进行分类的正确识别率,充分证明了本文所提方法可有效识别未知类别的局部放电信号样本,对局部放电信号的模式识别具有重要意义。

1 样本加权 FCM 聚类算法

1.1 FCM 聚类算法原理

FCM 算法是一种经典的基于目标函数的聚类算法,其目的是将 n 个样本无监督地划分为 c 个类别。FCM 算法在传统硬聚类算法中引入了模糊技术,通过极小化所有数据点与各个聚类中心的欧氏距离及模糊隶属度的加权和,不断迭代修正聚类中心和分类矩阵,直到符合终止准则,将具有相似特性的数据样本聚为一类^[13-14]。

设有一组含 n 个样本的样本集 $X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$,其中每个样本具有 p 个特征属性,即 $\mathbf{x}_k = [x_{k1}, x_{k2}, \dots, x_{kp}]$ ($k = 1, 2, \dots, n$)。各样本以一定的程度隶属于 c 个不同的区域,用隶属度 μ_{ij} 表示第 j 个样本隶属于第 i 类的程度, $\mu_{ij} \in [0, 1]$ 。

FCM 聚类的目标函数为:

$$J_m(\mathbf{U}, \mathbf{V}) = \sum_{k=1}^n \sum_{i=1}^c \mu_{ik}^m d_{ik}^2 \quad (1)$$

约束条件为:

$$\sum_{i=1}^c \mu_{ik} = 1 \quad (2)$$

其中, $\mathbf{U} = [\mu_{ik}]_{c \times n}$ 为隶属度矩阵; $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_c]$ 为聚类中心矩阵; $m \in [1, \infty)$ 为平滑参数,一般取 $m = 2$; d_{ik} 为样本 $\mathbf{x}_k$ 到聚类中心 $\mathbf{v}_i$ 的距离, $d_{ik} = \|\mathbf{x}_k - \mathbf{v}_i\|$ 。

FCM 算法的求解过程中,在式(2)所示约束条件下,算法通过不断迭代来更新聚类目标函数值、聚类中心和隶属度矩阵,以最小化目标函数,直至聚类中心不再变化或 2 次迭代的目标函数值之差在允许的范围内,其中目标函数值误差采用默认参数 10^{-5} 。得到的隶属度和聚类中心分别如式(3)、式(4)所示。

$$\mu_{ik} = \frac{1}{\sum_{h=1}^c (d_{ik}/d_{hk})^{2/(m-1)}} \quad (3)$$

$$\mathbf{v}_i = \frac{\sum_{k=1}^n \mu_{ik}^m \mathbf{x}_k}{\sum_{k=1}^n \mu_{ik}^m} \quad (4)$$

可以通过分离系数和分离熵 2 个性能指标进行聚类效果检验,其定义分别如式(5)、式(6)所示。

$$F = \frac{1}{n} \sum_{j=1}^n \sum_{i=1}^c \mu_{ij}^2 \quad (5)$$

$$H = -\frac{1}{n} \sum_{j=1}^n \sum_{i=1}^c \mu_{ij} \ln \mu_{ij} \quad (6)$$

分离系数 F 和分离熵 H 中包含了隶属度信息。分离系数 F 越接近于 1,则聚类效果越好;分离熵 H 越接近于 0,则聚类效果越好。

1.2 样本加权 FCM 聚类算法

在传统的聚类算法中,所有样本都同等地参与聚类过程。然而,待聚类样本集合中的每个样本在聚类过程中所起的作用并不一样,对聚类结果产生的影响也有所不同。样本加权 FCM 聚类算法通过为离群点赋予很小的权值,降低离群样本点参与聚类过程的程度,从而削弱离群样本点对聚类效果的影响。

样本权值的确定主要有 2 种方法,可以根据实际需要进行选择。

第 1 种权值计算方法是描述样本点相互之间的接近程度,计算公式如式(7)所示。

$$\varphi_j = \sum_{k=1}^n \exp(-\alpha \|\mathbf{x}_j - \mathbf{x}_k\|^2) \quad (7)$$

其中, α 为常数; $j = 1, 2, \dots, n$ 。

第 2 种权值计算方法描述各样本与各聚类中心的接近程度,计算公式如式(8)所示。

$$\varphi_j = \sum_{k=1}^c \exp(-\alpha \|x_j - v_k\|^2) \quad (8)$$

2 基于样本加权 FCM 聚类的未知类别局部放电信号识别方法

由于电力变压器的结构复杂,对于收集到的待识别局部放电信号集合中可能存在不属于已知放电类别的未知放电信号,若直接进行分类,会将其分到已知放电类别中,导致误诊断。因此,有必要在对待识别放电信号进行分类前挑选出不属于已知类别的放电信号。本文基于样本加权 FCM 聚类算法的思想,计算各待识别放电信号的权值,然后根据各样本权值的大小判断样本是否属于已知类别。待识别局部放电信号样本权值和权值阈值的确定是实现该方法的 2 个关键步骤。

2.1 样本权值计算

1.2 节给出了 2 种样本权值的计算方法,第 1 种方法(式(7))是通过计算每个样本点与其余样本点的距离来定量描述样本在整个样本集中的位置,该方法虽然在一定程度上可以刻画样本点相互之间的接近程度,但是计算量比较大,并且没有充分利用已知类别的样本信息;第 2 种方法(式(8))是通过计算样本与各个已知类别样本的聚类中心的距离来描述样本对各聚类中心的接近程度,该方法对聚类中心比较敏感,当聚类中心较为准确时权值才会更加合理。

本文所提基于样本加权 FCM 聚类的未知局部放电信号识别方法首先得到已知类局部放电信号的聚类中心,即可以确定准确的聚类中心,因此,本文选用第 2 种方法计算样本权值。

2.2 样本权值的自适应阈值选取方案

根据待识别局部放电信号样本权值的大小可以将待识别样本划分为已知类别样本和未知类别样本(与各聚类中心距离较远)2 类,其中样本权值阈值 T 的确定是该问题的关键。由于无法掌握待识别放电信号的信息(是否属于已知类别的放电信号),无论是设定统一的阈值还是根据经验手动设定,都无法满足待识别放电信号样本权值阈值确定的需要。本文提出一种基于 Otsu 准则^[15-16]的自适应确定样本权值阈值 T 的方法,即选取某阈值 T 令已知类别样本和未知类别样本间的方差最大,此时 2 类样本之间的差异也最大,因此,该阈值 T 是最优划分阈值^[17]。基于 Otsu 准则确定样本权值阈值 T 的基本原理如下^[15-16]。

对于 N 个局部放电信号样本,将各样本权值看作长度为 N 的离散序列 $\{w_i, i=1, 2, \dots, N\}$, $w_{\max}$ 、

$w_{\min}$ 分别为该序列的最大值和最小值。

为了方便描述,引入灰度概念,即对应放电样本权值的大小。设定灰度等级 L (即放电样本权值大小的等级),令 $d_w = (w_{\max} - w_{\min})/L$ 。统计权值大小落在 $[(l-1)d_w, ld_w]$ 范围内的值的个数 n_l ,其中 $l=1, 2, \dots, L$ 为灰度值, n_l 是灰度值为 l 时的像素。 n_l 出现的概率为 $p_l = n_l/N_s$, $N_s = n_1 + n_2 + \dots + n_L$ 为总像素数。

假设以 gd_w 作为阈值将序列 $\{w_i, i=1, 2, \dots, N\}$ 的样本权值分成 2 类,即样本权值大小落在 $[0, gd_w]$ 范围之内的构成一类,记作 C_0 ;样本权值大小落在 $[(g+1)d_w, Ld_w]$ 范围内的构成另一类,记作 C_1 。2 类的灰度均值 $\mu_0(g)$ 、 $\mu_1(g)$ 以及方差 $\sigma_0^2(g)$ 、 $\sigma_1^2(g)$ 的计算公式如下:

$$P_0(g) = \sum_{l=1}^g p_l \quad (9)$$

$$P_1(g) = \sum_{l=g+1}^L p_l = 1 - P_0(g) \quad (10)$$

$$\mu_0(g) = \sum_{l=1}^g l \frac{p_l}{P_0(g)} = \frac{1}{P_0(g)} \sum_{l=1}^g lp_l \quad (11)$$

$$\mu_1(g) = \sum_{l=g+1}^L l \frac{p_l}{P_1(g)} = \frac{1}{P_1(g)} \sum_{l=g+1}^L lp_l \quad (12)$$

$$\sigma_0^2(g) = \sum_{l=1}^g [l - \mu_0(g)]^2 \frac{p_l}{P_0(g)} \quad (13)$$

$$\sigma_1^2(g) = \sum_{l=g+1}^L [l - \mu_1(g)]^2 \frac{p_l}{P_1(g)} \quad (14)$$

C_0 和 C_1 间的类内方差之和为:

$$\sigma_b^2(g) = P_0(g) [\mu_0(g) - \mu(g)]^2 + P_1(g) [\mu_1(g) - \mu(g)]^2 \quad (15)$$

其中, μ 为整个序列的灰度均值,计算公式如式(16)所示。

$$\mu = \sum_{l=1}^L lp_l = P_0(g)\mu_0(g) + P_1(g)\mu_1(g) \quad (16)$$

则可以确定最优阈值为 $g^* d_w$, 令:

$$\sigma_b^2(g^*) = \max_{1 \leq g \leq L} \sigma_b^2(g) \quad (17)$$

综合式(9)——(17),可以得到最优样本权值的阈值,即 $T = g^* d_w$ 。

2.3 基于样本加权 FCM 聚类的未知局部放电信号识别方法

本文根据局部放电信号样本的权值大小来判断该信号是否属于已知类别。分别提取已知类别的 d 类局部放电信号和待识别局部放电信号的特征量,得到的特征量分别构成已知类别局部放电信号样本 $X = \{x_1, x_2, \dots, x_N\}$ 和待识别局部放电信号样本 $S = \{s_1, s_2, \dots, s_M\}$, 其中 N 和 M 分别为已知类别局部放电信号样本的个数和待识别局部放电信号样本的个数。

本文采用文献[18]中的变分模态分解 VMD (Variational Mode Decomposition) 和多尺度熵 MSE (MultiScale Entropy) 对局部放电信号进行特征提取。利用 VMD 方法对局部放电信号进行分解得到模态分量,利用 MSE 方法对得到的分解模态进行定量描述,形成特征向量,然后利用主成分分析法 PCA (Principle Component Analysis method) 对得到的特征向量进行降维处理,将其作为局部放电信号特征向量。该特征提取方法的流程图如图 1 所示,具体过程不再赘述。

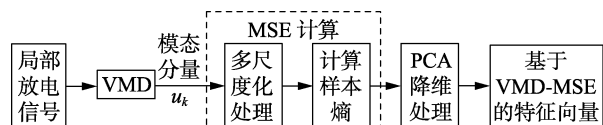


图 1 基于 VMD-MSE 方法的特征提取过程

Fig.1 Feature extraction process based on VMD-MSE method

基于样本加权 FCM 聚类的未知局部放电信号识别方法的流程图见附录中图 A1,具体实现步骤如下。

a. 采用 FCM 算法对已知类别的 d 类局部放电信号样本进行聚类,确定各类别的聚类中心 $\mathbf{v} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_d]$ 。

b. 根据步骤 a 得到的聚类中心和式(8)分别计算已知类别的局部放电信号样本和待识别的局部放电信号样本的样本权值,分别记为 $\mathbf{w}_X = [w_1, w_2, \dots, w_N]$ 和 $\mathbf{w}_S = [w_{N+1}, w_{N+2}, \dots, w_{N+M}]$ 。

c. 将待识别局部放电信号样本的权值和已知类别局部放电信号样本的权值进行组合,即 $\mathbf{w} = [\mathbf{w}_S, \mathbf{w}_X]$,根据 Otsu 准则确定样本权值的自适应阈值 T 。

d. 将待识别局部放电信号样本权值集合 $\mathbf{w}_S = [w_{N+1}, w_{N+2}, \dots, w_{N+M}]$ 中各样本权值分别与阈值 T 进行比较。若 $w_i \geq T (N+1 \leq i \leq N+M)$,则该局部放电信号样本 s_i 属于已知放电类别;若 $w_i < T$,则该局部放电信号样本 s_i 属于未知放电类别。

e. 用常规模式识别方法对待识别样本中属于已知类别的局部放电信号样本进行分类,对于属于未知类别的局部放电信号样本一一进行人为分析判断,确定其放电类型后将其归为已知类别的局部放电信号样本集中,对已知类别的局部放电信号样本集进行扩充。

3 实例分析

3.1 局部放电信号采集

本文的局部放电信号是在实验室环境下对不同放电模型进行局部放电信号检测得到的。试验采用 IEC60270—2000 标准,试验电路为基于脉冲电流法的并联测试电路,采用 TWPD-2F 局部放电综合分析仪。局部放电信号的采集频率为 20 MHz,带宽为 40~300 kHz。以每个工频周期记录到的数据作为一

个局部放电信号。

本文中已知类别的局部放电信号集合由电晕放电模型、油中多尖对板放电模型、油中板对板放电模型和油中悬浮颗粒放电模型 4 种放电模型产生的局部放电信号组成,4 种放电模型装置分别如附录中图 A2(a)—(d)所示。未知局部放电信号集合由上述 4 种放电模型之外的油中锥对板放电模型产生,如附录中图 A2(e)所示。

3.2 基于样本加权 FCM 聚类的未知局部放电信号识别

本文选取电晕放电模型、油中多尖对板放电模型、油中板对板放电模型和油中悬浮颗粒放电模型 4 种放电模型产生的 200 个局部放电信号,其中每种放电模型各产生 50 个局部放电信号。按照文献[18]提出的特征提取方法提取各局部放电信号的特征量,形成已知类别的局部放电信号样本集 X 。此外,收集 100 个未知类型的局部放电信号(包含 80 个已知类别的局部放电信号和 20 个未知类别的局部放电信号),进行特征提取后得到的特征量形成待识别局部放电信号样本集 S 。采用本文所提基于样本加权 FCM 聚类的未知类别局部放电信号识别方法对待识别局部放电信号样本集进行处理。其中采用基于样本加权 FCM 聚类方法的关键是权值阈值的确定,此处采用 2.2 节介绍的 Otsu 准则自适应地确定样本权值阈值 T 。

将未经处理的待识别局部放电信号样本集和经本文所提方法处理后的样本集采用 FCM 方法进行聚类。聚类效果可以通过分离系数 F 和分离熵 H 这 2 个指标进行检验。未经处理的待识别局部放电信号样本集和经本文所提方法处理后的样本集采用 FCM 方法进行聚类后的分离系数 F 和分离熵 H 如表 1 所示。

表 1 FCM 的聚类效果评价指标

Table 1 Evaluation indexes for clustering result of FCM

样本集	F	H
未经处理	0.833 5	0.368 3
经本文所提方法处理	0.751 0	0.523 9

由表 1 可知,对经本文所提方法处理后的待识别局部放电信号样本进行 FCM 聚类的聚类性能指标要远远优于对未经处理的待识别局部放电信号样本直接进行 FCM 聚类的聚类性能指标,这充分说明了采用样本加权 FCM 聚类方法可以有效地改善含未知类别的待识别样本的聚类效果,验证了所提方法的有效性。

采用基于样本加权 FCM 聚类的未知类别局部放电信号识别方法对待识别局部放电信号样本集进行处理后,属于已知类别和属于未知类别的局部放电信号样本判断的正误情况如表 2 所示。由表 2 可

知,本文所提未知类别局部放电信号识别方法将属于已知类别的局部放电信号样本全部正确分类到已知类别样本集中,属于未知类别的局部放电信号中有 4 个错误地分类到已知类别样本集合中。虽然本文所提方法仍然存在一定的判断误差,但已经能对绝大部分的待识别样本进行正确划分,正确率达 96%,这也充分说明了本文所提方法对未知类别样本识别的有效性,为进一步的模式识别创造了先决条件。

表 2 局部放电信号样本类别判断的正误情况

Table 2 Right and wrong situation of type judgment for partial discharge signal samples

实际输出的样本类别	期望的样本类别	
	已知放电类别	未知放电类别
已知放电类别	80	4
未知放电类别	0	16

3.3 局部放电信号模式识别

为了进一步验证本文所提方法的有效性,本文采用 SVM 分别对未经处理和采用本文所提方法处理后的局部放电信号样本集进行分类。SVM 中核函数选取径向基(RBF)核函数,参数均取最优,其中惩罚参数为 5,核参数为 0.2。分类结果如表 3 所示。由表 3 可知,对未经处理的待识别放电信号样本集直接利用 SVM 进行分类,虽然样本集中各已知类别样本的正确识别率均为 100%,但总体正确识别率却只有 80%。这是因为待识别样本集中含有一部分不属于已知类别的局部放电信号,直接用 SVM 进行分类会将这部分样本错分为已知类别,从而造成误判。而经本文所提方法处理后的待识别样本经 SVM 进行分类后,总体正确识别率提高到 95.24%。实验结果充分验证了本文所提基于样本加权 FCM 聚类的未知局部放电信号识别方法可有效地处理待识别样本集中的未知类别信号,提高局部放电信号识别的准确率。

表 3 基于 SVM 的识别结果

Table 3 Recognition results based on SVM

放电类型	正确识别率/%	
	待识别样本经本文方法处理	待识别样本未经处理
电晕放电	100	100
油中多尖对板放电	100	100
油中板对板放电	100	100
油中悬浮颗粒放电	100	100
总体正确识别率/%	95.24	80

4 结论

a. 本文提出了一种基于样本加权 FCM 聚类的未知类别局部放电信号识别方法,实现对待识别局部放电信号中未知类别样本的识别。采用 FCM 聚类确定已知类别局部放电信号样本的各类聚类中

心,计算局部放电信号的样本权值,将各待识别局部放电信号的权值与确定的阈值进行对比,判断其是否属于已知类别。

b. 提出了根据 Otsu 准则确定样本权值的自适应阈值,为样本权值阈值的确定提供了一种可行方案。

c. 采用基于样本加权 FCM 聚类的未知类别局部放电信号识别方法对待识别局部放电信号进行处理,将得到的属于已知类别的待识别局部放电信号采用 SVM 进行分类。将结果与未经处理的待识别局部放电信号直接进行 SVM 分类的结果进行比较,结果表明,采用本文所提方法处理后的局部放电信号具有较高的正确识别率,为未知类别局部放电信号的识别提供了一条有效的解决途径。

附录见本刊网络版(<http://www.epae.cn>)。

参考文献:

- [1] 严璋,朱德恒. 高电压绝缘技术[M]. 北京:中国电力出版社,2016:230-240.
- [2] 苑津莎,尚海昆. 基于主成分分析和概率神经网络的变压器局部放电模式识别[J]. 电力自动化设备,2013,33(6):27-31.
YUAN Jinsha, SHANG Haikun. Pattern recognition based on principal component analysis and probabilistic neural networks for partial discharge of power transformer[J]. Electric Power Automation Equipment, 2013, 33(6): 27-31.
- [3] 廖瑞金,段炼,汪可,等. 基于时频分析和 2DNMF 的局部放电模式识别[J]. 电力自动化设备,2013,33(3):20-25.
LIAO Ruijin, DUAN Lian, WANG Ke, et al. Partial discharge pattern recognition based on time-frequency analysis and 2DNMF[J]. Electric Power Automation Equipment, 2013, 33(3): 20-25.
- [4] 汪可,廖瑞金,王季宇,等. 局部放电 UHF 脉冲的时频特征提取与聚类分析[J]. 电工技术学报,2015,30(2):211-219.
WANG Ke, LIAO Ruijin, WANG Jiyu, et al. Time-frequency features extraction and clustering analysis of partial discharge UHF pulses[J]. Transactions of China Electrotechnical Society, 2015, 30(2): 211-219.
- [5] 靖小平,彭小圣,姜伟,等. 基于 K-Means 聚类算法的自动图谱识别在电缆局部放电在线监测系统中的应用[J]. 高电压技术,2012,38(9):2437-2446.
JING Xiaoping, PENG Xiaosheng, JIANG Wei, et al. Phase resolved partial discharge pattern recognition method for on-line cable condition monitoring system based on K-Means clustering[J]. High Voltage Engineering, 2012, 38(9): 2437-2446.
- [6] 陈攀,姚陈果,廖瑞金,等. 分频段能量谱及马氏聚类算法在开关柜局部放电模式识别中的应用[J]. 高电压技术,2015,41(10):3332-3341.
CHEN Pan, YAO Chenguo, LIAO Ruijin, et al. Application of signals separated band energy spectrum and mahalanobis clustering algorithm for switchgear partial discharge pattern recognition[J]. High Voltage Engineering, 2015, 41(10): 3332-3341.
- [7] 王瑜,苑津莎,尚海昆,等. 组合核支持向量机在放电模式识别中的优化策略[J]. 电工技术学报,2015,30(2):229-236.
WANG Yu, YUAN Jinsha, SHANG Haikun, et al. Optimization strategy research on combined-kernel support vector machine for partial discharge pattern recognition[J]. Transactions of China Electrotechnical Society, 2015, 30(2): 229-236.

- [8] 罗新,牛海清,来立永,等. 粒子群优化自适应小波神经网络在带电局放信号识别中的应用[J]. 电工技术学报, 2014, 29(10): 326-333.
LUO Xin, NIU Haiqing, LAI Liyong, et al. Application of adaptive wavelet neural network based on particle swarm optimization algorithm in online PD pattern recognition[J]. Transactions of China Electrotechnical Society, 2014, 29(10): 326-333.
- [9] 律方成,金虎,王子建,等. 基于主成分分析和多分类相关向量的GIS局部放电模式识别[J]. 电工技术学报, 2015, 30(6): 225-231.
LÜ Fangcheng, JIN Hu, WANG Zijian, et al. GIS partial discharge pattern recognition based on principal component analysis and multi-class relevance vector machine[J]. Transactions of China Electrotechnical Society, 2015, 30(6): 225-231.
- [10] HAO L, LEWIN P L. Partial discharge source discrimination using a support vector machine[J]. IEEE Transactions on Dielectrics and Electrical Insulation, 2010, 17(1): 189-197.
- [11] 尚海昆,苑津莎,王瑜,等. 多核多分类相关向量机在变压器局部放电模式识别中的应用[J]. 电工技术学报, 2014, 29(11): 221-228.
SHANG Haikun, YUAN Jinsha, WANG Yu, et al. Partial discharge pattern recognition in power transformer based on multi-kernel multi-class relevance vector machine[J]. Transactions of China Electrotechnical Society, 2014, 29(11): 221-228.
- [12] MAJIDI M, FADALI M S, ETEZADI-AMOLI M. Partial discharge pattern recognition via sparse representation and ANN[J]. IEEE Transactions on Dielectrics and Electrical Insulation, 2015, 22(2): 1061-1070.
- [13] 张淑清,孙国秀,李亮,等. 基于LMD近似熵和FCM聚类的机械故障诊断研究[J]. 仪器仪表学报, 2013, 34(3): 714-720.
ZHANG Shuqing, SUN Guoxiu, LI Liang, et al. Study on mechanical fault diagnosis method based on LMD approximate entropy and fuzzy C-means clustering[J]. Chinese Journal of Scientific Instrument, 2013, 34(3): 714-720.
- [14] LEI Yaguo, HE Zhengjia, ZI Yanyang, et al. New clustering algorithm-based fault diagnosis using compensation distance evaluation technique[J]. Mechanical Systems and Signal Processing, 2008, 22(2): 419-435.
- [15] 王刘旺,朱永利,李莉,等. 基于自适应双阈值的局部放电基本参数提取[J]. 高电压技术, 2016, 42(4): 1268-1274.
WANG Liuwang, ZHU Yongli, LI Li, et al. Extraction of fundamental parameters in partial discharge based on adaptive dual threshold[J]. High Voltage Engineering, 2016, 42(4): 1268-1274.
- [16] 朱永利,贾亚飞,王刘旺,等. 基于改进变分模态分解和Hilbert变换的变压器局部放电信号特征提取及分类[J]. 电工技术学报, 2017, 32(9): 221-235.
ZHU Yongli, JIA Yafei, WANG Liuwang, et al. Feature extraction and classification on partial discharge signals of power transformers based on improved variational mode decomposition and Hilbert transform[J]. Transactions of China Electrotechnical Society, 2017, 32(9): 221-235.
- [17] OTSU N. A threshold selection method from gray-level histograms[J]. IEEE Transactions on Systems Man & Cybernetics, 2007, 9(1): 62-66.
- [18] 贾亚飞,朱永利,王刘旺,等. 基于VMD和多尺度熵的变压器内绝缘局部放电信号特征提取及分类[J]. 电工技术学报, 2016, 31(19): 208-217.
JIA Yafei, ZHU Yongli, WANG Liuwang, et al. Feature extraction and classification on partial discharge signals of power transformers based on VMD and multiscale entropy[J]. Transactions of China Electrotechnical Society, 2016, 31(19): 208-217.

作者简介:



贾亚飞

贾亚飞(1988—),女,河北望都人,博士研究生,主要研究方向为高压电气设备局部放电信号分析和故障诊断(E-mail: jiaayafeiyanshan@163.com);

朱永利(1963—),男,河北冀州人,教授,博士研究生导师,主要研究方向为网络化监控与智能信息处理(E-mail: yonglipw@

163.com);

高佳程(1993—),男,河北高阳县人,硕士研究生,主要研究方向为电气设备故障诊断(E-mail: gaojiacheng1993@163.com)。

Recognition of unknown partial discharge signals based on sample-weighted FCM clustering

JIA Yafei^{1,2}, ZHU Yongli¹, GAO Jiacheng¹, YUAN Bo³

(1. State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, North China Electric Power University, Baoding 071003 China; 2. Xiongan New Area

Electric Power Supply Company, State Grid Hebei Electric Power Supply Co., Ltd., Xiongan New Area 071800, China;

3. Economic and Technology Research Institute, State Grid Hebei Electric Power Company, Shijiazhuang 050000, China)

Abstract: In the power transformers' PD(Partial Discharge) signals to be identified, there may be samples belonging to unknown types, to solve this problem, a recognition method of unknown PD signals based on sample-weighted FCM(Fuzzy C-Means) clustering is proposed. The PD signals of known types are clustered by FCM clustering to determine the clustering centers of known types. The PD signal sample weights of known types and the signals to be identified are calculated respectively, and the adaptive thresholds of sample weights are determined by Otsu criterion. Then, the weights of the PD signals to be identified are compared with the determined thresholds to judge whether they belong to known PD types. The PD signals which belong to known types are classified by SVM(Support Vector Machine) and the PD signals of unknown types are artificially analyzed and judged. The PD signals under laboratory conditions are analyzed by the proposed method and experimental results show that the proposed method can effectively recognize the PD signals with unknown types from the PD signals to be identified.

Key words: power transformers; partial discharge; pattern recognition; unknown sample; sample-weighted; FCM clustering; Otsu criterion; support vector machines

附录

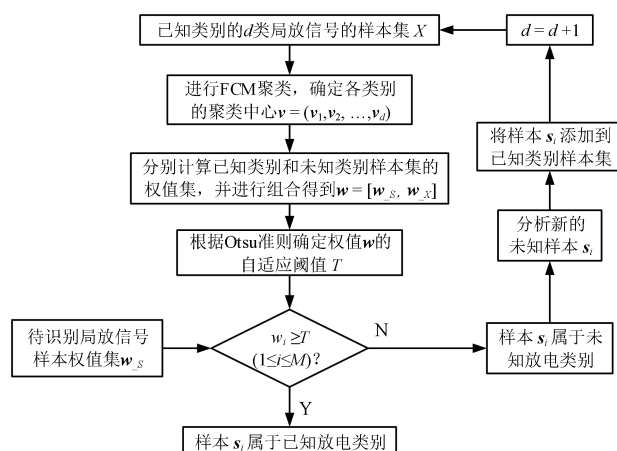
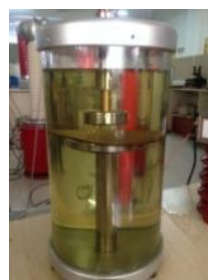


图 A1 基于样本加权 FCM 聚类的未知类别局部放电信号识别方法流程图

Fig.A1 Flowchart of recognition method for unknown partial discharge signal types based on sample weighting FCM clustering



(a) 电晕放电模



(b) 油中多尖对板放电模型



(c) 油中板对板放电模型



(d) 油中悬浮颗粒放电模型



(e) 锥对板放电模型

图 A2 局部放电模型

Fig.A2 Four types of partial discharge model