

基于时频分析和 2DNMF 的局部放电模式识别

廖瑞金,段 炼,汪 可,杨丽君

(重庆大学 输配电装备及系统安全与新技术国家重点实验室,重庆 400044)

摘要: 提出时频分析结合二维非负矩阵分解的混合特征提取算法识别不同局部放电类型。在实验室环境下采集了4种典型绝缘缺陷模型的局部放电超高频(UHF)波形,引入自适应最优径向高斯核时频分析挖掘局部放电 UHF 信号的时频信息,在对时频幅值矩阵进行二维非负矩阵分解提取降维特征后,采用模糊 k -近邻分类器对4种不同类型的局部放电信号进行识别。对试验样本的识别结果表明:自适应最优径向高斯核时频分布能较好地表征局部放电单次波形的时频信息;二维非负矩阵分解降维后的特征矩阵能保存原始时频矩阵的大部分有用信息;模糊 k -近邻分类器比 k -近邻分类器和3层反向传播神经网络具有更高的识别率,并较反向传播神经网络具有容易拓展的优点。

关键词: 绝缘;局部放电;模式识别;时频分析;二维非负矩阵分解;模糊 k -近邻分类器

中图分类号: TM 835

文献标识码: A

DOI: 10.3969/j.issn.1006-6047.2013.03.004

0 引言

电力设备制造和运行中产生的尖端、气隙等绝缘缺陷会引起绝缘局部场强集中,进而导致绝缘局部击穿并引发局部放电 PD(Partial Discharge)。绝缘系统在持续的 PD 作用下会逐渐劣化,而不同缺陷类型的 PD 对绝缘的危害程度及损伤机理不同,使基于 PD 信号分析的电气设备绝缘诊断及运行风险评估等项目面临较大难度^[1-2]。因此,识别 PD 缺陷类型对于电力设备绝缘状态的诊断和评估具有重要的理论意义和实用价值。

PD 超高频 UHF(Ultra High Frequency)天线测量技术因能有效避开低频的电磁干扰、更适用于现场应用而获得了较快发展^[3]。PD UHF 信号是典型的非平稳信号,具有时变的频谱。传统的傅里叶变换无法得到非平稳信号某种频率分量的发生时间及其变化的情况;而时频分析 TFA(Time-Frequency Analysis)方法能在时频平面上表征放电信号随时间变化的频率和幅值,因此采用时频分析方法来挖掘 PD 电/声脉冲信号的时/频域联合信息成为了近年来的研究热点^[4-9]。其中,刘尚合^[7]等应用时频分析对实际 PD UHF 信号进行处理,结果表明时频谱能准确反映 PD 信号的时频特性,有利于对信号特征的提取;郭灿新^[8]等对仿真 PD 信号和现场 PD 信号进行时频分析,结果显示时频分析方法可有效获取 PD 脉冲信息;X. Ma^[9]等通过对 PD UHF 信号的小波变换分析,结果显示了时频分析方法在 UHF 信号处理的

有效性,但也指出了单一的小波基无法对各种 PD 类型均适用,有必要研究自适应的小波基来智能地处理 PD 信号。针对时频分析应用于 PD UHF 信号模式识别,相关文献鲜有报道,但文献^[7,10]均指出时频分析有利于 PD 信号的特征提取,为其模式识别提供了一个准确、可靠的新方法。针对以上结果,本文引入自适应最优径向高斯核 AORGK(Adaptive Optimal Radially Gaussian Kernel)时频分析方法,它采用短时模糊函数 STAF(Short-Time Ambiguity Function)和随时间变化的自适应核函数,可以有效分离自分量信号和互分量信号,并具有良好的时域和频域分辨率,从而能较好地克服常规时频分析方法的缺点^[11],准确地表征 PD 脉冲的时频信息。

时频分析的结果为一个二维时频矩阵,直接用于 PD 模式识别时,会由于特征维数过高而无法在短时间内完成分类,并且可能由于时频特征之间的相关性影响分类器的识别效果^[12]。因此,在进行分类之前必须对时频矩阵进行降维处理。非负矩阵分解 NMF(Non-negative Matrix Factorization)降维方法的最大特点是通过非负性约束不允许矩阵中出现负数元素,在直观上具有“局部构成整体”的特性(纯“加”性计算)^[13]。然而,在实际应用 NMF 时,一般需要事先将二维矩阵转换为一维向量,这不仅会导致高维的向量空间,难以选择优良的基矩阵去表征原始矩阵,而且可能导致在转换过程中原始矩阵隐藏的结构信息丢失。为克服 NMF 的不足,文献^[14]提出二维非负矩阵分解 2DNMF(Two-Dimensional Non-negative Matrix Factorization)。2DNMF 直接对矩阵在列和行方向上同时进行投影运算,不仅保留了矩阵的原始结构信息,而且大幅降低了计算量,更适合对高维矩阵进行处理。

收稿日期:2012-02-23;修回日期:2013-01-04

基金项目:国家自然科学基金资助项目(51277187);国家重点基础研究发展计划(973计划)项目(2009CB724505-1)

Project supported by the National Natural Science Foundation of China(51277187) and the State Key Development Program for Basic Research of China(973 Program)(2009CB724505-1)

本文首次提出采用 AORGK 时频分析结合 2DNMF 的混合算法提取 PD 特征参量用于识别缺陷类型。首先在实验室环境下采集了 4 种典型的绝缘缺陷模型的 PD UHF 脉冲波形。然后,利用 AORGK 时频分布表征 PD 的时频信息,并应用 2DNMF 对时频矩阵进行降维提取特征参量。最后采用 3 种分类器校验了该方法的有效性。

1 AORGK 时频分析

1.1 最优径向高斯核时频分析

信号 $s(t)$ 的最优径向高斯核 ORGK(Optimal Radially Gaussian Kernel) 时频分布 $P(t, w)$ 定义为^[15]:

$$P_{\text{ORGK}}(t, w) = \frac{1}{4\pi^2} \iint A(\theta, \tau) \phi_{\text{opt}}(\theta, \tau) e^{-j\theta t - j\tau w} d\theta d\tau \quad (1)$$

其中, t, w 分别为时间和频率; j 为虚数单位; $A(\theta, \tau)$ 为模糊函数; $\phi_{\text{opt}}(\theta, \tau)$ 为最优径向高斯核函数; θ, τ 分别为频率偏差和延迟时间, 构成模糊平面的横坐标和纵坐标。 $A(\theta, \tau)$ 的表达式如下:

$$A(\theta, \tau) = \int_{-\infty}^{+\infty} s^*\left(t + \frac{\tau}{2}\right) s\left(t + \frac{\tau}{2}\right) e^{j\theta t} dt \quad (2)$$

其中, $*$ 表示复共轭。径向高斯核 $\phi(\theta, \tau)$ 定义为沿任意剖面都是 Gauss 型的二维函数, 表达式如下:

$$\phi(\theta, \tau) = \exp\left[-\frac{\theta^2 + \tau^2}{2\sigma^2(\psi)}\right] \quad (3)$$

其中, $\sigma(\psi)$ 为控制径向高斯核函数在径向角 ψ 方向的扩展函数, 径向角 $\psi = \arctan(\tau/\theta)$ 为径向与水平方向的夹角。为了得到高质量的时频分布结果, 要求核函数与信号相匹配, 即求解最优的径向高斯核 $\phi_{\text{opt}}(\theta, \tau)$ 。求解过程即实现以下最优化问题:

$$\max_{\phi} \int_0^{2\pi} \int_0^{+\infty} |A(r, \psi) \phi(r, \psi)|^2 r dr d\psi \quad (4)$$

约束条件为:

$$\begin{aligned} \phi(r, \psi) &= \exp\left[-\frac{r^2}{2\sigma^2(\psi)}\right] \\ \frac{1}{4\pi^2} \int_0^{2\pi} \int_0^{+\infty} |\phi(r, \psi)|^2 r dr d\psi &= \\ \frac{1}{4\pi^2} \int_0^{2\pi} \sigma^2(\psi) d\psi &\leq \alpha \quad \alpha \geq 0 \end{aligned} \quad (5)$$

其中, $A(r, \psi)$ 、 $\phi(r, \psi)$ 为极坐标形式的模糊函数和径向高斯核函数, $r = \sqrt{\theta^2 + \tau^2}$ 。式(4)中的模糊函数能最大限度地区分信号的自分量和互分量, 即自分量信号集中在原点附近, 而互分量信号远离原点。式(5)中的约束 1 为低通的径向高斯核, 能保证自分量信号通过; 约束 2 通过核函数体积 α 的设置来抑制交叉项的干扰, 一般取 $1 \leq \alpha \leq 5$ ^[16]。上述优化过程是抑制交叉项干扰并保证自分量信号的通过, 从而实现信号的最小畸变。

1.2 AORGK 时频分析

AORGK 时频分析^[11]的核函数随时间自适应变化, 在刻画信号细节方面优于基于信号的径向高斯核时频分布, 并能较好地抑制交叉项干扰。为得到随时间变化的自适应核函数, 首先计算 STAF:

$$A(t; \theta, \tau) = \int_{-\infty}^{+\infty} s^*\left(u + \frac{\tau}{2}\right) w^*\left(u - t - \frac{\tau}{2}\right) \times s\left(u + \frac{\tau}{2}\right) w\left(u - t - \frac{\tau}{2}\right) e^{j\theta u} du \quad (6)$$

其中, t 为对称窗函数 $w(u)$ 的中心位置, 令 $|u| > T$ 时 (T 为时间窗宽度), 对称窗函数 $w(u) = 0$ 。由此可知, $A(t; \theta, \tau)$ 只计算以时间 t 为中心的 $[t-T, t+T]$ 范围内的信号, 即 STAF 可以准确刻画出信号的任意细节部分。对于每个时刻 t , 在 STAF 的基础上再计算相应的 AORGK $\phi_{\text{opt}}(t; \theta, \tau)$, 求解过程类似于式(4)和式(5)。由此可以得到如式(7)所示的信号 $s(t)$ 的 AORGK 时频分布的计算公式。

$$P_{\text{AORGK}}(t, w) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} A(t; \theta, \tau) \phi_{\text{opt}}(t; \theta, \tau) e^{-j\theta t - j\tau w} d\theta d\tau \quad (7)$$

AORGK 时频分析的结果是一个二维时频矩阵, 该矩阵每行表征该频率点处信号随时间变化的分布规律, 每列则表征该时刻信号随频率变化的分布规律。本文采用 AORGK 时频幅值矩阵用于表征 PD UHF 脉冲波形的时频信息, 如式(8)所示。

$$A_{\text{AORGK}}(t, w) = |P_{\text{AORGK}}(t, w)| \quad (8)$$

2 二维 NMF

2.1 NMF

对给定的一个非负矩阵 $V_{n \times m}$, NMF 算法可将其表示为 2 个非负矩阵 $W_{n \times r}$ 和 $H_{r \times m}$ 的乘积:

$$V_{n \times m} \approx W_{n \times r} H_{r \times m} \quad (9)$$

其中, W 和 H 分别称为基矩阵和系数矩阵。当 r 满足 $(n+m)r < nm$ 时, 即可实现数据压缩。矩阵 W 和矩阵 H 的求解是逐步减少矩阵 V 和矩阵 WH 的差别的过程, 如式(10)所示。

$$\min_{W, H} f(W, H) = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^m [V_{ij} - (WH)_{ij}]^2 \quad (10)$$

$$W_{ia} \geq 0, H_{bj} \geq 0, \forall i, a, b, j$$

本文采用投影梯度方法 PGM(Projected Gradient Methods) 进行 NMF 迭代运算, PGM 具有优异的收敛速度和收敛结果, 具体算法详见参考文献[17]。

2.2 二维 NMF

NMF 处理的对象一般是一维的向量, 而 2DNMF 可直接对高维矩阵进行列和行方向上的压缩^[14], 具体步骤如下。

步骤 1 计算列基矩阵。矩阵 $X = [A_1, A_2, \dots, A_m]$

由 m 个 $p \times q$ 维非负时频幅值矩阵构成, A_k 表征第 k 个样本。类似于 NMF, 2DNMF 将矩阵 X 分解为 $p \times d_1$ 的非负矩阵 L 和 $d_1 \times qm$ 的非负矩阵 H , 如式(11)所示。

$$X_{p \times mq} \approx L_{p \times d_1} H_{d_1 \times mq} \quad (11)$$

其中, X 的每一列都对应着原始时频矩阵的每一列, 故称 L 为列基矩阵。

步骤 2 计算行基矩阵。参照列基矩阵, 构造一个 $q \times pm$ 的矩阵 $X' = [A'_1, A'_2, \dots, A'_m]$, 其中 A'_k 为 A_k 的转置。同理, X' 可表示为如下形式:

$$X'_{q \times pm} \approx R_{q \times d_2} C_{d_2 \times pm} \quad (12)$$

其中, 非负矩阵 R, C 的大小分别为 $q \times d_2, d_2 \times pm$ 维。 X' 的每一列都对应着原始时频矩阵的每一行, 故称 R 为行基矩阵。迭代过程采用 PGM 得到列基矩阵 L 和行基矩阵 R 。

步骤 3 提取特征矩阵。将原始时频矩阵 A_k 向列基矩阵的转置和行基矩阵组成的二维基上进行投影, 得到降维后的时频特征矩阵 D_k , 如下式所示:

$$D_k = L_k^T A_k R \quad (13)$$

其中, D_k 的维数为 $d_1 \times d_2$ 。与 NMF 算法相比, 2DNMF 提取的特征矩阵 D_k 保存了原始时频矩阵的行和列结构上的信息, 并且当选择 d_1, d_2 远小于 p, q 时, 特征维数得到了很大程度的约减, 更利于实现样本的快速分类。

3 模糊 k -近邻分类器

k -近邻法 kNN(k -Nearest Neighbor)是最近邻法 NN(Nearest Neighbor)的推广, 它可以描述为在已知类别表示的样本中, 找出未知样本 x 的 k (奇数) 个近邻, 然后把未知样本归为近邻样本出现频率最高的那一类。

kNN 分类器的固有缺陷是认为其 k 个近邻重要性相当而不考虑它们的差异性。模糊 k -近邻法 (FkNN) 采用模糊分类思想, 通过引入隶属度函数来克服这一缺陷, 能对未知样本隶属于每一类的程度进行量化。在得到待分类样本 x 的 k 个近邻 x_j 后, 可用式(14)计算样本 x 对于每一类的隶属度值^[18]。

$$\mu_i(x) = \frac{\sum_{j=1}^k \mu_i(x_j) / \{ [1 / \|x - x_j\|^{2/(b-1)}] \}}{\sum_{j=1}^k \{ [1 / \|x - x_j\|^{2/(b-1)}] \}} \quad (14)$$

其中, $\mu_i(x_j)$ 为已知类别样本 x_j 对第 i 类的隶属度值, $\mu_i(x)$ 为未知样本 x 隶属于第 i 类的程度。参数 b 用来决定对距离加权的程度: 当 $b=2$ 时, 每个近邻按距离的倒数分配权重系数; 当 b 增大时, 每个近邻的权重系数比较平均; 当 b 接近于 1 时, 最近邻给予的权重明显增大。

4 PD 模式识别

4.1 PD 模型及数据采集

在实验室中开展了固体绝缘内部缺陷(P1)、油中沿面放电(P2)、油中悬浮电极放电(P3)和油中电晕放电(P4)4种典型缺陷模型的 PD 试验, 电极结构参照 CIGRE II 方法电极结构制作, 如图 1 所示。图 1(a)为模拟绝缘内部 PD 的电极系统及缺陷模型结构, 采用绝缘薄膜在板电极表面支撑环氧板形成厚度为 0.15 mm、直径为 38 mm 的气隙。为防止变压器油渗入人工气隙, 还使用了 2 层环氧板将气隙模型夹住。图 1(b)、(c)和(d)分别是柱板放电系统、悬浮放电系统、针-板放电系统, 其中悬浮电极中的环氧板边缘放置有直径为 0.3 mm 的金属球颗粒, 针电极尖部与环氧板距离为 1 mm。试验中将 4 个模型均浸入绝缘油中。

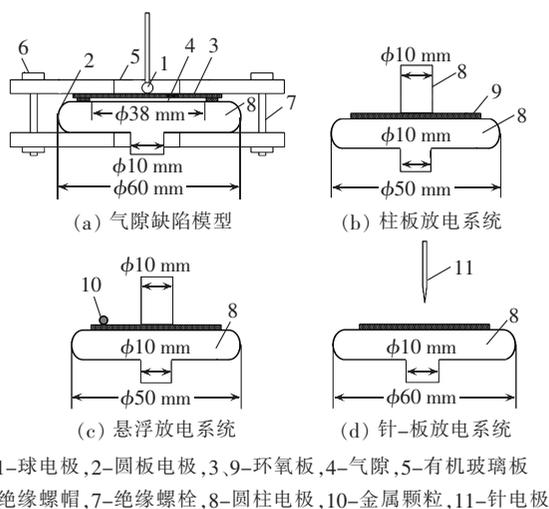


图 1 PD 人工缺陷模型

Fig.1 Artificial defect models of PD

试验中采用 4 阶 Hilbert 分形天线测量 PD 的 UHF 信号, UHF 传感器距局部缺陷约 15 cm, 信号经带宽为 300~1000 MHz 放大器后接入 Wavepro7100 示波器中进行显示和存储, 数据每秒采样 5×10^9 个点。采样电压为起始放电电压的 1.3~1.5 倍, 对每类缺陷下的 3 个电压等级分别采集了 50 个放电样本, 共采集了 600 组 PD 的 UHF 波形数据, 原始 PD 数据样本如表 1 所示。为减小油中水分对 PD 的影响, 本文采用未受潮的新油进行试验, 并且每次试验结束后都对容器密封以防止受潮。

4.2 PD 数据预处理

单次 PD 脉冲持续时间一般介于几十 ns 到数百 ns 之间^[19-20], 试验采集了 1 000 ns 间隔内的脉冲波形, 如图 2 所示。

参照文献[10], 从原始脉冲波形中提取能完全表征单次脉冲波形的 1 000 个采样点, 油中悬浮电极放电(P3)预筛选前、后的脉冲波形如图 3 所示。

表 1 原始 PD 数据样本
Tab.1 Original PD data samples

PD 类别	起始放电电压/kV	采样电压/kV	样本数
P1	5.2~5.6	7	50
		8	50
		9	50
P2	6.9~7.2	9	50
		10	50
		11	50
P3	7.4~8.2	10	50
		11	50
		12	50
P4	6.6~7.1	9	50
		10	50
		11	50

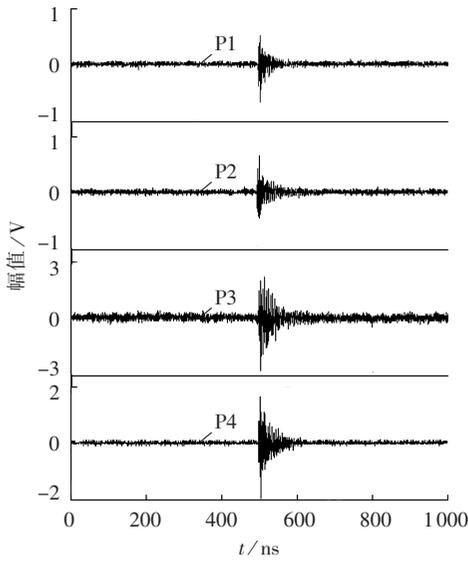


图 2 PD UHF 原始脉冲波形
Fig.2 Primitive PD UHF signals

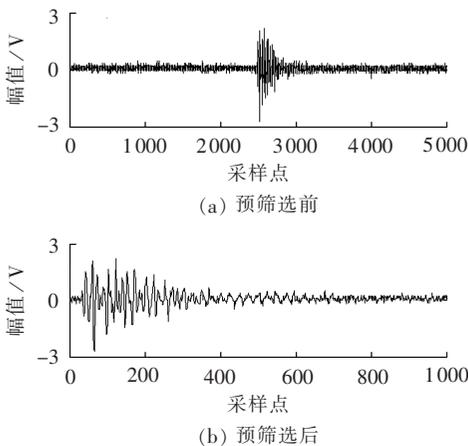


图 3 PD UHF 信号的预筛选
Fig.3 Pre-selection of PD UHF signals

最后,为有效降低白噪声和波形分散性对后续特征提取的影响,采用式(15)对数据进行归一化处理:

$$x_n(i) = \frac{x(i)}{\max(|x(i)|)} \quad 1 \leq i \leq 1000 \quad (15)$$

其中, $\max(|x(i)|)$ 是采样点中的模值最大者, $x_n(i)$

是 $x(i)$ 对应的归一化数据。

4.3 PD 特征提取

图 4 为 4 种类型 PD UHF 信号的时频分布谱图,由图可知不同缺陷类型对应的谱图存在一定差异,表明获取的时频分布矩阵包含了不同 PD 类型信息。由于原始时频矩阵的维数为 512×1000 (512 为脉冲波形信号快速傅里叶变换的采样点数),无法直接用于分类识别,需要对时频矩阵进行降维处理。

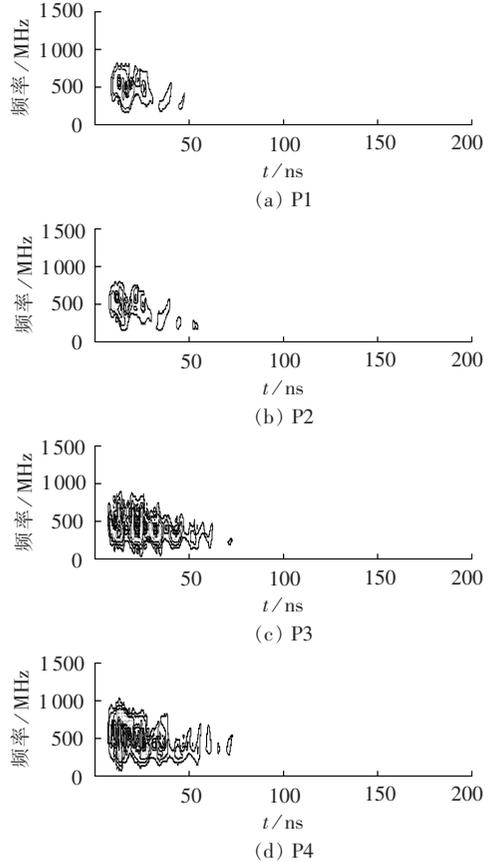


图 4 不同缺陷类型 PD 信号的时频分布
Fig.4 Time-frequency representation of PD signals for different defect models

参照 2.2 节的降维算法,从 600 个 512×1000 维时频矩阵中随机选取 24 个(每类放电类型 4 个)作为时频训练矩阵;通过对 24 个样本矩阵进行 2DNMF 运算,即可得到 600 个 $d_1 \times d_2$ 维的特征矩阵。为讨论不同 d_1 和 d_2 的取值对后续分类结果的影响,本文选取了 (d_1, d_2) 为 $(5, 5)$ 、 $(5, 10)$ 、 $(10, 5)$ 、 $(10, 10)$ 的 4 种组合方式。

4.4 分类识别结果与讨论

在测试阶段,随机选取每种放电类型样本的一半作为训练样本,另一半作为待识别样本;并选取 10 次识别率的平均值作为最后的识别结果。测试结果表明 $b=2$ 能够取得较好的 PD 识别率,即相似度采用欧氏距离度量,因此后续测试过程中 b 值选为 2。表 2 为不同 k 值下 kNN 分类器和 FkNN 分类器的识别

率,其中 (d_1, d_2) 取 $(10, 5)$ 。从表 2 可知,在 k 值从 1 到 11 变化的过程中, FkNN 的识别率均在 85% 以上,高于 kNN 的识别率。当 $k=3$ 时, kNN 取得最好的识别效果;当 $k=7$ 时, FkNN 的识别率最高。

表 2 不同 k 值的 PD 信号的识别效果
Tab.2 Classification results of PD signals
with different k values

k	识别率/%		k	识别率/%	
	kNN	FkNN		kNN	FkNN
1	81.00	85.23	7	78.63	93.73
3	84.07	87.73	9	75.27	86.97
5	77.57	89.43	11	72.67	85.93

3 层反向传播神经网络 BPNN (Back Propagation Neural Network) 是以往 PD 模式识别中应用最广泛的分类器。表 3 对比了不同约减维数的 kNN、FkNN 和 BPNN 的识别率,其中 k 参照表 2 选择识别率最高的奇数值。从识别结果可以看出,4 种不同约减维数的组合中, (d_1, d_2) 取 $(10, 5)$ 时 kNN、FkNN、BPNN 均取得了最好的识别效果。究其原因, $(5, 5)$ 组合特征维数虽然较低,更利于分类器的设计,但损失了较多的原始信息,导致识别率不太理想; $(10, 10)$ 组合虽然保留了更多的原始信息,但是由于特征维数较高,在降维的过程中也引入了一定的干扰和噪声,使识别率受到一定的影响。 $(5, 10)$ 组合和 $(10, 5)$ 组合的特征矩阵由于计算过程并不等效,使得两者的识别率存在差异。由此可知,采用 2DNMF 提取时频矩阵的特征时,需要综合考虑两方面的因素:保证降维后的特征维数尽可能地小以减少分类器的压力;在不引入干扰及冗余信息的同时最大限度地保留原始矩阵的有效信息。

除此之外,从表 3 还可以看出, FkNN 分类器较 BPNN 具有更好的识别效果。并且, FkNN 较 BPNN 还具有一个优势:当增加新的放电类型样本时, FkNN 只需要增加一类放电的训练样本即可进行分类识别,而不需要像 BPNN 那样重新调整网络结构并对整个网络进行训练,具有较好的拓展性。

表 3 不同约减维数的 PD 信号的识别效果
Tab.3 Classification results of PD signals
with different compressed dimensions

(d_1, d_2)	识别率/%		
	kNN	FkNN	BPNN
$(5, 5)$	81.07	87.50	83.57
$(5, 10)$	82.20	91.47	84.70
$(10, 5)$	84.07	93.73	89.37
$(10, 10)$	83.50	90.73	88.67

5 结论

本文重点研究 PD UHF 信号的特征提取算法与分类识别策略,主要得到以下结论:

a. 应用 AORGK 获取 PD 脉冲波形的时频分布谱图,有效地表征并区分了 4 种绝缘缺陷类型的 PD 信号;

b. 与 NMF 相比, 2DNMF 对高维时频矩阵在列和行 2 个方向上投影,不仅解决了维数危机,也较好地保留了原始时频矩阵的结构信息;

c. 4 种 (d_1, d_2) 组合中, $(10, 5)$ 组合的识别率最高,并且 FkNN 分类器的识别效果在不同降维组合下均优于 kNN 和 BPNN 分类器;

d. 当增加新的放电类型时, FkNN 只需要添加新类型的训练样本即可进行分类识别,而无需像 BPNN 那样重新训练网络,具有较好的拓展性。

需要指出的是,对于 PD UHF 信号,本文提出的混合算法在实验室条件下取得了较好的识别效果,为电力设备绝缘状态的诊断和评估提供有益参考。但现场条件下 UHF 信号受信号传播的路径、传感器的位置、传感器和放大器的性能等诸多因素的影响,现场应用仍需进一步研究。

参考文献:

- [1] 杜林, 颜梁钦, 甘德刚, 等. 超高频扫描比较式局部放电在线检测系统[J]. 电力自动化设备, 2012, 32(6): 120-124.
DU Lin, YAN Liangqin, GAN Degang, et al. Online partial discharge monitoring system by UHF scanning and comparison[J]. Electric Power Automation Equipment, 2012, 32(6): 120-124.
- [2] 李剑, 宁佳欣, 金卓睿, 等. 变压器局部放电在线监测超高频 Hilbert 分形天线研究[J]. 电力自动化设备, 2007, 27(6): 31-35.
LI Jian, NING Jiaxin, JIN Zhuorui, et al. Research on UHF Hilbert fractal antenna for online transformer PD monitoring[J]. Electric Power Automation Equipment, 2007, 27(6): 31-35.
- [3] CLEARY G P, JUDD M D. UHF and current pulse measurements of partial discharge activity in mineral oil[J]. IEE Proceedings on Science, Measurement and Technology, 2006, 153(2): 47-54.
- [4] 王辉, 黄成军, 姚林朋, 等. 重排 Cohen 类时频分布用于 GIS 局部放电声信号时频分析[J]. 高电压技术, 2010, 36(9): 2236-2241.
WANG Hui, HUANG Chengjun, YAO Linpeng, et al. Time-frequency analysis of partial discharge for GIS using the theory of reassignment distribution[J]. High Voltage Engineering, 2010, 36(9): 2236-2241.
- [5] BOCZAR T, ZMARZLY D. Application of wavelet analysis to acoustic emission pulses generated by partial discharges[J]. IEEE Transactions on Dielectrics and Electrical Insulation, 2004, 11(3): 433-449.
- [6] CHAI M L, THAYOOB Y H M, GHOSH P S, et al. Identification of different types of partial discharge sources from acoustic emission signals in the time-frequency representation[C]//IEEE Conference on Power and Energy. Putra Jaya, Malaysia: IEEE, 2006: 581-586.
- [7] 刘卫东, 刘尚合, 王雷. 采用 Gabor 变换的局部放电信号时频分析[J]. 高电压技术, 2007, 33(8): 40-43.
LIU Weidong, LIU Shanghe, WANG Lei. Time-frequency analysis of partial discharge signal by gabor transform[J]. High Voltage

- Engineering, 2007, 33(8):40-43.
- [8] 郭灿新, 勇明, 徐敏骅, 等. S变换在电力电缆局部放电信号时频分析中的应用[J]. 电工技术学报, 2010, 25(11):9-14.
- GUO Canxin, YONG Ming, XU Minhua, et al. Application of S transform for time-frequency analysis of partial discharge signals in power cables[J]. Transactions of China Electrotechnical Society, 2010, 25(11):9-14.
- [9] MA X, ZHOU C, KEMP I J. Interpretation of wavelet analysis and its application in partial discharge detection[J]. IEEE Transactions on Dielectrics and Electrical Insulation, 2002, 9(3):446-457.
- [10] YANG L, JUDD M D. Recognizing multiple partial discharge sources in power transformers by wavelet analysis of UHF signals[J]. IEE Proceedings on Science, Measurement and Technology, 2003, 150(3):119-126.
- [11] JONES D L, BARANIUK R G. An adaptive optimal-kernel time-frequency representation[J]. IEEE Transactions on Signal Processing, 1995, 43(10):2361-2371.
- [12] 杨丽君, 廖瑞金, 孙才新, 等. 矢量量化在局部放电模式识别中的应用[J]. 中国电机工程学报, 2009, 29(31):122-127.
- YANG Lijun, LIAO Ruijin, SUN Caixin, et al. Application of vector quantization to partial discharge pattern recognition[J]. Proceedings of the CSEE, 2009, 29(31):122-127.
- [13] LEE D D, SEUANG H S. Learning the parts of objects by non-negative matrix factorization[J]. Nature, 1999, 401(6755):788-791.
- [14] LI Bing, ZHANG Peilin, LIU Dongsheng, et al. Feature extraction for rolling element bearing fault diagnosis utilizing generalized S transform and two-dimensional non-negative matrix factorization[J]. Journal of Sound and Vibration, 2011, 330(10):2388-2399.
- [15] BARANIUK R G, JONES D L. A radially Gaussian, signal-dependent time-frequency representation[J]. IEEE Transactions on Signal Processing, 1993, 32(10):263-284.
- [16] BARANIUK R G, JONES D L. A radially Gaussian, signal-dependent time-frequency representation[C]//IEEE Conference on Signal Processing, Toronto, Canada:IEEE, 1991:3181-3184.
- [17] LIN C J. Projected gradient methods for non-negative matrix factorization[J]. Neural Computation, 2007, 19(10):2756-2779.
- [18] KELLER J M, GRAY M R, GIVENS J A. A fuzzy k -nearest neighbor algorithm[J]. IEEE Transactions on Systems, Man, and Cybernetics, 1985, 15(4):580-585.
- [19] CHANG C S, JIN J, CHANG C, et al. Online source recognition of partial discharge for gas insulated substations using independent components analysis[J]. IEEE Transactions on Dielectrics and Electrical Insulation, 2006, 13(4):892-902.
- [20] 郭俊, 吴广宁, 张血琴, 等. 局部放电检测技术的限制和发展[J]. 电工技术学报, 2005, 20(2):29-35.
- GUO Jun, WU Guangning, ZHANG Xueqin, et al. The actuality and prespective of partial discharge detection techniques[J]. Transactions of China Electrotechnical Society, 2005, 20(2):29-35.

作者简介:

廖瑞金(1963-), 男, 四川遂宁人, 教授, 长江学者特聘教授, 博士研究生导师, 长期从事电气设备状态监测与故障智能诊断技术、高电压测试技术、变压器老化与诊断技术研究工作(E-mail: rjliao@cqu.edu.cn);

段炼(1987-), 男, 湖北黄冈人, 硕士研究生, 主要研究方向为电气设备绝缘在线监测及其故障诊断(E-mail: dylan.duan2008@gmail.com)。

Partial discharge pattern recognition based on time-frequency analysis and 2DNMF

LIAO Ruijin, DUAN Lian, WANG Ke, YANG Lijun

(State Key Laboratory of Power Transmission Equipment & System Security and New Technology, Chongqing University, Chongqing 400044, China)

Abstract: A hybrid feature extraction algorithm based on TFA(Time-Frequency Analysis) combined with 2DNMF(Two-Dimensional Non-negative Matrix Factorization) is proposed to identify the defect types of PD(Partial Discharge). The UHF(Ultra High Frequency) signals of various defect models are measured in laboratory and AORGK(Adaptive Optimal Radially Gaussian Kernel) time-frequency analysis is then introduced to represent the UHF signals. The obtained time-frequency amplitude matrices are further compressed by 2DNMF and FkNN(Fuzzy k -Nearest Neighbor) classifier is then applied to recognize the four typical PD defects. Experimental results show that, TFA describes the time-frequency information of PD UHF signals effectively; the extracted features reserves most useful information of original time-frequency matrices; FkNN classifier has a higher recognition rate than those of kNN(k -Nearest Neighbor) classifier and BPNN(Back Propagation Neural Network), and is easier to expand than BPNN.

Key words: insulation; partial discharge; pattern recognition; time-frequency analysis; two-dimensional non-negative matrix factorization; fuzzy k -nearest neighbor classifier