Application of Grey Clustering Approach and Genetic Algorithm to Partial Discharge Pattern Recognition

WEN-YEAU CHANG
Department of Electrical Engineering
St. John's University
No. 499, Sec. 4, Tam King Road, Tamsui, Taipei 251, Taiwan
TAIWAN
changwy@mail.sju.edu.tw

Abstract: Partial discharge (PD) measurement and recognition is a significant tool for potential failure diagnosis of the high-voltage equipment. This paper proposes the application of grey clustering approach (GCA) to recognize partial discharge patterns of the high-voltage equipment. The PD patterns are measured by using a commercial PD detector. A set of features, used as operators, for each PD pattern is extracted through statistical schemes. The significant features of PD patterns are extracted by using the genetic algorithm (GA). The proposed grey clustering approach has the advantages of high robustness and effectiveness to ambiguous patterns and is useful in recognizing the PD patterns of the high-voltage equipment. To verify the effectiveness of the proposed method, the grey clustering approach was verified on two types of high-voltage equipments. The test results show that the proposed approach may achieve quite satisfactory recognition of PD patterns.

Key-Words: Partial discharge, Pattern recognition, Grey clustering approach, Genetic algorithm

1 Introduction

Partial discharge measurement and pattern recognition are important tools for improving the reliability of high-voltage insulation systems. The pattern recognition of PD aims at identifying potential insulation defects from the measured data. The potential defects can then be used for estimating the risk of insulation failure of the high-voltage equipment [1].

In the presence of a sufficiently strong electric field, a sudden local displacement of electrons and ions will lead to a PD if there exists a defect in an insulator [2]. A PD event that occurs in the epoxy resin insulator of high-voltage equipment would have harmful effects on insulation that may finally cause service failure. A defect in high-voltage equipment, resulting in PD, will have a corresponding particular pattern. Therefore, pattern recognition of PD is significant for insulation condition evaluation of high-voltage equipment.

Thanks to physical understanding of PD made substantial progress in the last decade, it can now be exploited to support interpretation of insulation defects. Recently, several methods have been employed for the pattern recognition of PD, including neural networks, expert systems, self organizing map, wavelet analysis, and fuzzy classification methods.

The application of neural networks to pattern recognition and system identification has become a major trend in the fault diagnosis [3-4]. Neural networks has been applied for PD classification of epoxy resin power transformer (CRCT) [5], PD pattern recognition of current transformers [6], and PD monitoring technique of gas insulated substation [7]. Although the speed of neural networks allows real-time operation with comparable accuracy, the training process of multilayer neural networks is often very slow, and the training data must be sufficient and compatible.

The recognition of PD pattern and the estimation of insulation performance are relatively complicated, a task which is often completed by experienced experts. Several expert systems for the diagnostics of insulation systems have been developed [8]. The expert system method acquires the knowledge of human expertise to build knowledge base. However, it needs to build and maintain the base with efforts.

The self organizing map is a typical unsupervised neural network, which maps the multidimensional space onto a two dimensional space by preserving the original order [9]. It simulates the self-organizing feature map's function of the human cerebrum. The self organizing map is a two-layer neural network that consists of an input layer in a line and an output layer constructed of neurons in a two-dimensional grid.

Different from other clustering mapping methods for unsupervised data, mapping relationship of self organizing map can be highly nonlinear, directly showing the similar input vectors in the source space by points close in the two-dimensional target space [10]. Along with the similarity of the input data, self organizing map potentially leads to a classification result. It has been applied for PD pattern recognition of CRCT [11].

The wavelet analysis method is a useful tool in fault detection [12] and de-noising [13]. Wavelet analysis method has also been applied to identify the PD characteristics by decomposition of acoustic emission signals [14] and PD signal de-noising [15-17].

Another method is the fuzzy clustering algorithm [18]. The fuzzy c-means clustering algorithm is one of the most popular fuzzy clustering algorithms [19]. It has been applied for PD pattern recognition of CRCT [20]

This paper proposes the application of GCA to recognize partial discharge patterns of the high-voltage equipment. The PD patterns are measured by using a commercial PD detector. A set of features, used as operators, for each PD pattern is extracted through statistical schemes. The significant features of PD patterns are extracted by using the GA.

This paper is organized as follows. Creation of the PD pattern dataset and the extraction of phase-related distributions are described in Section 2. The development of the algorithm of statistical feature extraction is described in Section 3. The GA features extraction algorithm is described in Section 4. The principles of GCA and the operation flowchart of the proposed pattern recognition scheme are given subsequently. The experimental results and the analysis using 500 sets of field-test PD patterns from 2 kinds of high-voltage equipments are presented in Section 6. From the test results, the effectiveness of the proposed scheme to improve the recognition accuracy has been demonstrated. The paper is concluded in Section 7.

2 PD Pattern Dataset Creation

In order to investigate the PD features and to verify the classification capabilities of the proposed GCA based pattern recognition technique for different PD types commonly occurring in high-voltage equipments, a PD dataset is needed. The PD dataset was collected from laboratory tests on a series of model cast-resin current transformers and insulators. The material and process used to manufacture the CRCTs and insulators were exactly the same as that of making a field equipment. The appearance of model CRCT is shown in Fig. 1. The specifications of model CRCT are shown in Table 1.

Five types of experimental models with artificial defects embedded were made to produce five common PD events in the CRCT. The five PD activities include (1) normal PD activity in standard CRCT (NM), (2) internal cavity discharge caused by

Fig. 1 The appearance of model CRCT

Table 1 The specifications of model CRCT

Service Voltage	Primary Current	Secondary Current	Burden
12000 V	20 A	5 A	40VA

an air cavity inside the epoxy resin insulator on the high-voltage side (VH), as shown in Fig. 2, (3) internal cavity discharge caused by two cavities inside the epoxy resin insulator on the low-voltage side (VL), as shown in Fig. 3, (4) internal fissure discharge caused by an air fissure inside the epoxy resin insulator on the high-voltage side (FH), as shown in Fig. 4, (5) internal discharge caused by a metal-line impurity inside the epoxy resin insulator on the high-voltage side (MH), as shown in Fig. 5.

The appearance of model insulator is shown in Fig. 6. Five types of experimental models with artificial defect are purposely manufactured to produce five common PD activities on insulators. Five PD activities include (1) normal PD activity in standard epoxy resin insulator (NM), (2) internal cavity discharge caused by the air cavity inside epoxy resin insulator (VD), (3) internal discharge caused by a metal-line impurity inside epoxy resin insulator (MD), (4) external discharge in air between two plane electrodes (ED), (5) corona discharge in air between a needle electrode and a plane electrode (CD).

Fig. 7 shows the insulator model for NM, VD, and MD, the distance between two plane electrodes is fixed at 10mm. The epoxy resin insulator is normal for NM, with air cavities for VD, and with metal-line impurities for MD. Fig. 8 shows the model for ED, the distance between two plane electrodes is also fixed at 10mm. Fig. 9 shows the model for CD, the distance between the needle tip and the plane electrode is also fixed at 10mm.

The PD events were detected by a PD detecting system set up in our laboratory. The structure of the PD detecting system is shown in Fig. 10. It includes a step-up transformer, capacitor coupling circuit, PD detector, and the equipment under test. Through the testing processes, all the data measured were digitally converted in order to save them in the computer memory.

Then, the phase-related distributions of PD derived from the original PD data are obtained in relation to the waveform of the field test high voltage. The high voltage in the field tests is assumed to be held constant and the voltage phase angle is divided into a suitable number of windows (blocks). The PD detector, shown in Fig. 10, is used for acquisition of all the individual quasi-integrated pulses and quantifying each of these PD pulses by their discharge magnitude (q), the corresponding phase angle (ϕ), at which PD pulses occur and the number of discharge (n) over the chosen block. The analysis software (DDX DA3) plots these data as functions of the phase positions.

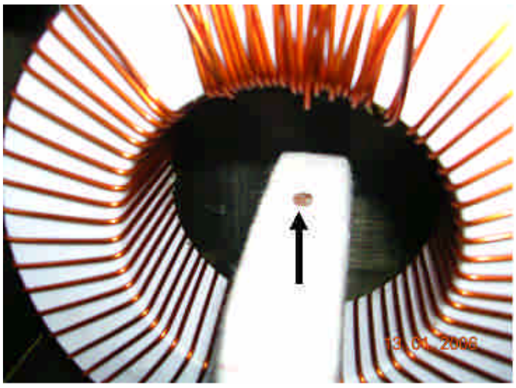

Fig. 2 VH on the high-voltage side of CRCT

Fig. 3 VL on the low-voltage side of CRCT

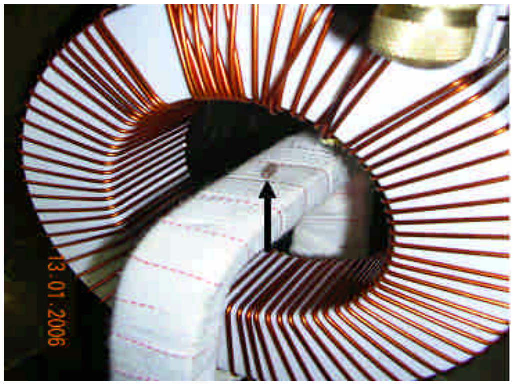

Fig. 4 FH on the high-voltage side of CRCT

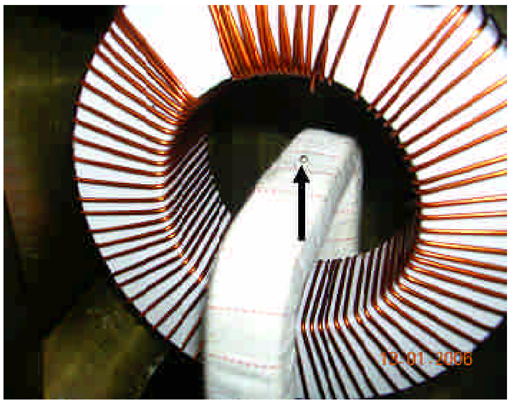

Fig. 5 MH on the high-voltage side of CRCT

The three phase-related distributions refer to the peak pulse magnitude distribution $H_{qmax}(\phi)$, the average pulse magnitude distribution $H_{qn}(\phi)$, and the number of pulse distribution $H_n(\phi)$. The typical phase-related distributions of PD patterns for the four kinds of defects (VH, VL, FH, and MH) of CRCTs are shown in Figs. 11 to 14, respectively. As shown in Figs. 11 to 14, the PD patterns of deferent defects display discriminative features.

3 Statistical Feature Extraction

Feature extraction is a technique essential in PD pattern recognition to reduce the dimension of the original data. The features are intended to denote the characteristics of different PD statuses. Several statistical methods of feature extraction are described in this section; five statistical operators are extracted from phase-related distributions. Definitions of the operators are described below. The profile of all these discrete distribution functions can be put in a general framework, i.e., $y_i = f(x_i)$ [21].

The statistical operators of mean (μ) and variance (σ^2) can be computed as follows:

$$\mu = \frac{\sum x_i f(x_i)}{\sum f(x_i)} \tag{1}$$

$$\sigma^2 = \frac{\sum (x_i - \mu)^2 f(x_i)}{\sum f(x_i)} \tag{2}$$

Skewness (S_k) is extracted from each phase-related distribution of PD to denote the asymmetry of the distribution. It can be represented as:

$$S_k = \frac{\sum (x_i - \mu)^3 p_i}{\sigma^3} \tag{3}$$

Kurtosis (K_u) is extracted to describe the sharpness of the distribution as:

Fig. 6 The appearance of model insulator

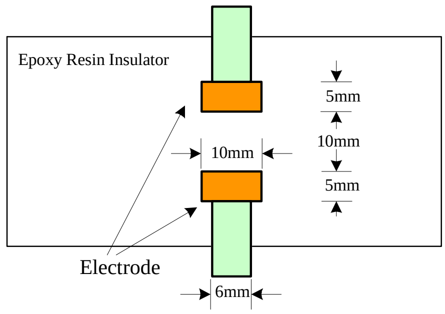

Fig. 7 Structure of model for NM, VD, MD defects

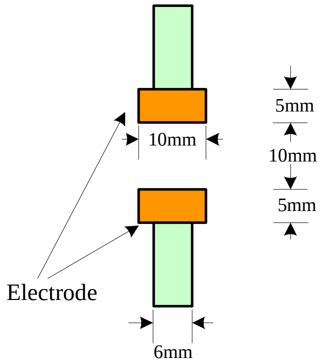

Fig. 8 Structure of model for ED defect

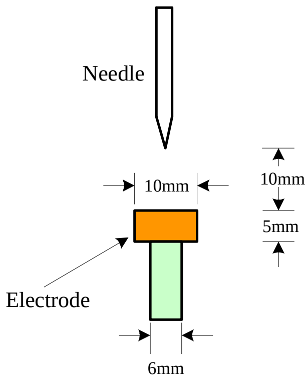

Fig. 9 Structure of model for CD defect

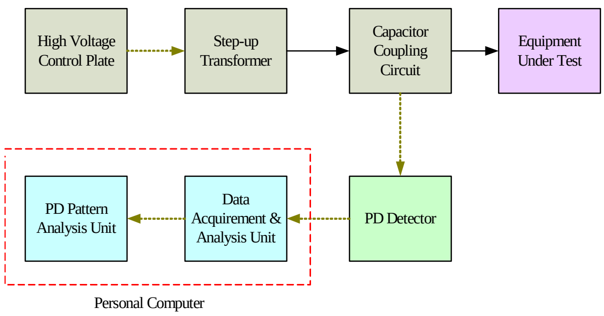

Fig. 10 System configuration of the PD detecting system

$$K_u = \frac{\sum (x_i - \mu)^4 p_i}{\sigma^4} - 3 \quad (4)$$

In (1) and (2), x_i is the statistical value in the phase window i, p_i is the related probability of appearance.

Skewness is a measure of asymmetry degree with respect to normal distribution. If the distribution is totally symmetric, then $S_k = 0$; if the distribution is asymmetric to the left of mean, $S_k > 0$; and if it is asymmetric to the right of mean, $S_k < 0$. Kurtosis is an indicator of sharpness of distribution. If the distribution has the same sharpness as a normal distribution, $K_u = 0$; and if it is sharper than normal, $K_u > 0$; and if it is flatter than normal, $K_u < 0$ [21].

Peaks (P_e) count the number of peaks in the positive or negative half of a cycle of the distribution.

Asymmetry (D_a) represents the asymmetrical characteristic of partial pulses in both positive and negative cycles. It is given by:

$$D_a = \frac{N^+ \sum q_i^-}{N^- \sum q_i^+} \quad (5)$$

where N^- is the number of PD pulses in the negative cycle, N^+ is the number of PD pulses in the positive cycle. q_i^- is the amplitude of the PD pulse at a phase window i in the negative cycle, and q_i^+ is the amplitude of the PD pulse at a phase window i in the positive cycle.

The cross correlation factor (C_c) can be expressed as:

$$C_c = \frac{\sum x_i \cdot y_i - \sum x_i \cdot \sum y_i / n}{\sqrt{(\sum x_i^2 - (\sum x_i)^2 / n) \cdot (\sum y_i^2 - (\sum y_i)^2 / n)}} \quad (6)$$

where x_i is the statistical value in the phase window i of the positive half cycle, y_i is the statistical value in the corresponding window of the negative half cycle, and n is the number of phase window per half cycle.

Cross correlation factor indicates the difference in the distribution sharps of both positive and negative half cycles. $C_c = 1$ means the sharps are totally symmetric, $C_c = 0$ means sharps are totally asymmetric.

As S_k, K_u and P_e are applied to both positive and negative cycles of $H_{qmax}(\phi)$, $H_{qn}(\phi)$, and $H_n(\phi)$, a total of 18 features can be extracted from a PD pattern. D_a and C_c are applied to indicate the difference or asymmetry in positive and negative cycles of $H_{qmax}(\phi)$, $H_{qn}(\phi)$, and $H_n(\phi)$, and a total of 6 features can be extracted from a PD pattern. Therefore, after the feature extraction procedure, a feature vector of 24 statistical features is built for each PD pattern.

The use of statistical featuring operators for the patterns instead of the distribution profiles can significantly reduce the dimension of the database. To a certain degree, they can characterize the PD patterns with reasonable discrimination [22].

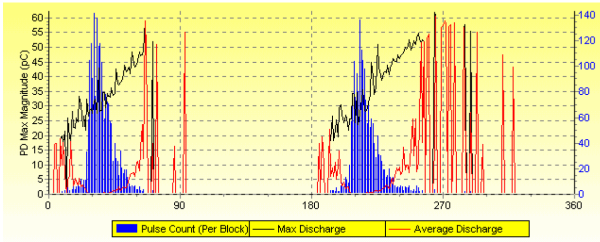

Fig. 11 Typical phase-related distributions of PD for the VH defect

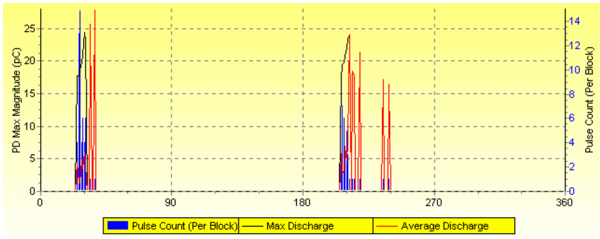

Fig. 12 Typical phase-related distributions of PD for the VL defect

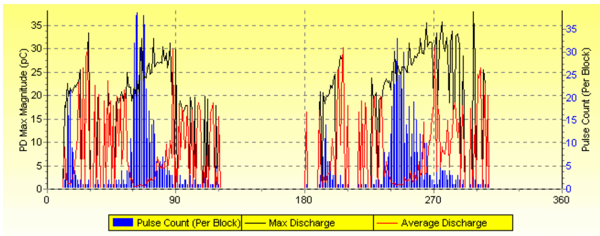

Fig. 13 Typical phase-related distributions of PD for the FH defect

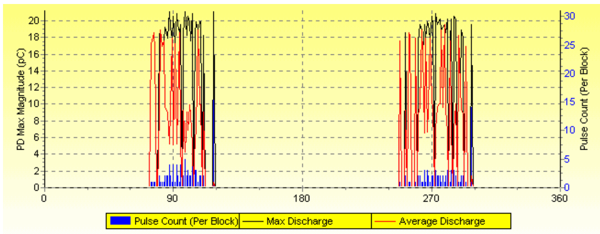

Fig. 14 Typical phase-related distributions of PD for the MH defect

4 GA Feature Extraction Method

Feature extraction is necessary in the PD pattern recognition to reduce dimension of original data and make effective discrimination of the statistical feature patterns for different PD status. In this paper, the significant features are extracted from statistical features by using GA method. Genetic algorithm is a search method utilizing the mechanism of natural selection and genetics. The application of genetic algorithm to optimization has become a useful tool in many fields [23-24].

The statistical feature extraction methods were used to extract 24 statistical features for patterns. But some of the statistical features are futile for pattern recognition. So, the combination of feature subset will influence the accuracy of pattern recognition. GA is a nonlinear optimization technique and has been widely used as a tool for feature selection in pattern recognition. In this paper, we employ GA for optimal feature vector combination selection. The proposed GA-based feature subset selection scheme has been successfully implemented using PC-based software (MATLAB). The proposed selection scheme is described briefly in the following steps:

Step1 The initial population has been randomly chosen from all possible combination of features.

Step2 The fitness function of the proposed scheme is defined as follows:

$$F(c)=\frac{\left[\sum_{i=1}^{24}\sum_{j=1}^{5}b_i(\mu_{i,j}-\mu_{i,else})^2 / \sum_{i=1}^{24}\sum_{j=1}^{5}b_i\sigma_{i,j}^2\right]-P(x)}{\sum_{i=1}^{24}b_i} \tag{7}$$

where $P(x)=\begin{cases}0 & if \quad 7\le \sum_{i=1}^{24}b_i \le 13\\ \infty & otherwise\end{cases}$

c is a chromosome, b_i is the value of bit i in c, $\mu_{i,j}$ is the mean of ith features belong to class j, $\mu_{i,else}$ is the mean of ith feature not belong to pattern class j, and $\sigma_{i,j}$ is the variance of ith feature belong to pattern class j.

Step3 According to the fitness values of individuals, the pairs from current population are selected for crossover to generate offspring.

Step4 The crossover operation generates new population from the selected pairs at a randomly chosen point. The probability of crossover in this paper is 0.6.

Step5 For all bits of the offspring string, the mutation operation is then performed. The probability of mutation in this paper is 0.002.

Step6 The optimization process repeats Step 2 to Step 5 until the stopping rule is satisfied. Then the optimal feature subset is selected.

5 GCA-Based PD Pattern Recognition Method

Grey system theory is useful methodology for systems with incomplete information. Grey relational analysis can be used to analysis the relationships between one major (reference) sequence and the other comparative ones in a given set [25-26].

In this section, the algorithms of grey relational analysis and GCA-based PD pattern recognition scheme are described. The PD recognition through GCA in multidimensional feature space is also validated on the basis of the features extracted by GA method as mentioned above.

5.1 Principals of Grey relational analysis

In this subsection, the principals of the grey relational analysis are described as follows [27]: Assume that the reference sequence is $\overline{x_i} = (x_i(1), x_i(2),..., x_i(n))$, and denote m comparative sequences as $\overline{x_j} = (x_j(l), x_j(2),..., x_j(n))$, $j = 1,2,.....,m$, where $x_j(k)$ stands for the k-th entry in $\overline{x_j}$, $k = 1,2,...,n$.

The grey relational coefficient of $\overline{x_j}$ with respect to $\overline{x_i}$, at the k-th entry can be obtained from following equation

$$r(x_i(k),x_j(k))=\frac{\Delta_{\min}+\xi\ \Delta_{\max}}{\Delta_{ij}+\xi\ \Delta_{\max}} \tag{8}$$

where ξ is the distinguished coefficient, $0\le \xi \le 1$.

$$\Delta_{ij}=\left|x_i(k)-x_j(k)\right| \tag{9}$$

$$\Delta_{\max}=\max_j \max_k\left|x_i(k)-x_j(k)\right| \tag{10}$$

$$\Delta_{\min}=\min_j \min_k\left|x_i(k)-x_j(k)\right| \tag{11}$$

where $j = 1,2,\ldots,n$, $k = 1,2,\ldots,m$.

From the grey relational coefficient, the grey relational grade of each comparative sequence and the reference sequence can be obtained. The grey relational grade is denoted as

$$r(\overline{x_i},\overline{x_j})=\frac{1}{n}\sum_{k=1}^{n}r(x_i(k),x_j(k)) \tag{12}$$

where $r(\overline{x_i}, \overline{x_j})$ represents the degree of grey relation between the reference sequence $\overline{x_i}$ and the comparative sequence $\overline{x_j}$.

The higher degree of relation means the comparative sequence is more similar to the reference sequence than the others. Using the grey relational measure to train the grey cluster among of the given data points, and clustering the data according to the grey relational grade. The detail procedure of grey cluster training is described in the next subsection.

5.2 Training Procedure for Grey Cluster

Let $X = \{\overline{x_1}, \overline{x_2}, \ldots\ldots, \overline{x_n}\}$ be a given data set of m objects, and $\overline{x_i} = (x_i(1), x_i(2), \ldots, x_i(n))$ is regard as a vector that is described by n real-valued measurements of their features. Consider these vectors as sequences and choose an object as a reference sequence by turns and all objects as the comparative sequences [27]. The detail steps of grey cluster training can be proposed as follows [28]:

Step1 Define m moveable vectors $\overline{v_i} = (v_i(1), v_i(2), \ldots, v_i(n))$, $i = 1,2,\ldots,m$ and let $\overline{v_i} = \overline{x_i}$. Choose an appropriate threshold ω, $0 < \omega < 1$.

Step2 The relational grades between the reference vector $\overline{v_i}$ and the comparative vectors $\overline{v_j} \in \{v_1, v_2, \ldots, v_m\}$ can be obtained from following equations

$$r(\overline{v_i}, \overline{v_j}) = \frac{1}{n}\sum_{k=1}^{n} r(v_i(k), v_j(k)) \qquad (13)$$

$$r(v_i(k), v_j(k)) = \frac{\Delta_{\min} + \xi\ \Delta_{\max}}{\Delta_{ij} + \xi\ \Delta_{\max}} \qquad (14)$$

where the distinguished coefficient $\xi = 0.2$.

Step3 Calculate new reference vector $\overline{v_i}' = (v_i'(1), v_i'(2), \ldots, v_i'(n))$, $i = 1,2,\ldots,m$ by

$$v_i'(k) = \frac{\sum_{j=1}^{m} n_{ij} v_i(k)}{\sum_{j=1}^{m} n_{ij}}, \quad k = 1,2,\ldots,n \qquad (15)$$

where

$$n_{ij} = \begin{cases} 0, & if\ r(\overline{v_i}, \overline{v_j}) < \omega \\ 1, & otherwise \end{cases} \qquad (16)$$

$i = 1,2,\ldots,m$ and $j = 1,2,\ldots,m$

Step4 If all the new reference vectors does not change, $\overline{v_i}' = \overline{v_i}$ for $i = 1,2,\ldots,m$, then go to Step 5. Otherwise, let $\overline{v_i} = \overline{v_i}'$ and go to Step 2.

Step5 The convergent vector is viewed as the cluster center. Save the centres of trained grey clusters v_i, $i = 1,2,\ldots,m$,, as training procedure is finished.

5.3 GCA-based PD Pattern Recognizing Procedure

The proposed GCA-based PD pattern recognition scheme has been successfully implemented using PC-based software (MATLAB) for the PD recognition. The overall flowchart is shown in Fig. 15. The proposed recognition scheme is described briefly in the following steps:

Step1 Creating data base of the phase-related distributions of PD patterns.

Step2 Extracting the statistical features from phase-related distributions.

Step3 Extracting the significant features from statistical features by using GA method.

Step4 Prepare the training set for grey cluster.

Step5 Using the training set to train the grey cluster for PD pattern recognition.

Step6 Save the centres of trained grey clusters, as training procedure is finished.

Step7 Use (12) to calculate the grey relational grade to identify the defect types of PD patterns.

6 Experimental Results

To verify the proposed approach, a practical experiment is conducted to demonstrate the effectiveness of the PD pattern recognition scheme. The experimental tests were carried out on two kinds of high-voltage equipments: CRCT and insulator. The test results show that the proposed method is able to accurately recognize the testing defects.

6.1 Tests for CRCT

Five types of experimental models with artificial defects are purposely embedded to produce five common PD events in CRCT.

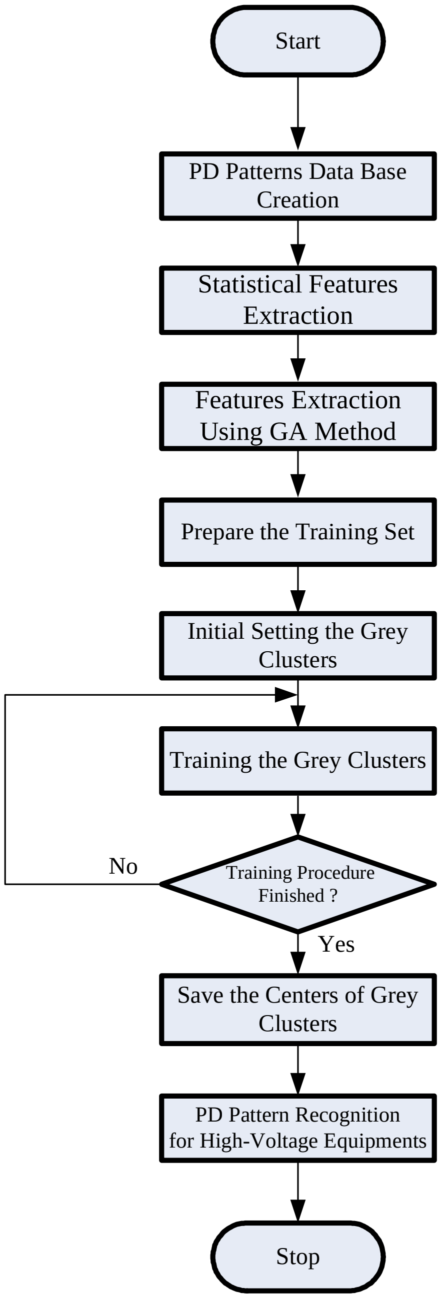

Fig. 15 Flowchart of the GCA-based recognition scheme

The proposed method has been implemented according to the field-test PD patterns collected from our laboratory. The input data to a PD recognition system are the peak pulse magnitude distribution $H_{qmax}(\phi)$, the average pulse magnitude distribution $H_{qn}(\phi)$, and the number of pulse distribution $H_n(\phi)$.

Associated with their real defect types, there are a total of 250 sample data for different PD events. Each PD event contains 50 patterns of sample data, of which 30 patterns are training data and 20 patterns are testing data.

The statistical feature extraction methods are used to extract 24 statistical features for each pattern. The significant features are extracted from statistical features by using GA method. The number of features for feature vector extracted by GA method is 11 in this experiment. After feature extraction process, all the features in the feature vectors were normalized to set up the training sets.

After setting up the training sets, the training procedure of grey clustering is started. The training data consist of 150 feature vectors, which are randomly chosen from the 250 feature vectors of sample data. The rest of 100 feature vectors were used as the testing data.

To verify the training results of grey clusters, the training data were applied to the trained grey clusters again. Table 2 shows the test result of the training data. The data in Table 2 shows that the proposed method has 100% accuracy for the 150 training feature vectors. Table 3 demonstrates the promising performance when 100 testing patterns were tested. It is shown in Table 3 that among the 100 testing patterns, there are only 3 errors of recognition, one for NM, one for VL and the other for MH defects.

6.2 Tests for Insulator

Five types of experimental models with artificial defect are purposely manufactured to produce five common PD activities on insulator.

The proposed method has been implemented according to the field-test PD patterns collected from our laboratory. Associated with their real defect types, there are a total of 250 sample data for different PD events. Each PD event contains 50 patterns of sample data, of which 30 patterns are training data and 20 patterns are testing data.

The statistical feature extraction methods are used to extract 24 statistical features for each pattern. The significant features are extracted from statistical features by using GA method. The number of features for feature vector extracted by GA method is 10 in this experiment. After feature extraction process, all the features in the feature vectors were normalized to set up the training sets.

After setting up the training sets, the training procedure of grey clustering is started. The training data consist of 150 feature vectors, which are randomly chosen from the 250 feature vectors of sample data. The rest of 100 feature vectors were used as the testing data.

To verify the training results of grey clusters, the training data were applied to the trained grey clusters again. Table 4 shows the test result of the training data. The data in Table 4 shows that the proposed method has 100% accuracy for the 150 training feature vectors. Table 5 demonstrates the promising performance when 100 testing patterns were tested. It is shown in Table 5 that among the 100 testing patterns, there are only 4 errors of recognition, one for VD, one for MD, one for ED, and the other for CD defects.

7 Conclusions

This paper has proposed a GCA based pattern recognition technique for PD of high-voltage equipments. The effectiveness of the proposed technique has been verified using experimental results. It has been shown that through the GA feature extraction procedure, the extracted feature vectors can significantly reduce the size of the PD pattern database. Also, the GCA based PD pattern recognition scheme is very effective for clustering the defects of high-voltage equipments.

The content of PD dataset influences the accuracy of pattern recognition. To ameliorate further the recognition accuracy of the proposed approach, the more plenteous PD dataset creation methods will be studied in the future researches.

8 Acknowledgements

The author would like to express his acknowledgements to the National Science Council of ROC for the financial support under Grant No. NSC 98-2221-E-129 -014.

Table 2 Recognition performance of training data in CRCT tests

Pattern	Defect Types	Accuracy Rate
Training Data (150 patterns)	NM	100%
	VH	100%
	VL	100%
	FH	100%
	MH	100%

Table 3 Recognition performance of testing data in CRCT tests

Pattern	Defect Types	Accuracy Rate
Testing Data (100 patterns)	NM	95%
	VH	100%
	VL	95%
	FH	100%
	MH	95%

Table 4 Recognition performance of training data in insulator tests

Pattern	Defect Types	Accuracy Rate
Training Data (150 patterns)	NM	100%
	VD	100%
	MD	100%
	ED	100%
	CD	100%

Table 5 Recognition performance of testing data in insulator tests

Pattern	Defect Types	Accuracy Rate
Testing Data (100 patterns)	NM	100%
	VD	95%
	MD	95%
	ED	95%
	CD	95%

References:

[1] L. Niemeyer, A Generalized Approach to Partial Discharge Modeling, *IEEE Transactions on Dielectrics and Electrical Insulation*, Vol. 2, No. 4, August 1995, pp. 510-528**.**

[2] C. Cachin and H.J. Wiesmann, PD Recognition with Knowledge-Based Preprocessing and Neural Networks, *IEEE Transactions on Dielectrics and Electrical Insulation*, Vol. 2, No. 4, 1995, pp. 578-589**.**

[3] M.G.D. Bono, G. Pieri, and O. Salvetti, A Tool for System Monitoring Based on Artificial Neural Networks, *WSEAS Transactions on Systems*, Volume 3, Issue 2, April 2004, pp. 746-751.

[4] G. Mousalli-Kayat, J. Calderón-Vielma, F. Rivas-Echeverría, and A. Ríos-Bolívar, Faults Detection and Isolation computational tool Using Neural Networks and State Observers, *WSEAS Transactions on Systems*, Volume 3, Issue 2, April 2004, pp. 773-777.

[5] H.C. Chen, P.H. Chen, and M.H. Wang, Partial Discharge Classification Using Neural Networks and Statistical Parameters,

Proceedings of the 6th WSEAS International Conference on Instrumentation, Measurement, Circuits & Systems, Hangzhou, China, April 15-17, 2007, pp. 84-88.

[6] M.H. Wang, Partial Discharge Pattern Recognition of Current Transformers Using an ENN, *IEEE Transactions on Power Delivery*, Vol. 20, No. 3, 2005, pp. 1984-1990.

[7] I. Oki, T. Haida, S. Wakabayashi, R. Tsuge, T. Sakakibarb, and H. Muraseg, Development of Partial Discharge Monitoring Technique Using a Neural Network in a Gas Insulated Substation, *IEEE Transactions on Power Systems*, Vol. 12, No. 2, May 1997, pp. 1014-1021.

[8] K. Zalis, Applications of Expert Systems in Evaluation of Data from Partial Discharge Diagnostic Measurement, *Proceedings of the 7th International Conference on Properties and Applications of Dielectric Materials*, 2003, pp. 331-334.

[9] A.F. Kuri-Morales, Automatic Clustering with Self-Organizing Maps and Genetic Algorithms II: an Improved Approach, *WSEAS Transactions on Systems*, Volume 3, Issue 2, April 2004, pp. 551-556.

[10] Y. Han and Y.H. Song, Using Improved Self-organizing Map for Partial Discharge Diagnosis of Large Turbogenerators, *IEEE Transactions on Energy Conversion*, Vol. 18, No. 3, 2003, pp. 392-399.

[11] W.Y. Chang and H.T. Yang, Application of Self Organizing Map Approach to Partial Discharge Pattern Recognition of Cast-Resin Current Transformers, *WSEAS Transactions on Computer Research*, Vol. 3, Issue 3, 2008, pp. 142-151.

[12] S. Postalcioglu, K. Erkan, and E.D. Bolat, Discrete Wavelet Analysis Based Fault Detection, *WSEAS Transactions on Systems*, Volume 5, Issue 10, October 2006, pp. 2391-2397.

[13] Y. Zhang and L. Wu, Research on Time Series Modeling by Genetic Programming and Wavelet De-noising Performance of the Model, *WSEAS Transactions on Computer Research*, Volume 2, Issue 1, January 2007, pp. 44-49.

[14] Y. Tian, P.L. Lewin, S.J. Sutton, and S.G. Swingler, PD Characterization Using Wavelet Decomposition of Acoustic Emission Signals, *Proceedings of the 2004 International Conference on Solid Dielectrics*, Toulouse, France, July 5-9, 2004.

[15] Y. Tian, P.L. Lewin, A.E. Davies, S.G. Swingler, S.J. Sutton, and G.H. Hathaway, Comparison of On-Line PD Detection Methods for High Voltage Cable Joints, *IEEE Transactions on Dielectrics and Electrical Insulation*, Vol. 9, No. 3, 2002, pp. 604-615.

[16] L. Satish and B. Nazneen, Wavelet-Based Denoising of Partial Discharge Signals Buried in Excessive Noise and Interference, *IEEE Transactions on Dielectrics and Electrical Insulation*, Vol. 10, No. 2, 2003, pp. 354-367.

[17] P. Wang, P.L. Lewin, Y. Tian,, S.J. Sutton, and S.G. Swingler, Application of Wavelet-Based Denoising to Online Measurement of Partial Discharge, *Proceedings of the 2004 International Conference on Solid Dielectrics,* Toulouse, France, July 5-9, 2004.

[18] A. Park, I.W. Jung, S.J. Baek, J.Y. Kim, and S.Y. Na, An Improvement of Skin Cancer Detection by the Fuzzy Algorithm with Ambiguous Pattern, *WSEAS Transactions on Systems*, Volume 6, Issue 9, September 2007, pp. 1265-1269.

[19] M. Ziaii, A.A. Pouyan, and M. Ziaii, Geochemical Anomaly Recognition Using Fuzzy C-Means Cluster Analysis, *WSEAS Transactions on Systems*, Volume 5, Issue 10, October 2006, pp. 2424-2429.

[20] W.Y. Chang and H.T. Yang, Partial Discharge Pattern Recognition of Cast-Resin Current Transformers Using Fuzzy C-Means Clustering Approach, *WSEAS Transactions on Computer Research*, Vol. 3, Issue 3, 2008, pp. 172-181.

[21] N.C. Sahoo and M.M.A. Salama, Trends in Partial Discharge Pattern Classification: A Survey, *IEEE Transactions on Dielectrics and Electrical Insulation*, Vol. 12, No. 2, 2005, pp. 248-264.

[22] R.E. James and B.T. Phung, Development of Computer-based Measurements and Their Application to PD Pattern Analysis, *IEEE Transactions on Dielectrics and Electrical Insulation*, Vol. 2, No. 5, 1995, pp. 838-856.

[23] N. Kovshov and V. Riazanov, About One Approach for Detecting Logical Dependencies in Recognition by Precedents Based on the Genetic Algorithm, *WSEAS Transactions on Computer Research*, Volume 1, Issue 2, December 2006, pp. 152-155.

[24] W. Ziomek, M. Reformat, and E. Kuffel, Application of Genetic Algorithm to Pattern Recognition of Defects in GIS, *IEEE Transactions on Dielectrics and Electrical Insulation*, Vol. 7, No. 2, 2000, pp. 161-168.

[25] K.C. Chang and M.F. Yeh, “Grey Relational Analysis Based Approach for Data Clustering,” *IEE Proceedings-Visual Image Processing*, Vol. 152, No. 2, 2005, pp. 165-172.

[26] M.F. Yeh, Y.J. Chen and K.C. Chang, ECG Signal Pattern Recognition Using Grey Relational Analysis, *Proceedings of the 2004 IEEE International Conference on Networking, Sensing & Control*, 2004, pp. 725-730.
[27] C.C. Wong and H.R. Lai, Generating Fuzzy Control Rules by a Clustering Algorithm Based on a Grey Relational Measure, *Proceedings of the 1999 IEEE International Fuzzy Systems Conference*, Seoul, Korea, 1999, pp. 470-473.
[28] C.C. Wong, C.C. Chen, and H.R. Lai, *Grey Systems-Basic Methods and Applications,* ISBN 957-584-874-8, Gau-Lih Book Co., 2001.