Partial Discharge Pattern Recognition of Cast-Resin Current Transformers Using Fuzzy C-Means Clustering Approach

WEN-YEAU CHANG
Department of Electrical Engineering
St. John's University

HONG-TZER YANG *
* Department of Electrical Engineering
* National Cheng Kung University

No. 499, Sec. 4, Tam King Road, Tamsui, Taipei 251, Taiwan
* No. 1, University Road, Tainan City 701, Taiwan
TAIWAN
changwy@mail.sju.edu.tw
* htyang@mail.ncku.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 fuzzy c-means (FCM) clustering approach to recognize partial discharge patterns of cast-resin current transformer (CRCT). 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 nonlinear principal component analysis (NLPCA) method. The proposed FCM classifier 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 classifier was verified on 250 sets of field-test PD patterns of CRCTs. The test results show that the proposed approach may achieve quite satisfactory recognition of PD patterns.

Key-Words: Cast-Resin Current Transformer, Partial Discharge, Pattern Recognition, Fuzzy C-Means Clustering, Nonlinear Principal Component Analysis

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 [1]. Recently, several methods have been employed for the pattern recognition of PD, including neural networks [3-5], genetic algorithm [6], 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. Neural networks has been applied for spatial variability identification of greenhouse [7], PD pattern recognition of current transformers [8], and PD monitoring technique of gas insulated substation [9]. 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 [10]. 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. 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 SOM can be highly nonlinear, directly showing the similar input vectors in the source space by points close in the two-dimensional target space [11]. 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 [12].

The wavelet analysis method has been used to carry out time-frequency analysis in fault diagnosis [13] and de-noising [14]. Wavelet analysis method has also been applied to identify the PD characteristics by decomposition of acoustic emission signals [15] and PD signal de-noising [16-18].

Genetic algorithm is a search method utilizing the mechanism of natural selection and genetics. The application of genetic algorithm to recognition has become a useful tool in many fields [19]. It has been applied for PD pattern recognition of gas-insulated system [20].

Another method is the fuzzy clustering algorithm [21]. The FCM clustering algorithm is one of the most popular fuzzy clustering algorithms [22]. In this paper, a novel FCM based pattern recognition technique for the PD identification of CRCT is proposed with more effectiveness and robustness than the conventional pattern recognition methods.

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 NLPCA features extraction algorithm is described in Section 4. The principles of FCM and the operation flowchart of the proposed pattern recognition scheme are given subsequently. The experimental results and the analysis using 250 sets of field-test PD patterns from five artificial defect types of CRCTs 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 the last Section.

2 PD Pattern Dataset Creation

In order to investigate the PD features and to verify the classification capabilities of the proposed FCM based pattern recognition technique for different PD types commonly occurring in CRCTs, a PD dataset is needed. The PD dataset was collected from laboratory tests on a series of model CRCTs. The material and process used to manufacture the CRCTs were exactly the same as that of making a field CRCT. The appearance of the model CRCT is shown in Fig. 1. The specifications of model CRCTs are shown in Table 1. Five types of experimental models with artificial defects embedded were made to produce five common PD events in the CRCTs.

The five PD activities include (a) normal PD activity (NM) in standard CRCT, (b) internal cavity discharge (VH) caused by an air cavity inside the epoxy resin insulator on the high-voltage side, as shown in Fig. 2, (c) internal cavity discharge (VL) caused by two cavities inside the epoxy resin insulator on the low-voltage side, as shown in Fig. 3, (d) internal fissure discharge (FH) caused by an air fissure inside the epoxy resin insulator on the high-voltage side, as shown in Fig. 4, (e) internal discharge (MH) caused by a metal-line impurity inside the epoxy resin insulator on the high-voltage side, as shown in Fig. 5.

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. 6. It includes a step-up transformer, capacitor coupling circuit, PD

Fig. 1 The appearance of model CRCT

Table 1 The specifications of model CRCTs

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

detector, and the CRCT 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. 6, 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 [23].

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) are shown in Figs. 7 to 10, respectively. As shown in Figs. 7 to 10, 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 [11]. 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., $y_i = f(x_i)$ [23].

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:

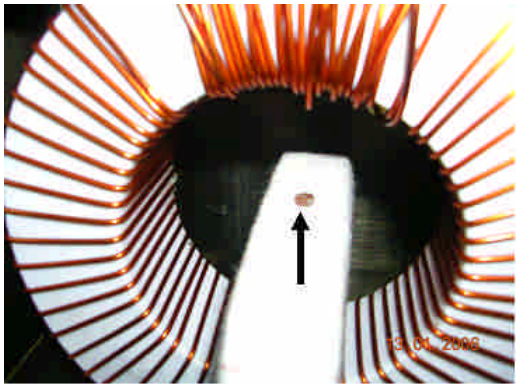

Fig. 2 VH on the high-voltage side of CRCT

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

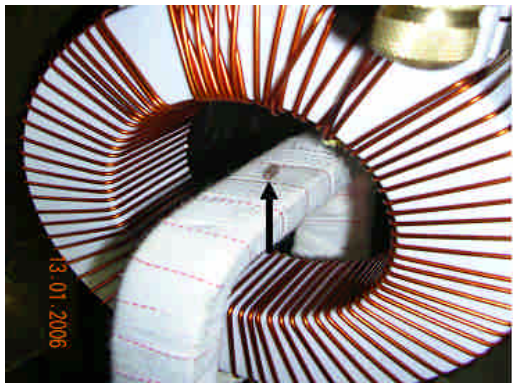

Fig. 4 FH on the high-voltage side of CRCT

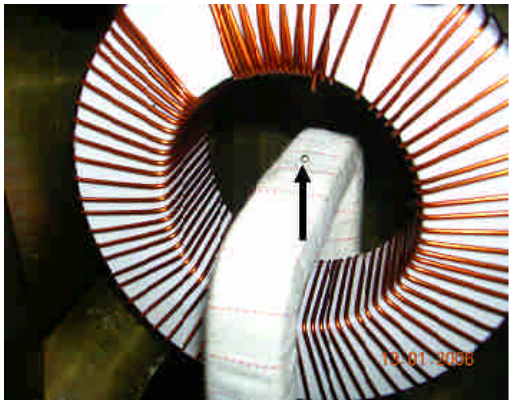

Fig. 5 MH on the high-voltage side of CRCT

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

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

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

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 [23].

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

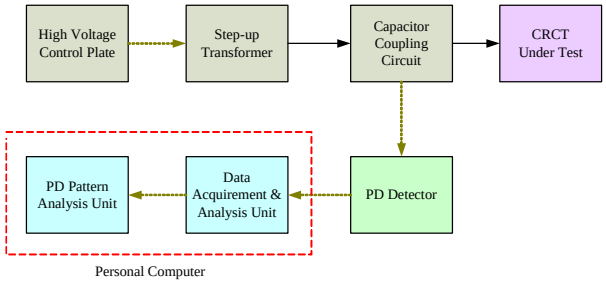

Fig. 6 System configuration of the PD detecting system

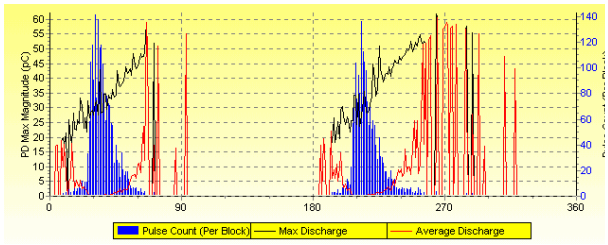

Fig. 7 Typical phase-related distributions of PD for the VH defect

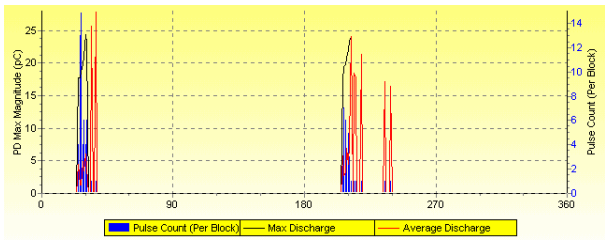

Fig. 8 Typical phase-related distributions of PD for the VL defect

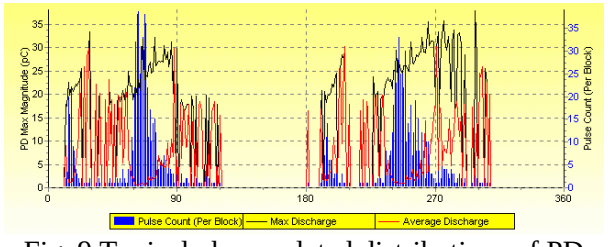

Fig. 9 Typical phase-related distributions of PD for the FH defect

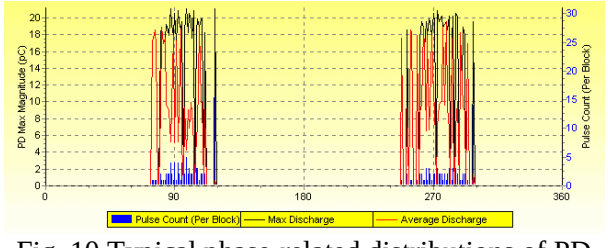

Fig. 10 Typical phase-related distributions of PD for the MH defect

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 typical statistical features extracted by the analysis software (DDX DA3) from PD patterns for the four kinds of defects (VH, VL, FH, and MH) are shown in Figs. 11 to 14, respectively.

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 [24].

4 NLPCA 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 NLPCA method [25-26]. The NLPCA is based on the structure of dual multiplayer neural networks model (DMNN), which contains five layers of neurons, as shown in Fig. 15.

In Fig.15, the DMNN for NLPCA contains two subnetworks of mapping and demapping networks. The mapping from data space to feature space is referred to as the mapping network and the reverse mapping as the demapping network. The neurons at layers 1 and 3 of the network have sigmoid activation functions.

In training, the output vector $\bar{x} = [\bar{x}_1, \bar{x}_2, ..., \bar{x}_n]$, where n is the number of the neurons at the output and input layers, is anticipated to approach to the input data vector $x = [x_1, x_2, ..., x_n]$ at the input layer. As noted, the input layer of the mapping network has neurons equal to the dimensionality of the input data. In this paper n is set to be 24 which is the number of statistical features. After the network is trained, the m neurons at layer 2 represent lower-dimensional nonlinear features $f = [f_1, f_2, ..., f_m]$ extracted from the input data set.

The NLPCA attempts to find the mappings from multidimensional data space to lower-dimensional feature space. In the process, the reconstruction error between input x and output $\bar{x}$ of the dual networks is minimized [25].

$$J = \left\| x - \bar{x} \right\|^2 \tag{7}$$

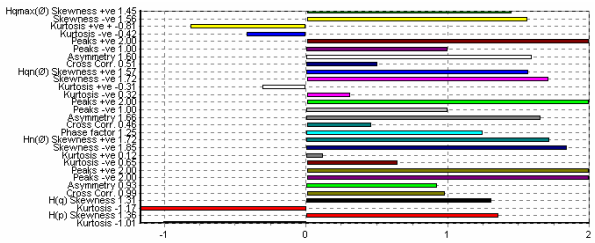

Fig. 11 Typical statistical features of PD for VH

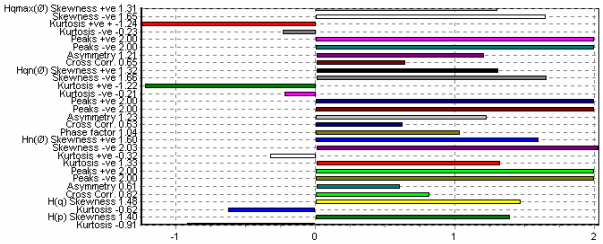

Fig. 12 Typical statistical features of PD for VL

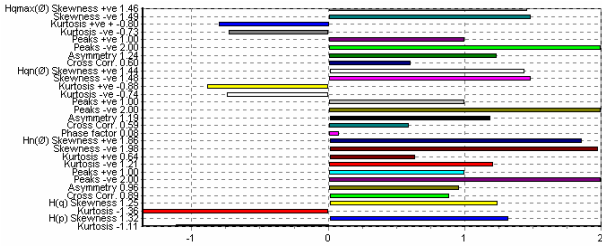

Fig. 13 Typical statistical features of PD for FH

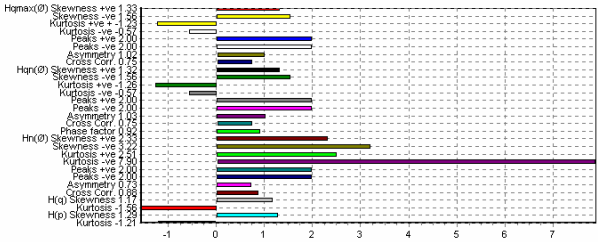

Fig. 14 Typical statistical features of PD for MH

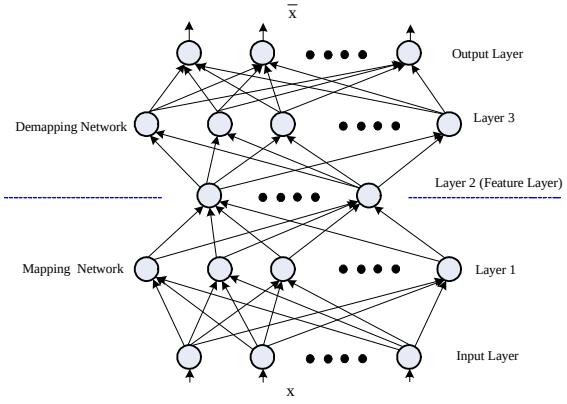

Fig. 15 Architecture of the DMNN in the NLPCA.

The whole network, consisting of the dual networks in the NLPCA, is an autoassociative network where the output vector corresponds to the input vector. The main advantage of NLPCA over principal component analysis is that NLPCA has the ability to stand for nonlinear relationships among the data set of variables.

5 FCM-Based PD Pattern Recognition Method

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

5.1 FCM Algorithm

The FCM has been successfully employed for the data reduction task by providing a tool that recognizes the inherent structure of a given set of data. One of the most important benefits of applying FCM to automate the data reduction task is its mathematical basis for grouping data, rather than subjective one [27].

The FCM algorithm is based on the following objective function [28]:

$$J_m(U,V) = \sum_{k=1}^{N}\sum_{i=1}^{c}(u_{ik})^m d_{ik}^2 \tag{8}$$

Where

$$d_{ik}^2 = \left\|y_k - v_i\right\|_A^2 = (y_k - v_i)^T A(y_k - v_i) \tag{9}$$

In which $y = \{y_1, y_2, \ldots\ldots, y_N\} \subset R^n$ is the data set, N is the number of the data patterns, c is the number of clusters, m is the weighting exponent on each fuzzy membership, u_{ik} is the membership value of the k-th feature vector to cluster i, U is a matrix whose elements are u_{ik}, $V = (v_1, v_2, \ldots.., v_c)$ is the vector of cluster centres, $v_i = (v_{i1}, v_{i2}, \ldots\ldots., v_{in})$ is the centre of cluster i, $\| \bullet \|_A$ is the A-norm on R^n , and A is positive-definite ($n \times n$) weight matrix.

The FCM algorithm tries to minimize J_m, by iteratively updating the partition matrix via the following equations:

$$v_i = \frac{\sum_{k=1}^{N}(u_{ik})^m y_k}{\sum_{k=1}^{N}(u_{ik})^m} \tag{10}$$

$$u_{ik} = \frac{1}{\sum_{j=1}^{c}(d_{ik}/d_{jk})^{2/(m-1)}} \tag{11}$$

The FCM has been applied to power system coherency [22], automated dynamic strain gage data reduction [27], and image segmentation [28]. In this paper, FCM scheme is provided with the training set of PD patterns. Each pattern is represented by a feature vector. The set of feature vectors is clustered for subsequent use in the PD pattern recognition system.

5.2 FCM-based PD Pattern Recognizing Procedure

The proposed FCM-based PD pattern recognition scheme has been successfully implemented using PC-based software (MATLAB) for the PD recognition of CRCTs. The overall flowchart is shown in Fig. 16. The proposed recognition scheme is described briefly in the following steps:

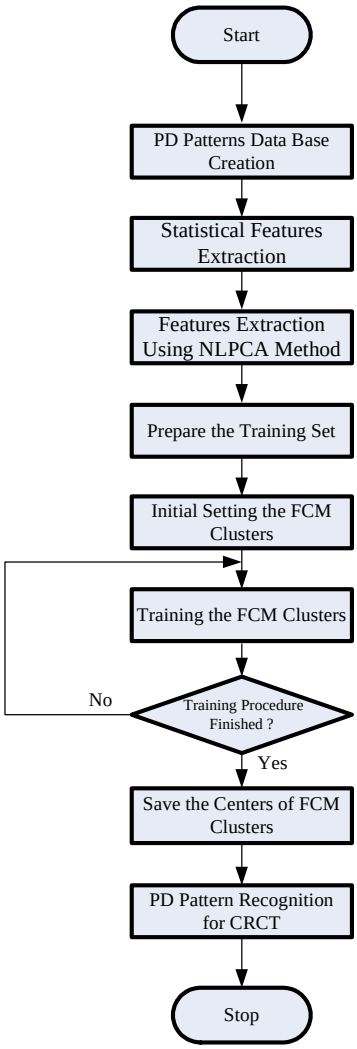

Fig. 16 Flowchart of the FCM-based recognition scheme

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 NLPCA method.
Step4 Prepare the training set for FCM.
Step5 Set the number of clusters c and the weighting exponent m of FCM clusters, and initialize $U^{(0)}$.
Step6 Calculate the centres of clusters $v_i^{(l)}$ using equation (10).
Step7 Calculate $U^{(l)}$ using (11).
Step8 Iterate the training procedure from Steps 6 to 7, till $\| U^{(l)} - U^{(l-1)} \| < \varepsilon$.
Step9 Save the centres of trained FCM clusters $v_i^{(l)}$, as training procedure is finished.
Step10 Use (11) to calculate the membership value to identify the defect types of CRCTs.

6 Experimental Results

To verify the proposed approach, a practical experiment is conducted to demonstrate the effectiveness of the PD pattern recognition scheme. Five types of experimental models with artificial defects are purposely embedded to produce five common PD events in CRCTs.

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. But, some of the statistical features are futile for pattern recognition. So, the features extraction of feature vector from statistical features will influence the accuracy of pattern recognition. In this paper, the significant features are extracted from statistical features by using NLPCA method. To evaluate the optimal number of features for feature vector, we set up three systems of feature vectors. In System 1, the feature vector includes 10 features. In System 2, the feature vector includes 12 features; and in System 3 the feature vector includes 14 features.

To extract different number of the features from statistical features, the structure of NLPCA must be determined. In Systems 1 to 3, the number of neurons in the input layers and output layer of NLPCA is designed to comprise the 24 statistical featuring operators mentioned above. In System 1, the number of neurons in the layer 2 of NLPCA is set to be 10. The numbers of neurons in the layers 2 of NLPCA are set to be 12 and 14 for System 2 and System 3, respectively. The structures of 3 types of NLPCA are shown in Table 2. 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 of three systems, the training procedure of FCM clustering starts. 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. During the training process, the number of clusters c was set to be 5 for five types of defects, and the weighting exponent m was set to be 1.6 based on experience. After the training process, the centres of trained FCM clusters were saved. The centres of 3 trained FCM clusters are shown in Tables 3 to 5, respectively.

Table 2 The structures of 3 types of NLPCA

System	Neurons in Input Layer	Neurons in Layer 1	Neurons in Layer 2	Neurons in Layer 3	Neurons in Output Layer
System 1	24	17	10	17	24
System 2	24	18	12	18	24
System 3	24	19	14	19	24

Table 3 The centres of trained FCM clusters for System 1

Feature	Defect Types				
	NM	VH	VL	FH	MH
#1	0.332	0.246	0.356	0.335	0.672
#2	0.021	0.358	0.201	0.447	0.012
#3	0.238	0.112	0.056	0.668	0.278
#4	0.452	0.569	0.012	0.447	0.732
#5	0.107	0.338	0.109	0.208	0.663
#6	0.545	0.721	0.443	0.443	0.432
#7	0.012	0.743	0.342	0.379	0.278
#8	0.783	0.390	0.334	0.294	0.443
#9	0.211	0.226	0.279	0.106	0.221
#10	0.202	0.621	0.167	0.523	0.390

To verify the training results of FCM clusters, the training data were applied to the trained FCM clusters again. Tables 6 to 8 show the test results of the training data for Systems 1 to 3, respectively. The data in Tables 6 to 8 shows that the proposed method has 100% accuracy for the 150 training feature vectors in three systems. Tables 9 to 11 demonstrate the promising performance when 100 testing patterns of three systems were tested. It is shown in Table 9 that among the 100 testing patterns of System 1, there are only 2 errors of recognition, one for FH and the other for MH defects. It is shown in Table 10 that the proposed method has 100% accuracy for the 100 testing patterns of System 2. As shown in Table 11 among the 100 testing patterns of System 3, there is only one error of recognition for MH defect.

The test results show that the proposed method is able to accurately recognize the testing defects for three systems. The number of features in the feature vector will influence the accuracy of pattern recognition. The optimal combination of feature vector is System 2.

7 Conclusions

This paper has proposed an FCM based pattern recognition technique for PD of CRCTs. The effectiveness of the proposed technique has been verified using experimental results. It has been shown that through the NLPCA feature extraction procedure, the extracted feature vectors can significantly reduce the size of the PD pattern database. Also, the FCM based PD pattern recognition scheme is very effective for clustering the defects of CRCTs. The FCM based PD pattern recognition scheme can be applied to other high-voltage equipments such as transformer, circuit breaker, and generator.

The experimental results show that the number of features in the feature vector influences the accuracy of pattern recognition. The directions for future research of the FCM based PD pattern recognition scheme can be described as follow: To further improve the recognition accuracy of the proposed approach, the optimal search methods for the optimal combination selection of feature vectors can be investigated and integrated in the proposed FCM based PD pattern recognition for the CRCTs and other high-voltage equipment. Besides, the proposed recognition approach is based on the PD dataset collected from a series of model CRCTs. The content of PD dataset influences the accuracy of pattern recognition. To ameliorate further the recognition accuracy of the proposed approach, the

Table 4 The centres of trained FCM clusters for System 2

Feature	Defect Types				
	NM	VH	VL	FH	MH
#1	0.394	0.643	0.348	0.523	0.379
#2	0.227	0.389	0.237	0.431	0.278
#3	0.478	0.467	0.521	0.623	0.602
#4	0.642	0.211	0.233	0.321	0.368
#5	0.277	0.732	0.456	0.345	0.233
#6	0.186	0.197	0.201	0.327	0.378
#7	0.568	0.378	0.721	0.540	0.504
#8	0.297	0.289	0.367	0.419	0.467
#9	0.442	0.397	0.421	0.728	0.325
#10	0.397	0.489	0.236	0.212	0.275
#11	0.197	0.752	0.629	0.356	0.449
#12	0.624	0.233	0.189	0.228	0.208

Table 5 The centres of trained FCM clusters for System 3

Feature	Defect Types				
	NM	VH	VL	FH	MH
#1	0.425	0.512	0.216	0.219	0.265
#2	0.209	0.219	0.427	0.286	0.331
#3	0.563	0.431	0.222	0.318	0.294
#4	0.217	0.189	0.381	0.641	0.186
#5	0.167	0.318	0.287	0.324	0.528
#6	0.201	0.443	0.329	0.218	0.428
#7	0.408	0.209	0.210	0.228	0.143
#8	0.372	0.317	0.482	0.339	0.327
#9	0.184	0.228	0.308	0.497	0.374
#10	0.201	0.189	0.219	0.129	0.216
#11	0.310	0.286	0.312	0.175	0.286
#12	0.298	0.323	0.107	0.336	0.249
#13	0.119	0.186	0.186	0.329	0.335
#14	0.286	0.332	0.218	0.215	0.381

Table 6 Recognition performance of training data for System 1 (150 patterns)

Pattern	Defect Types	Accuracy Rate
Training Data	NM	100%
	VH	100%
	VL	100%
	FH	100%
	MH	100%

more plenteous PD dataset creation methods will be studied in the future researches.

References:

[1] L. Niemeyer, A Generalized Approach to Partial Discharge Modeling, *IEEE Transactions on Dielectrics and Electrical Insulation*, Vol. 2, No. 4, 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. Cabrerizo, M. Adjouadi, M. Ayala, I. Yaylali, A. Barreto, and N. Rishe, EEG Analysis Using Neural Networks for Seizure Detection, *Proceedings of the 6th WSEAS International Conference on Signal Processing, Robotics and Automation*, Corfu Island, Greece, 2007, pp. 121-126.

[4] S.D. Sawarkar, A.A. Ghatol, and A.P. Pande, Neural Network Aided Breast Cancer Detection and Diagnosis Using Support Vector Machine, *Proceedings of the 7th WSEAS International Conference on Neural Networks*, Cavtat, Croatia, 2006, pp.158-163.

[5] M.M.A. Salama and R. Bartnikas, Determination of Neural Network Topology for Partial Discharge Pulse Pattern Recognition, *IEEE Transactions on Neural Networks*, Vol. 13, No. 2, 2002, pp. 446-456.

[6] H.P. Chiu, K.L. Hsieh, Y.T. Tang, and W.J. Chien, A Knowledge-Based Genetic Algorithm for the Job Shop Scheduling Problem, *Proceedings of the 6th WSEAS Int. Conf. on Artificial Intelligence, Knowledge Engineering and Data Bases*, Corfu Island, Greece, 2007, pp. 81-85.

[7] M.A. Bussab, J.I. Bernardo, and A. R. Hirakawa, Neural Networks Modeling in Greenhouse with Spatial Variability Identification, *WSEAS Transactions on Computer Research*, Volume 2, Issue 2, 2007, pp. 214-219.

[8] 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.

[9] 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, 1997, pp. 1014-1021.

Table 7 Recognition performance of training data for System 2 (150 patterns)

Pattern	Defect Types	Accuracy Rate
Training Data	NM	100%
	VH	100%
	VL	100%
	FH	100%
	MH	100%

Table 8 Recognition performance of training data for System 3 (150 patterns)

Pattern	Defect Types	Accuracy Rate
Training Data	NM	100%
	VH	100%
	VL	100%
	FH	100%
	MH	100%

Table 9 Recognition performance of testing data for System 1 (100 patterns)

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

Table 10 Recognition performance of testing data for System 2 (100 patterns)

Pattern	Defect Types	Accuracy Rate
Testing Data	NM	100%
	VH	100%
	VL	100%
	FH	100%
	MH	100%

Table 11 Recognition performance of testing data for System 3 (100 patterns)

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

[10] 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.

[11] 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.

[12] W.Y. Chang and H.T. Yang, Partial Discharge Pattern Recognition of Molded Type Transformers Using Self Organizing Map, *Proceedings of the 8th International Conference on Properties and Application of Dielectric Materials*, ICPADM 2006, Bali, Indonesia, 2006, pp. 246-249.

[13] J. Pi and M.F. Liao, Rolling Bearing Fault Diagnosis with Wavelet-Based Method, *WSEAS Transactions on Computer Research*, Volume 2, Issue 1, 2007, pp. 1-7.

[14] 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, 2007, pp. 44-49.

[15] 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, 2004.

[16] 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.

[17] 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.

[18] 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, 2004.

[19] 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, 2006, pp. 152-155.

[20] 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.

[21] M.S. Yang, W.L. Hung, and C.H. Chang, A Penalized Fuzzy Clustering Algorithm, *WSEAS Transactions on Computer Research*, Volume 1, Issue 2, 2006, pp. 83-88.

[22] S.C. Wang and P.H. Huang, Fuzzy C-Means Clustering for Power System Coherency, *Proceedings of the 2005 IEEE International Conference on Systems, Man and Cybernetics*, 2005, pp. 2850-2855.

[23] 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.

[24] 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.

[25] T. Stamkopoulos, K. Diamantaras, N. Maglaveras, and M. Strintzis, ECG Analysis Using Nonlinear PCA Neural Networks for Ischemia Detection, *IEEE Transactions on Signal Processing*, Vol. 46, 1998, pp. 3058-3066.

[26] H.T. Yang, S.C. Chen and W.N. Tsai, Classification of Direct Load Control Curves for Performance Evaluation, *IEEE Transactions on Power Systems*, Vol. 19, 2004, pp. 811-817.

[27] G.J. Mascoli, Automated Dynamic Strain Gage Data Reduction Using Fuzzy C-Means Clustering, *Proceedings of the 1995 IEEE International Conference on Fuzzy Systems*, 1995, pp. 2207-2214.

[28] T. Pham, M. Wagner, and D. Clark, Applications of Genetic Algorithms, Geostatistics, and Fuzzy C-Means Clustering to Image Segmentation, *Proceedings of the 2001 Congress on Evolutionary Computation*, 2001, pp. 741-746