

Partial Discharge Pattern Recognition of Cast Resin Current Transformers Using Radial Basis Function Neural Network

Wen-Yeou Chang[†]

Abstract – This paper proposes a novel pattern recognition approach based on the radial basis function (RBF) neural network for identifying insulation defects of high-voltage electrical apparatus arising from partial discharge (PD). Pattern recognition of PD is used for identifying defects causing the PD, such as internal discharge, external discharge, corona, etc. This information is vital for estimating the harmfulness of the discharge in the insulation. Since an insulation defect, such as one resulting from PD, would have a corresponding particular pattern, pattern recognition of PD is significant means to discriminate insulation conditions of high-voltage electrical apparatus. To verify the proposed approach, experiments were conducted to demonstrate the field-test PD pattern recognition of cast resin current transformer (CRCT) models. These tests used artificial defects created in order to produce the common PD activities of CRCTs by using feature vectors of field-test PD patterns. The significant features are extracted by using nonlinear principal component analysis (NLPCA) method. The experimental data are found to be in close agreement with the recognized data. The test results show that the proposed approach is efficient and reliable.

Keywords: Partial discharge, Pattern recognition, Radial basis function neural network, Cast resin current transformer, Defect of insulation, Nonlinear principal component analysis

1. Introduction

PD measurement and pattern recognition are important tools for improving the reliability of high-voltage insulation systems [1]. The pattern recognition of PD aims to identify potential insulation defects from the measured data, and the potential defects can then be used for estimating the risk of insulation failure in high-voltage electrical apparatus.

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

Because there has been substantial progress in the physical understanding of PD during the last decade, this knowledge can now be exploited to support the interpretation of insulation defects. Recently, several methods have been employed for the pattern recognition of PD, including neural network, expert systems, self organizing maps, wavelet analysis, and the grey clustering

method.

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

The recognition of PD pattern and the evaluation of insulation performance are relatively complicated, a task which often must be completed by decision tree method. Decision tree method for the diagnostics of SF₆ decomposition products have been developed [4].

In contrast to other clustering mapping methods for unsupervised data, the mapping relationship of a self organizing map can be highly nonlinear, directly showing the similar input vectors in the source space by points close to the two-dimensional target space. Along with the similarity of the input data, a self organizing map may lead to classification results, and this technique has been applied for PD pattern recognition of CRCT [5].

The wavelet analysis method is a useful tool in fault detection and de-noising, and this method has also been applied to analysis of power transformer partial discharge signals [6]. Grey system theory is a useful methodology for systems with incomplete information. Grey relational analysis can be used to analysis the relationships between

[†] Corresponding Author: Dept. of Electrical Engineering, St. John's University, Taiwan. (changwy@mail.sju.edu.tw)

Received: September 20, 2012; Accepted: September 24, 2012

one major sequence and the other comparative ones in a given set. The application of grey clustering approach has been proposed for recognizing partial discharge patterns of the high-voltage equipment [7].

In this paper the PD patterns are measured 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 statistical operators are extended extracted by using the NLPCA scheme. After feature extraction, this paper proposes the application of RBF neural network to recognize partial discharge patterns of CRCT.

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 algorithm of statistical feature extraction is described in Section 3. The NLPCA features extraction algorithm is described in Section 4. The principles of RBF neural network and the operation flowchart of the proposed pattern recognition scheme are given in the next section. The experimental results and the analysis using 250 sets of field-test PD patterns from high-voltage 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 Section 7.

2. PD Patterns Database Creation

In order to investigate the PD features and to verify the classification capabilities of the proposed RBF neural network based pattern recognition approach for different PD types commonly occurring in high-voltage electrical apparatus, a PD dataset is needed. The PD dataset for this study was collected from laboratory PD tests on a series of model CRCTs. The materials and process used to manufacture these high-voltage CRCTs were exactly the same as that of making the field equipment.

The appearance of a 12kV, 40VA model CRCT is shown in Fig. 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



Fig. 1. The appearance of model CRCT

cavity discharge caused by 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 Figs. 4; and (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 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 detector, and the CRCT under test. Through the testing processes, all the data measured were digitally converted in order to store them in

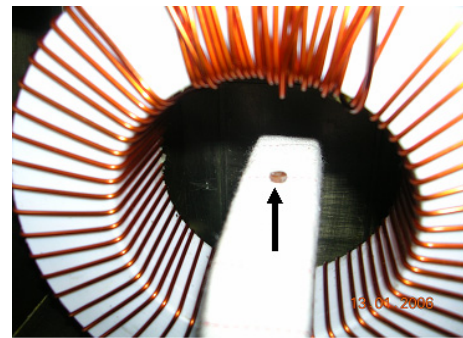


Fig. 2. VH on the high-voltage side of CRCT



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

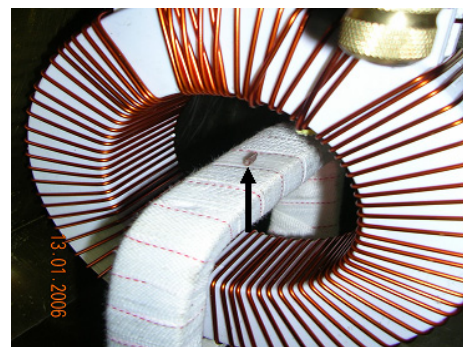


Fig. 4. FH on the high-voltage side of CRCT

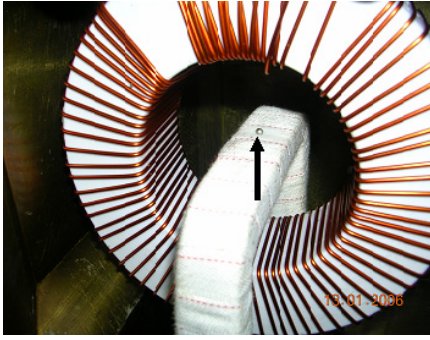


Fig. 5. MH on the high-voltage side of CRCT

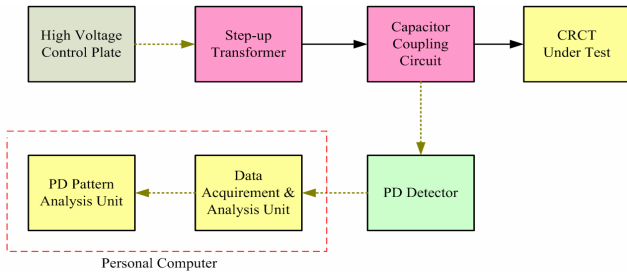


Fig. 6. System configuration of the PD detecting system

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 detecting system, 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 plots these data as functions of the phase positions.

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 types of defects (VH, VL, FH, and MH) of the insulation models are shown in Figs. 7 to 10, respectively. As shown in Figs. 7 to 10, the PD patterns of different defects display discriminative features.

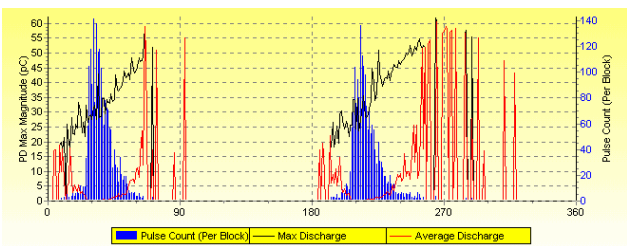


Fig. 7. Typical phase-related distributions of PD for VH

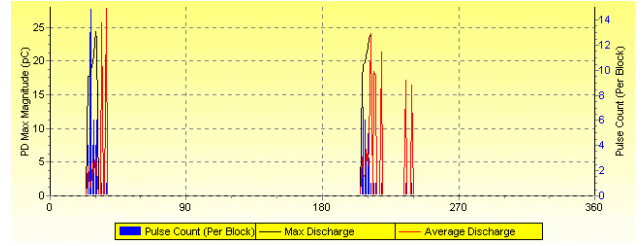


Fig. 8. Typical phase-related distributions of PD for VL

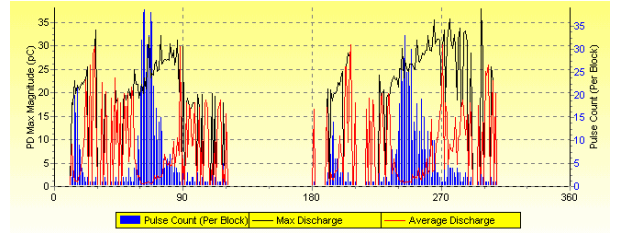


Fig. 9. Typical phase-related distributions of PD for FH

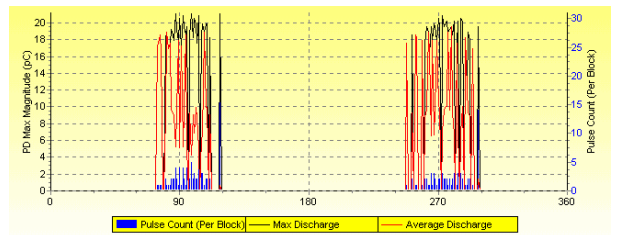


Fig. 10. Typical phase-related distributions of PD for MH

3. Statistical Feature Extraction

In PD pattern recognition, feature extraction is a technique essential to reduce the dimension of the original data [8]. The features are intended to denote the characteristics of different PD statuses. Several statistical methods of feature extraction are described in this section, and five statistical operators are extracted from the 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., $y_i = f(x_i)$ [9].

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

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

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

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

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

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

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

In (3) and (4), x_i is the statistical value in the phase window i , p_i is the related probability of appearance.

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 positive and negative cycles. It can be expressed as:

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

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 the phase window i in the negative cycle, and q_i^+ is the amplitude of the PD pulse at the phase window i in the positive cycle.

Cross correlation factor (C_c) indicated the difference in sharp of the distributions in the positive and negative half cycles. $C_c = 1$ means that the sharps are totally symmetric and $C_c = 0$ means that sharps are totally asymmetric. The cross correlation factor 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 windows per half cycle.

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

The use of statistical features rather than recording the distribution profiles can significantly reduce the dimension of the database. To a certain extent, they can be used for characterizing PD patterns with reasonable discrimination [10].

4. NLPCA Based Feature Extraction Scheme

The statistical feature extraction methods were used to extract 24 statistical features for patterns. But since some of the statistical features are futile for pattern recognition, 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 the NLPCA scheme [11]. The NLPCA is based on the structure of dual multiplayer neural networks model (DMNN), which contains five layers of neurons, as shown in Fig. 11.

In Fig. 11, the DMNN for NLPCA contains two subnetworks of mapping network and demapping network. 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 $x' = [x'_1, x'_2, \dots, x'_n]$, where n is the number of the neurons at the output and input layers, is anticipated to approach to the input data vector $x = [x_1, x_2, \dots, x_n]$ 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

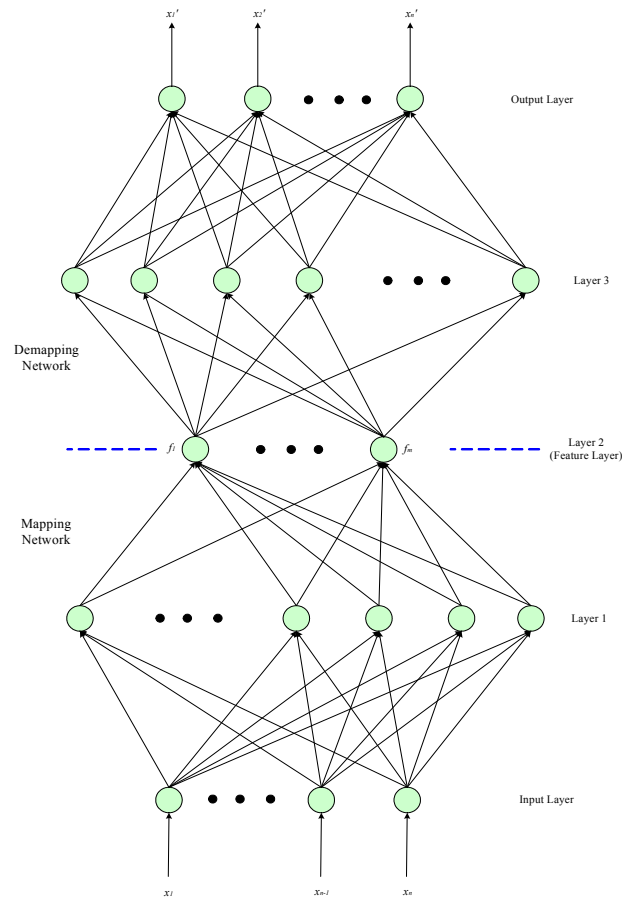


Fig. 11. Architecture of the DMNN in the NLPCA

24 which is the number of statistical features. After the network is trained, the m neurons at layer 2 (feature layer) represent lower-dimensional nonlinear features $f = [f_1, f_2, \dots, f_m]$ 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 x' of the dual networks is minimized [12].

$$J = \|x - x'\|^2 \quad (7)$$

The whole network, consisting of the dual networks in the NLPCA, is an auto-associative 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. RBF Neural Network Based PD Pattern Recognition Approach

The RBF neural network is a useful methodology for systems with incomplete information. It can be used to analyze the relationships between one major sequence and the other comparative ones in a given set. In this section, the algorithms of RBF neural network and the RBF neural network-based PD pattern recognition scheme are described. The PD recognition through a RBF neural network in multidimensional feature space is also validated on the basis of the features extracted by the NLPCA scheme, as mentioned above.

5.1 Principals of RBF neural network

The RBF neural network is a forward network models with universal approximation capabilities, and which is employed to approximate the function [13]. It is a multi-input, multi-output system consisting of an input layer, a hidden layer, and an output layer. During data processing, the hidden layer performs nonlinear transforms for the feature extraction and the output layer gives a linear combination of output weights [14]. The structure is shown in Fig. 12.

The network actually performs a nonlinear mapping from the input space R^d to the output space R^n . The mapping relationship between input vector and output vector of RBF neural network is based on the following function:

$$RBF \text{ Neural Network} : \begin{cases} R^d \rightarrow R^n \\ \bar{x} \rightarrow \bar{o} \end{cases} \quad (8)$$

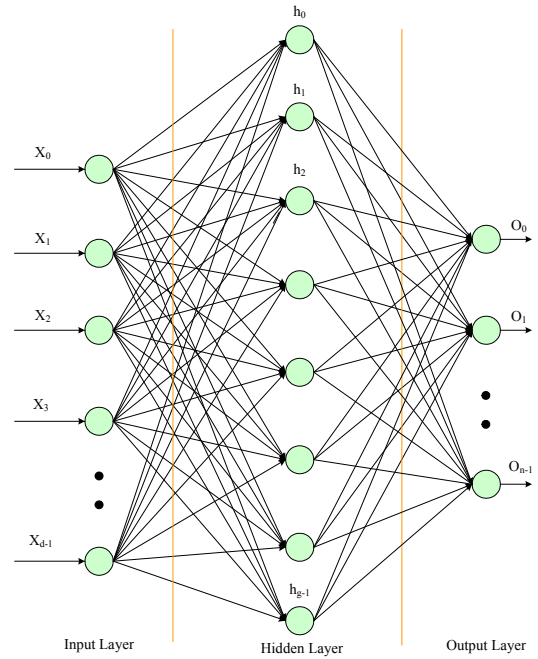


Fig. 12. Architecture of the RBF neural network system

where input vector $\bar{x}_i = \{x_i, \text{ for } i = 0, 1, 2, \dots, d-1\}$, output vector $\bar{o}_i = \{o_i, \text{ for } i = 0, 1, 2, \dots, n-1\}$.

Each hidden neuron computes a Gaussian function in the following equation

$$h_j(\bar{x}) = \exp\left[-\frac{(\bar{x} - \bar{\mu}_j)^T (\bar{x} - \bar{\mu}_j)}{2\sigma_j^2}\right], \text{ for } j = 0, 1, 2, \dots, g-1 \quad (9)$$

where μ_j and σ_j are, respectively, the center and the width of the Gaussian potential function of the j th neuron in the hidden layer.

Each output neuron of the RBF neural network computes a linear function in the following form:

$$o_k = \sum_{j=0}^{g-1} w_{kj} h_j(\bar{x}) - \theta_k, \text{ for } k = 0, 1, 2, \dots, n-1 \quad (10)$$

where o_k is output of the k th node in the output layer, w_{kj} is weight between j th node in the hidden layer and k th node in the output layer, $h_j(\bar{x})$ is output from the j th node in the hidden layer, θ_k is bias of the k th node in the output layer.

5.2 Training procedure of RBF neural network

The training procedure of RBF neural network is composed of a two-step decomposition: estimating μ_j and σ_j and estimating the weights between the hidden layer and output layer. The procedure of μ_j and σ_j estimation is described briefly in the following steps:

Step 1 Take any point μ_j and its associated width σ_j (initially $\sigma_j = 0$);

- Step 2 Use the Euclidean distance to find the nearest point μ_l of the same class;
- Step 3 Compute the mean of these two points to obtain a new point with its associated width by using following equation $\sigma = (\|\mu_j, \mu_l\|) / 2 + \sigma_j$;
- Step 4 Compute the distance D from the new mean to the nearest point of all other classes;
- Step 5 If $D < 2\sigma$, then accept the merge of μ_j and μ_l and start again from Step 2; if $D \geq 2\sigma$, reject the merge and recover the two original points and their widths, restart from Step 1;
- Step 6 Repeat Steps 1–5 until all clusters of each class are used.

The procedure of estimating the weights between the hidden and output layer is described briefly as follows. After the Gaussian function centres and widths are computed from training vectors, the connection weights between the hidden and output layers can be calculated using the pseudo-inverse matrix method.

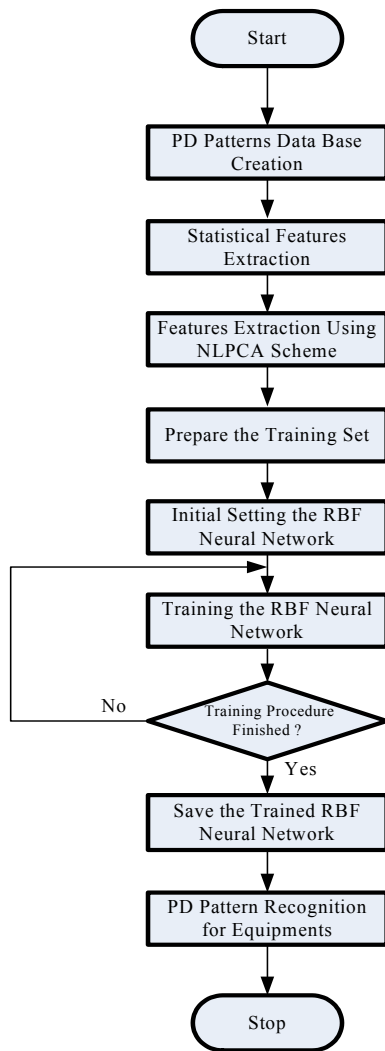


Fig. 13. Flowchart of the RBF neural network based recognition scheme

5.3 RBF neural network based PD pattern recognizing procedure

The proposed PD pattern recognition scheme based on RBF neural network has been successfully implemented using PC-based software for the PD recognition. The overall flowchart is shown in Fig. 13, and the proposed recognition scheme is described briefly in the following steps:

- Step 1 Create data base of the phase-related distributions of PD patterns;
- Step 2 Extract the statistical features from phase-related distributions;
- Step 3 Extract the significant features from statistical features by using the NLPCA scheme;
- Step 4 Prepare the training set for the RBF neural network;
- Step 5 Use the training set to train the RBF neural network for PD pattern recognition;
- Step 6 Save the Gaussian functions centres, widths and connection weights between the hidden and output layers of trained RBF neural network when the training procedure is finished;
- Step 7 Use the trained RBF neural network to identify the defect types of PD patterns.

Even though a conclusion may review the main results or contributions of the paper, do not duplicate the abstract or the introduction. For a conclusion, you might elaborate on the importance of the work or suggest the potential applications and extensions.

6. Experiment 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 model CRCTs. The test results show that the proposed approach is able to accurately recognize the testing defects.

Five types of experimental models with artificial defects are purposely embedded to produce five common PD events in CRCT. The proposed approach has been implemented according to the field-test PD patterns collected from our laboratory. The input data for this 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 distributions $H_n(\phi)$.

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.

Statistical feature extraction methods are used to extract 24 statistical features for each pattern. Upon applying S_k , K_u , P_e , D_a and C_c to both positive and negative cycles of

Table 1. 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 2. Recognition performance of testing data in CRCT tests

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

$H_{qmax}(\phi)$, $H_{qn}(\phi)$, and $H_n(\phi)$, a total of 24 statistical features have been extracted from a PD pattern. The significant features are extracted from statistical features by using the NLPCA scheme. There are 10 extracted features for feature vector extracted from 24 statistical features using the NLPCA scheme in this experiment. After the 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 RBF neural network was started. The training data consist of 150 feature vectors, which are randomly chosen from the 250 feature vectors of sample data. The remaining 100 feature vectors were used as the testing data.

To verify the training results of RBF neural network, the training data were applied to the trained RBF neural network again. Table 1 shows the test results of the training data. The data in Table 1 show that the proposed approach has 100% accuracy for the 150 training feature vectors, because the training process of RBF neural network was stop under the error is lower than 0.0001. Table 2 demonstrates the promising performance when 100 testing patterns were tested. Table 2 shows that among the 100 testing patterns, there were only 3 errors of recognition, one for VH, one for FH, and the other for MH defects. The total accuracy rate of 100 testing patterns is 97%.

7. Conclusion

This paper has proposed an RBF neural network based pattern recognition technique for PD of high-voltage equipment. 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. In addition, the PD pattern recognition scheme, based on RBF neural network is very effective for clustering the defects of high-

voltage equipment. The content of PD dataset influences the accuracy of pattern recognition. To further improve the recognition accuracy of the proposed approach, more extensive PD dataset creation methods will be examined in future studies.

References

- [1] J. Y. Jeong, D. S. Kang, J. H. Sun, J. C. Heo and C. H. Park, "Assessment of 23 kV Capacitive Coupler for On-line Partial Discharge Measurements," Journal of Electrical Engineering & Technology, vol. 4, no.1, pp. 123-130, March 2009.
- [2] A. Rodrigo, P. Llovera, V. Fuster and A. Quijano, "Study of Partial Discharge Charge Evaluation and the Associated Uncertainty by Means of High Frequency Current Transformers," IEEE Trans. Dielectrics and Electrical Insulation, vol. 19, no. 2, pp. 434-442, April 2012.
- [3] M. Oskooee, A.R. Yazdizadeh and H.R. Mahdiani, "A New Feature Extraction and Pattern Recognition of Partial Discharge in Solid Material by Neural Network," in Proceedings of the Eighth International Conference on Natural Computation, Chongqing, China, May 2012.
- [4] J. Tang, F. Liu, Q.H. Meng, X.X. Zhang and J.G. Tao, "Partial Discharge Recognition Through an Analysis of SF6 Decomposition Products Part 2: Feature Extraction and Decision Tree-Based Pattern Recognition," IEEE Trans. Dielectrics and Electrical Insulation, vol. 19, no. 1, pp. 37-44, February 2012.
- [5] W.Y. Chang and H.T. Yang, "Application of Self Organizing Map Approach to Partial Discharge Pattern Recognition of Cast-Resin Current Transformers," WSEAS Trans. Computer Research, vol. 3, no. 3, pp. 142-151, March 2008.
- [6] S.W. Gao, M. Zhang, C.H. Yuan, X.T. Sun and C. Zhang, "Wavelet Packet Analyzing of Power Transformer Partial Discharge Signals," in Proceedings of the 2011 International Conference on Control, Automation and Systems Engineering, Singapore, July 2011.
- [7] W.Y. Chang, "Application of Grey Clustering Approach and Genetic Algorithm to Partial Discharge Pattern Recognition," WSEAS Trans. Systems, vol. 8, no. 12, pp. 1273-1283, December 2009.
- [8] R.J. Liao, K. Wang, L.J. Yang, T.C. Zhou and S.X. Zheng, "Study on Thermal Aging Condition Assessment of Oil-paper Insulation Based on Statistical Features of Partial Discharge," in Proceedings of the IEEE 9th International Conference on Properties and Applications of Dielectric Materials, Harbin, China, July 2009.
- [9] N.C. Sahoo and M.M.A. Salama, "Trends in Partial Discharge Pattern Classification: A Survey," IEEE Trans. Dielectrics and Electrical Insulation, vol. 12, no. 2, pp. 248-264, April 2005.

- [10] R.E. James and B.T. Phung, "Development of Computer-based Measurements and Their Application to PD Pattern Analysis," *IEEE Trans. Dielectrics and Electrical Insulation*, vol. 2, no. 5, pp. 838-856, October 1995.
- [11] T.A. Reddy, K.R. Devi and S.V. Gangashetty, "Nonlinear Principal Component Analysis for Seismic Data Compression," in *Proceedings of the 1st International Conference on Recent Advances in Information Technology*, Dhanbad, India, March 2012.
- [12] T. Stamkopoulos, K. Diamantaras, N. Maglaveras, and M. Strintzis, "ECG Analysis Using Nonlinear PCA Neural Networks for Ischemia Detection," *IEEE Trans. Signal Processing*, vol. 46, no. 11, pp. 3058-3066, November 1998.
- [13] Ali Karami, "Radial Basis Function Neural Network for Power System Transient Energy Margin Estimation," *Journal of Electrical Engineering & Technology*, vol. 3, no.4, pp. 468-475, December 2008.
- [14] Y. Zhang, Q. Zhou, C. Sun, S. Lei, Y. Liu and Y. Song, "RBF Neural Network and ANFIS-based Short-term Load Forecasting Approach in Real-time Price Environment," *IEEE Trans. Power Systems*, vol. 23 no. 3, pp. 853-858, August 2008.



Wen-Yeau Chang He received the Ph.D. degree in electrical engineering from National Cheng-Kung University, Tainan, Taiwan in 1994. His current research interests are high voltage engineering, renewable generation systems, power electronics, and distributed generators.