

文章编号:1000-582X(2006)10-0001-04

一种基于线性鉴别分析的 GIS 局部放电模式识别*

张晓星,唐炬,孙才新,许中荣,周倩

(重庆大学 高电压与电工新技术教育部重点实验室,重庆 400030)

摘要:根据 GIS 设备绝缘缺陷放电形式和特点,设计了 4 种典型的 GIS 缺陷模型,构造了局部放电灰度谱图;针对 GIS 局部放电及其缺陷特点,提出一种基于局部放电图像的主分量分析-线性鉴别方法,即首先进行主分量分析,将数据从超高维空间降至低维空间,再提取统计不相关的最优鉴别矢量集,采用最小距离分类器进行模式识别,识别结果表明该方法对 GIS 各类模拟缺陷的正确识别率较高,效果良好。

关键词:GIS;局部放电;模式识别;主分量分析;线性鉴别分析

中图分类号:TM835

文献标识码:A

由于传统脉冲电流法易受现场电晕等低频电磁干扰,目前大量采用超高频法检测 GIS (Gas Insulated Switchgear) 局部放电 (Partial Discharge) 信号。由于谱图中含有局部放电丰富的信息,因此局放特征通常从二维和三维谱图($\varphi-q-n$) 中提取,提取特征的方法有:统计特征^[1]、分形特征^[2-4]、小波特征^[5-6]等。

在模式识别领域中, Fisher 线性判别方法 (FDA) 有着重大的影响,其基本思想是,选择使得 Fisher 准则函数-广义 Rayleigh 商达到最大值的向量(称最优鉴别向量)作为最优投影方向,从而使得高维输入空间中的模式样本在该轴投影后,类间散度达到最大的同时类内散度达到最小^[7]。Foley 和 Sammon 发展了 Fisher 线性判别方法(在 Fisher 鉴别准则函数取最大值的条件下,求得与 Fisher 最佳鉴别方向正交的第 2 个最佳鉴别方向,用于解决两类问题^[8])。Duchene 和 Leclercq 给出了对多类问题的 Foley-Sammon 最佳鉴别矢量集的计算公式^[9]。杨静宇等提出了统计不相关的最优矢量集的解析算法,并用于人脸识别,取得了良好的效果^[10-11]。

尽管 Fisher 及改进算法在模式识别领域得到一定的应用,但也存在一定的问题:由于局部放电灰度图像维数通常很大(高达数万维),导致构造最优鉴别矢量集的计算量非常大;而训练样本又远远小于图像维数,

导致类内散布矩阵常常为奇异阵,使得算法难以进行。为避免上述问题,局放数据在进行线性鉴别前,先进行主分量分析 (PCA), PCA 将数据从超高维空间降至低维空间,不仅可以解决维数危机,消除类内散度矩阵的奇异性,而且能最大限度的保持原有样本模式的结构分布。

笔者设计了 4 种 GIS 典型缺陷模型,构造出超高频局部放电灰度图像,采用 PCA-FDA 算法,将高维的局部放电灰度图像压缩至低维;在此基础上,提取了统计不相关的最优鉴别矢量集,最后在特征空间中采用最小距离分类器进行模式识别,取得了良好的识别效果。

1 人工缺陷物理模型与信号检测

1.1 人工缺陷物理模型的设计

根据 GIS 设备绝缘缺陷放电形式和特点,设计了 4 种 GIS 模拟装置内 PD 物理模型(见图 1):

- 1) 金属突出物缺陷:在内导体表面径向安装一根可调节长短的银针来模拟这种放电类型,根据试验装置内的气压和外施电压高低,选择银针突出的高度;
- 2) 自由金属微粒缺陷:微粒运动的程度取决于微粒的材料性质、形状大小和外施电压高低等因素。通过在模拟装置上大量试验研究,采用一定数量的矩形铝金属

* 收稿日期:2006-06-08

基金项目:国家自然科学基金资助项目(50577069);重庆市自然科学基金资助(CSTC2005BB3170)

作者简介:张晓星(1972-),男,湖北潜江人,重庆大学副教授,博士研究生,主要从事高压电气设备绝缘在线监测和故障诊断研究工作。

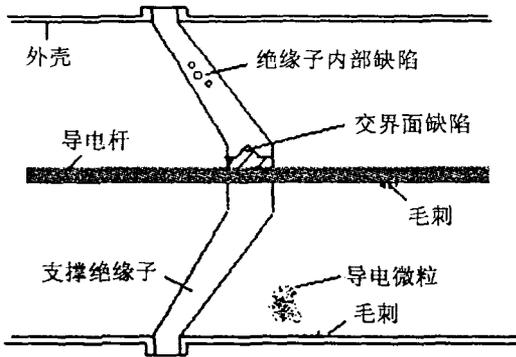


图1 GIS模拟装置内绝缘缺陷物理模型

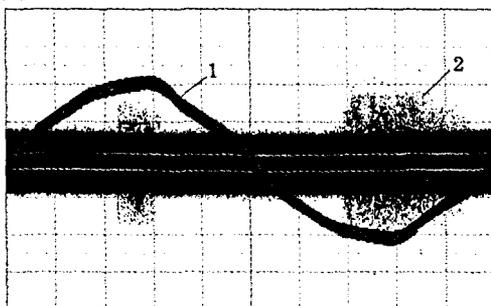
微粒放入装置中央;3)绝缘子表面固定金属微粒缺陷:作者用细长的锡箔纸粘于绝缘子上面模拟此缺陷;4)气隙缺陷:在GIS内部,当绝缘子和内导体间的交界处出现一定的间隙时,会产生损害固体绝缘子的局部放电.作者选用有机绝缘材料与同轴内导体之间形成一定气隙来模拟该缺陷类型.

1.2 GIS模拟装置的局放信号检测

将上述4种人工物理缺陷模型分别置于研制的GIS模拟装置中^[12],并充以0.5 MPa的SF₆与N₂的混合气体(体积比4:1),通过研制的内置超高频传感器和WavePro7100高速数字存储示波器等组成的检测系统采集PD信号的数据和波形,同时从电容分压器上取工频信号提供相位信息.

由于GIS中产生的PD等效频率可高达1 GHz以上,采样频率越高,生成的局放信号灰度图像越丰富,有利于提高模式识别率,但是数据存储、分析处理量随之上升,综合考虑我们将采样率定为250 MS/s,分别采样100个工频周期为一个样本.

图2为对针缺陷产生的局放信号,用示波器对100个工频周期的信号叠加生成的余辉图,我们对采集的局部放电序列,根据测定的噪声设置阈值,提取大于阈值的放电信号,构造 $\varphi-q-n$ (放电工频相位 φ 、放电幅值 q 和放电次数 n)灰度图像,分辨率为128×256,灰度级为0-255.



1. 工频波形;2. 放电脉冲

图2 针缺陷下100个工频周期叠加后波形与相位之间的关系

2 局部放电图像的PCA-FDA分析

2.1 局放图像的PCA分解算法

主分量分析是近年来研究较多的一种统计方法,由于PCA中主分量对应于数据均方差重建误差曲面的最小点,因此对数据有较强的描述能力,PCA充分利用数据中的二阶统计信息进行特征提取和降维,算法简单,运算量小.

主成分分解步骤如下:

1)将局放图像按正负半周分为两部分,分别进行主分量分析.

2)将局放正(负)半周图像像素排成 n 维列向量 a_i ,则 P 个样本组成 $n \times P$ 矩阵 A :

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{bmatrix}. \quad (1)$$

3)计算其协方差矩阵

$$C_a = \frac{1}{p} \sum_{i=1}^p [a_i - E(a_i)][a_i - E(a_i)]^T. \quad (2)$$

协方差矩阵 C_a 的 n 个特征值 $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ 对应的标准正交的本征向量 $\xi_1, \xi_2, \cdots, \xi_n$ 构成特征空间的一组基,如果前 $d (< n)$ 个基向量可以描述 a_i 的主要信息,则由这 d 个基向量张成的子空间就形成了描述的特征空间.

2.2 Foley-Sammon最佳鉴别矢量集

在 n 维空间中, $X = \{x_i\}, i = 1, 2, \cdots, N$ 为 n 维训练样本集, w_1, w_2, \cdots, w_m 为 m 个模式类, $x_i \in w_j$,设 w_j 类的平均矢量、先验概率为 $m_i, \sum_{i=1}^m P(w_i)$,则类间散布矩阵 S_b 、类内散布矩阵 S_w 和总散布矩阵 S_t 分别为:

$$m_i = \frac{1}{N} \sum_{x \in w_i} x, i = 1, 2, \cdots, m. \quad (3)$$

$m_0 = \sum_{i=1}^m P(w_i) m_i$ 为全体训练样本的平均矢量;

$$S_b = \sum_{i=1}^m P(w_i) (m_i - m_0)(m_i - m_0)^T, \quad (4)$$

$$S_w = \sum_{i=1}^m P(w_i) E\{(x - m_i)(x - m_i)^T / w_i\}, \quad (5)$$

$$S_t = S_b + S_w. \quad (6)$$

Fisher鉴别准则是,选择使广义Rayleigh商取最大值时的向量 φ 为Fisher最佳鉴别方向:

$$J(\varphi) = \frac{\varphi^T S_b \varphi}{\varphi^T S_w \varphi} \quad (7)$$

训练样本在方向 φ_1 上的投影集在一维子空间 $\text{Span}\{\varphi_1\}$ 中有最小的类内距离和最大的类间距离.求

解上述问题可以转换为求解特征方程 $S_b \varphi = \lambda S_w \varphi$ 的最大特征值对应的特征向量问题。

在求出 r 个最优鉴别矢量 $\varphi_1, \varphi_2, \dots, \varphi_r$ 后,第 $r+1$ 个最佳鉴别方向可取为满足下列共轭条件下使(7)式取最大值的向量 φ_{r+1} :

$$\varphi_{r+1}^T \varphi_i = 0 \quad (i = 1, 2, \dots, r) \quad (8)$$

模式识别中希望抽取的特征矢量集是统计不相关的,但 Foley-Sammon 最佳鉴别变换中,任意 2 个最佳鉴别方向是正交的,但投影所得到的 2 个鉴别特征却可能是统计相关的。为了获得具有统计不相关的最佳鉴别矢量集,文献[10]提出一种改进的方法。

2.3 统计不相关最优鉴别矢量集

如果 2 个随机变量 ξ 和 η 的协方差满足:

$$E[(\xi - E\xi)(\eta - E\eta)] = 0, \quad (9)$$

则称 ξ 和 η 是统计不相关的。

为了得到统计不相关的鉴别特征,在求出第一个最佳鉴别矢量 φ_1 后,第 $r+1$ 个最佳鉴别方向可取为满足下列共轭条件下使式(7)取最大值的向量 φ_{r+1} :

$$\varphi_{r+1}^T S_i \varphi_i = 0 \quad (i = 1, 2, \dots, r) \quad (10)$$

由于共轭正交条件(10),改进的最佳鉴别矢量集中任意两个鉴别矢量都是统计不相关的,则第 $r+1$ 个最佳鉴别方向 φ_{r+1} 是下列广义特征方程中最大的特征值对应的特征向量:

$$PSb\varphi_{r+1} = \lambda S_w \varphi_{r+1} \quad (11)$$

$$P = I - S_i D^T (D S_i S_w^{-1} S_i D^T)^{-1} D S_i S_w^{-1} \quad (12)$$

$$\text{其中, } I \text{ 为单位矩阵, } D = [\varphi_1, \varphi_2, \dots, \varphi_r]^T \quad (13)$$

对于 n 维向量 X ,在求出 k (通常为 $m-1$ 个, m 为训练样本类数) 个最佳鉴别方向 $\varphi_1, \varphi_2, \dots, \varphi_k$ 后,将 X 向这 k 个最佳鉴别方向上投影可得 k 维特征向量 Y 。

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{bmatrix} = \begin{bmatrix} \varphi_1^T \\ \varphi_2^T \\ \vdots \\ \varphi_k^T \end{bmatrix} X \quad (14)$$

通过以上的统计不相关最佳鉴别变换,就可以将 n 维向量降维得到 k 维的鉴别特征向量。

3 试验结果

试验样本由 4 种放电缺陷模型在不同的试验电压下测得,每类 200 个有效放电样本,取其中 50 个样本作为训练样本,其余 150 个为测试样本。

根据所提出的算法,结合局部放电灰度图像的特性,其步骤如下:

1) 将局部放电灰度图像分按正负半周分为两部分,分别进行主分量分析。由于主分量分析中各分量对

图像特征的描述能力依次减弱,在试验中,笔者采用 PCA 抽取 60~100 个主分量,测试不同数量的主分量对识别率的影响,表 1 为对绝缘子金属污染缺陷局放图像选取 60~100 个投影轴数时,再使用第 2、3 步进行模式识别的结果比较。由表 1 可知当选取 80 个主分量时,识别率已经足够高。经过主分量变换后,正(负)半周的局放图像降维至 80 维。

表 1 主分量投影轴数对识别率的影响

投影轴数	识别率/%
60	89.3
65	90.6
70	91.3
75	92.6
80	94.0
85	94.0
90	93.3
95	93.3
100	93.3

2) 对第 1 步得到的正(负)半周的 80 维向量,利用统计不相关最优鉴别矢量准则提取 3 个(类别数减 1)最优鉴别向量,正负半周合计 6 个最优鉴别向量。

3) 在投影空间中,利用最小距离分类器进行模式识别,其识别结果如表 2 所示。

表 2 GIS 局部放电模式识别结果

识别结果	缺陷类型			
	金属突出物	自由金属微粒	绝缘子金属污染	气隙
试	23.0	29.0	15.8	15.8
验	24.8	30.0	18.7	16.7
电	26.6	31.5	20.8	17.6
压	27.3	32.7	26.4	18.6
(kV)	28.8	33.1		
正确识别率%	88.0	90.6	94.0	84.0

从表 2 可以看出,该算法方法对各种缺陷的识别率都在 80% 以上,说明该方法对 GIS 各类模拟缺陷的正确识别率较高,效果良好。

4 结论

设计了 4 种 GIS 典型缺陷模型,构造出局部放电灰度图像。针对局部放电灰度图像的特点,提出一种基于局部放电图像的主分量分析-线性鉴别(PCA-FDA)方法,即首先进行主分量分析,再提取统计不相关的最优鉴别矢量集,最后采用最小距离分类器进行模式识别,识别结果表明该方法对 GIS 各类模拟缺陷的正确识别率较高,效果良好。

参考文献:

- [1] 李剑,孙才新,廖瑞金,等.用于局部放电图像识别的统计特征研究[J].中国电机工程学报,2002,22(9):104-107.
- [2] 高凯,谈克雄,李福祺,等.基于散点集分形特征的局部放电模式识别研究[J].中国电机工程学报,2002,22(5):22-26.
- [3] 李剑,孙才新,杜林,等.局部放电灰度图象分维数的研究[J].中国电机工程学报,2002,22(8):123-127.
- [4] 孙才新,许高峰,唐炬,等.以盒维数和信息维数为识别特征量的GIS局部放电模式识别方法[J].中国电机工程学报,2005,25(3):100-104.
- [5] 谈文刚,陈祥训,郑健超.采用小波分析与神经网络技术的局部放电统计识别方法[J].中国电机工程学报,2002,22(9):1-5.
- [6] 孙才新,李新,李俭,等.小波与分形理论的互补性及其在局部放电模式识别中的应用研究[J].中国电机工程学报,2001,21(12):73-76.
- [7] 边肇祺,张学工.模式识别[M].北京:清华大学出版社,2000.
- [8] FOLEY D H, SAMMON J W. An optimal Set of Discriminant Vectors [J]. IEEE Trans on Computers, 1975, 24(3): 281-289.
- [9] DUCHENE J, LECLERCQ S. An Optimal Transformation for Discriminant and Principal Component Analysis [J]. IEEE Trans On PAMI, 1988, 10(6): 978-983.
- [10] 金忠,杨静宇,陆建峰.一种具有统计不相关性的最优鉴别矢量集[J].计算机学报,1999,22(10):1105-1108.
- [11] 郭跃飞,杨静宇.求解广义最佳鉴别矢量的一种迭代算法及人脸识别[J].计算机学报,2000,23(11):1189-1195.
- [12] 孙才新,许高峰,唐炬等.检测GIS局部放电的内置传感器的模型及性能研究[J].中国电机工程学报,2004,24(8):89-94.

Pd Pattern Recognition Based on Linear Discriminant Analysis in GIS

ZHANG Xiao-xing, TANG Ju, SU Cai-xin, XU Zhong-rong, ZHOU Qian

(Key Laboratory of High Voltage Engineering and Electrical New Technology, Ministry of Education, Chongqing University, Chongqing 400030, China)

Abstract: According to the character of PD in GIS, the authors design four kinds of GIS defection models. The GIS gray intensity images are constructed based on mass specimens gathered by the ultra-high frequency and high speeds systems. Aiming at the PD characteristics and its defections, A PCA-FDA method is put forward based on PD images. The principal component analysis is employed to condense the dimension of PD images, then the optimal sets of statistically uncorrelated discriminant vectors are extracted, and the minimum distance classifier is constructed as classifier. The identified results show that this method can effectively elevated the discrimination of the four kinds of defects in GIS PD.

Key words: GIS ;PD; pattern recognition; Linear Discriminant Analysis; principal component analysis

(编辑 姚 飞)