Scientific Research Publishing

Applied Mathematics, 2022, 13, 896-916
https://www.scirp.org/journal/am
ISSN Online: 2152-7393
ISSN Print: 2152-7385

Research on Intelligent Identification of PD Patterns Based on the Fingerprint Features

Qiuping Zheng[1], Ting Chen[2], Haitao Hu[2], Yingli Wang[2], Dawei Zhao[2], Chuntian Chen[2], Dianchun Zheng[2]

[1]Instrumentation Technology & Economy Institute, Beijing, China
[2]Harbin University of Science and Technology, Harbin, China
Email: zhengqiupingnicole@hotmail.com, zhengdianchun@126.com

How to cite this paper: Zheng, Q.P., Chen, T., Hu, H.T., Wang, Y.L., Zhao, D.W., Chen, C.T. and Zheng, D.C. (2022) Research on Intelligent Identification of PD Patterns Based on the Fingerprint Features. *Applied Mathematics*, **13**, 896-916.
https://doi.org/10.4236/am.2022.1311057

Received: October 21, 2022
Accepted: November 27, 2022
Published: November 30, 2022

Abstract

Five-electrode configurations were designed to simulate the distribution inhomogeneity of electric field intensities in the air-insulating medium, and the characteristic data waveforms of partial discharge generated by different electrode configurations under the excitation of power frequency AC voltage were carefully collected in this paper. Furthermore, the feature vectors of the corresponding fingerprint, contained in partial discharge data, were extracted by rigorous mathematical algorithms, and the artificial neural network was employed to realize the pattern recognition of partial discharge caused by the inhomogeneity of electric field intensity with different electrode configurations. The results indicate that the J_4 value in the space of 7 feature quantities is 1905.6, and the recognition rate is 100% when the hidden layer neuron of the network is 19. However, the J_5 value of 9 feature quantities is 1589.9, and the purpose of recognition has been achieved when the number of hidden layer neurons of the network is 6. Increasing the number of hidden layer neurons will only waste computing resources. Of course, PD information collection mode, feature quantity selection, optimal feature space composition, network structure and classification algorithm are the key to realizing PD fault intelligence identification.

Keywords

PD, Fingerprint, Feature Extraction, Pattern Recognition, Class Separability

1. Introduction

Practical experience and rigorous theoretical studies have proved that the Partial Discharge (abbreviated PD) phenomena are closely related to the insulation defects or failures that happen in operating equipment such as switchgear, power

transformers, cables, as well as motor and generator stator windings [1] [2] [3]. Therefore, in addition to detection technology, data mining algorithms and processing techniques of PD signals and information have become a research hotspot in recent years.

With the advancement of high-precision detection technology and the wide application of artificial intelligence algorithms, the great improvement of online monitoring and diagnosis technology for insulation faults of high-voltage equipment has been achieved to a large extent [4] [5].

Over the past couple of decades, several classification algorithms of PD patterns have been proposed and tested, which adopt different approaches, using one or more of the different PD data patterns [6].

As well known, the PD phenomenon of gas insulation systems strongly depends on the non-uniformity of the electric field and the characteristics of PD signals also obey certain mathematical statistical laws, so five kinds of electrode structures are designed to artificially simulate non-uniformity of the electric field, and PD signals of time-domain for five models are collected by specially designed experimental circuit, and the corresponding PD fingerprints are obtained through analysis and calculation of the data information. Furthermore, different combinations of fingerprint features extracted from histograms are taken as input feature vectors to neural networks for training the classifier of BPNN and identifying the corresponding five PD modes. The above research work and the results obtained are fully described in this paper.

The article is organized as follows: Section 1 is a summary of the content related to this research subject. Experimental models of PD, original data of PD and fingerprint of PD are elaborated on in Section 2, and the feature extraction, rationality verification and class separability criterion are all analyzed and demonstrated in Section 3. Section 4 includes the discussion of recognition results and analysis of the effect of hidden layers in BPNN on recognition rate, and Section 5 illustrates some valuable conclusions. The contribution of this paper is to demonstrate the separability of partial discharge signal modes in gas dielectrics caused by the inhomogeneity of electric field intensity by experiments and theory.

2. PD Model and Fingerprint Features

As known by dielectric physics, the primary cause of triggering PD is the uneven electric field distribution in the local region of the insulation system, especially in high-voltage electrical equipment whose gas medium is the main insulation dielectrics, such as GIS (Gas Insulated Switchgear), gas-insulated power transformers and so on, this type of problem is particularly obvious. Excluding the influence of impurities contained in the gas, the gas dielectric itself is especially sensitive to the unevenness of the electric field, and most of the triggers for PD are protrusions existing on electrode surface, metal fine particles freely moving between electrodes, and cracks rest in the supporting insulator [7] [8]. Although

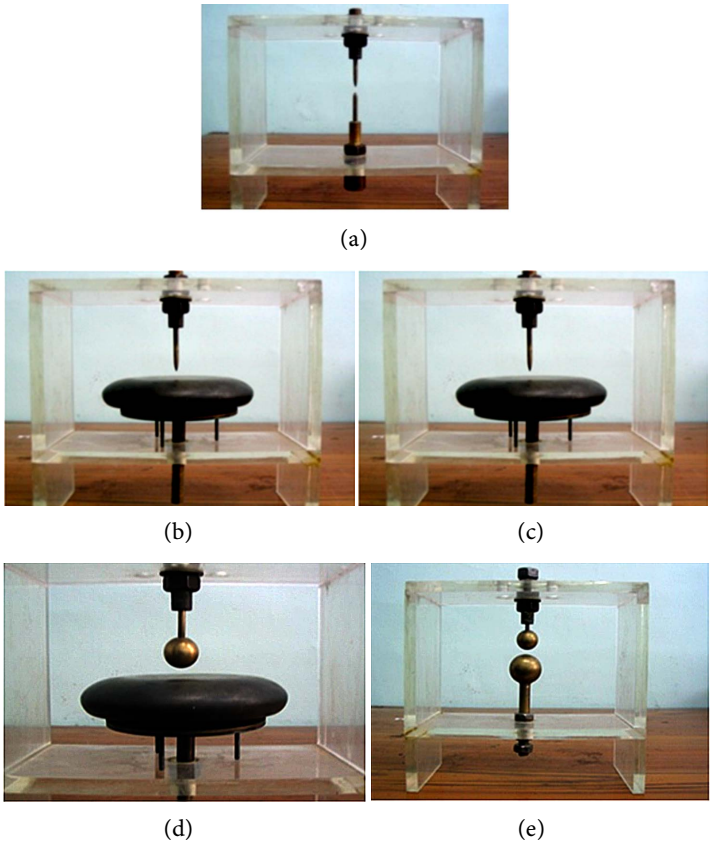
different detection techniques can obtain various manifestations of the PD signals, "fingerprints" of the PD occurred due to defective insulation always having been deposited in the data pool. Hence, it is very practical significance for modeling the PD triggered by the uneven electric field, collecting the signal waveform of PD, extracting fingerprint characteristics with the data mining techniques, establishing a set of vector spaces depicting the heterogeneity of each specific electric field, employing artificial intelligence algorithms to train the classifier ANN and finally realizing the intelligent identification of inhomogeneity of electric field intensity. The research methods and experimental results in this paper have reference significance for the diagnosis, prediction and residual life prediction of partial discharge faults of gas-insulated electrical equipment.

2.1. PD Model and Experimental Procedure

Based on the design principle of SCR series dry-type transformer insulation structure, this paper collects PD data information online during its operation and conducts on-site anatomy of the equipment [9]. Meanwhile, one-to-one correspondence related the geometric configuration of the PD fault location to the corresponding PD information characteristics is analyzed and summarized in details. Therefore, five types of PD models are designed to provide experimental and theoretical basis for establishing standard database of PD fault diagnosis, location, classification and life prediction of SCR series dry-type transformer in service, and the aim is to realize finally the intelligent monitoring and warning for the equipment operating status [10].

For artificially simulating the unevenness of the electric field distribution, five pairs of experimental models are designed and made from the industry standard brass material, they are needle-needle, needle-plate, needle-sphere, sphere-plate and sphere-sphere electrode system respectively, the clearance distances of needle-needle, needle-plate and needle-sphere are all 12 mm, while gaps of the sphere-plate and sphere-sphere equals to 10 mm. The needle electrodes are all copper cylindrical with a diameter of 4 mm, and the tip angle is 30˚. Then the plate electrodes are made from copper circular plate, diameter is of 100 mm, surface polishing and edge chamfering, and two sphere electrodes are all solid copper ball, and their diameters are of 10 mm and 20 mm respectively while the sphere surface polished. These five pairs of PD model are shown as in **Figure 1**.

The discharge chamber is a cylinder made of plexiglass material with a diameter of 450 mm, a height of 600 mm, and a wall thickness of 15 mm, as shown in **Figure 2**. Moreover, the chamber must be cleaned, dried and evacuated before the experiment, and then injected with clean and dry air. The experiment is carried out under the conditions of IEC and GB, and the standard specifies that the atmospheric pressure equals to 0.1013 MPa, temperature is $T = 20$˚C and absolute humidity 11 g/m^3. In the field of electrical insulation technology, the four parameters that describe the insulating properties of gas dielectrics respectively are relative permittivity ε_r, electrical conductivity γ, dielectric loss angle tangent



(a) (b) (c) (d) (e)

Figure 1. Five pairs of the PD model. (a) Needle-needle electrode; (b) Needle-plate electrode; (c) Needle-sphere electrode; (d) Sphere-plate electrode; (e) Sphere-sphere electrode.

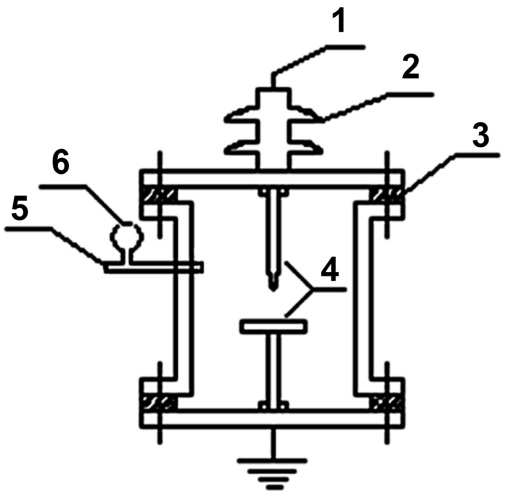


Figure 2. Schematic diagram of plexiglass chamber. 1. High voltage leads; 2. Insulator; 3. Seal ring; 4. Electrode system; 5. Air filler valve; 6. Pressure gauge and thermometer.

$\mathrm{tg}\,\delta$ and breakdown field strength $\vec{E}$. Under IEC and GB conditions, the relative permittivity of air is $\varepsilon_r \approx 1$, and the electrical conductivity is $\gamma \approx 1.8\times10^{-17}$ S/m, the dielectric loss angle tangent under power frequency is $\mathrm{tg}\,\delta = 10^{-9}$, and the electric field strength of air is $\vec{E} = 30$ kV/cm.

The pressure of the gas in the chamber depends on the needs of the experiment, and the temperature is always maintained at 20 degrees Celsius. According to the requirements of the experiment, five pairs of PD electrode are fixed in sequence on the base of the chamber, not only the electrode gap can be adjusted free, but the electrode itself should also be easy to install and remove.

The experiment work is carried out in an electromagnetic shielded laboratory, and the experimental principle is shown in **Figure 3**. The acquisition of the PD signals is performed by the Hipotronics DDX-7000 digital PD detector. In order to meet the needs of signal analysis, processing and PD identification classification, PD signals are quickly acquired, stored and intercepted in integer multiple cycle lengths. In the whole experiment process of this project, 1000 original PD data are collected, that is to say 200 basic data were collected for each PD model, and the length of each data was 100 complete cycles.

Therefore, the original PD data, corresponding to the five electrode structures described herein, are illustrated in **Figure 4**, which are triggered by power frequency AC voltage.

2.2. Basic Quantities of PD

For the registration of basic quantities, the momentary values of the experimental voltage and the discharge signal were processed [11]. It is known that using conventional detection methods and the electrical activity of PD is represented by three independent quantities only: the discharge magnitude q_i, the ignition voltage u_i and the position of the discharge related to the phase angle φ_i of the test voltage. If during one half cycle of the test voltage more discharges occur, on the basis of q_i, u_i and φ_i several quantities can be calculated [12]. In this study, the following basic quantities were measured and processed as shown in **Figure 5**: the inception voltage U_{inc} as a voltage at the sample at which discharge pattern in a half cycle of the test voltage starts, and N_q as the number of discharges for each halve period of the voltage cycle. Therefore, finally PD $\varphi - q - n$ maps generated by the corresponding electrode have been gained shown in **Figure A1** (See **Appendix** to this article).

According to reference [13] and reference [14], it is known that the statistical

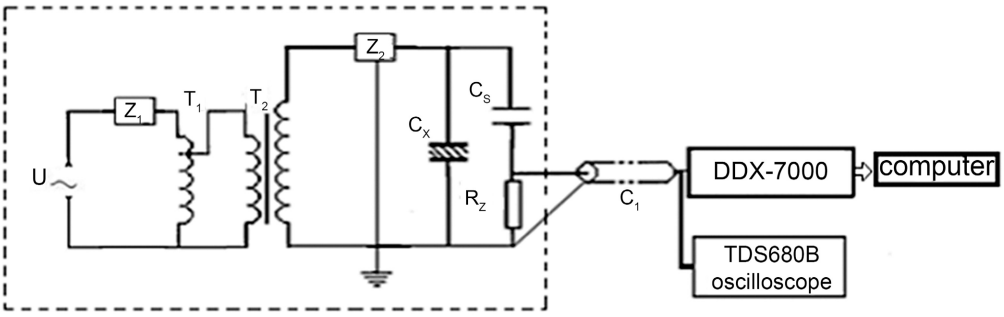


Figure 3. Schematic diagram of PD signal acquisition. Z_1, Z_2—low pass filter; T_1—regulator; T_2—transfmrmer; Cs—standard capacitor; Cx—Sample capacitance in the chamber; C_1—coaxial signal cable; Rz—Detecting impedance.

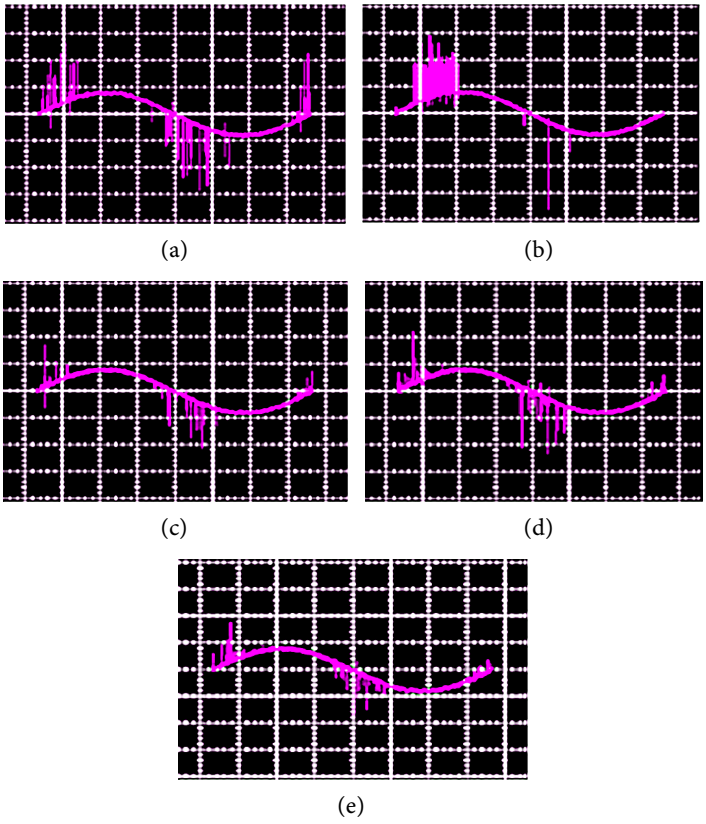


Figure 4. PD waveforms in one period of AC. (a) PD signal of the needle-needle; (b) PD signal of the needle-plate; (c) PD signal of the needle-sphere; (d) PD signal of the sphere-plate; (e) PD signal of the sphere-sphere.

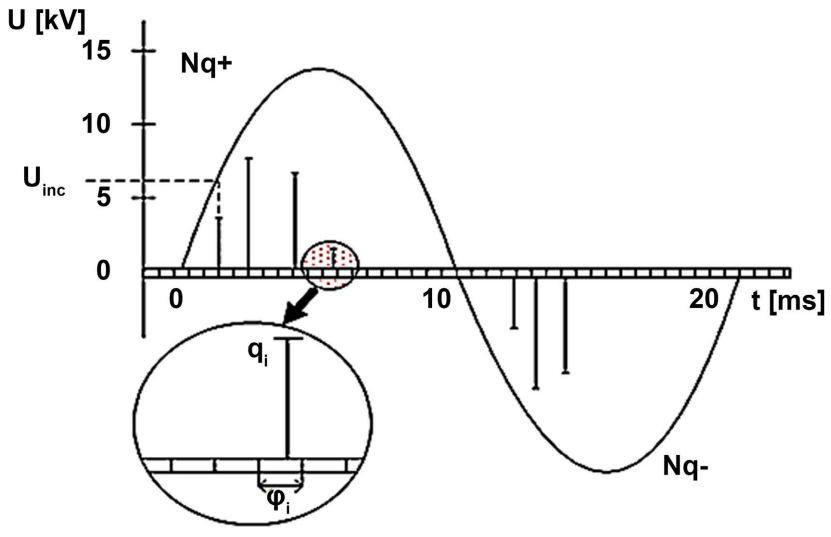


Figure 5. Distribution of discharge of single cycle.

variation in PD occurs, both in magnitude and in the temporal behavior of partial discharges. This variation in the time is partly caused by statistical variations in the discharge phenomenon itself and is also partly the result of the changes in the discharge site. Therefore, the voltage cycle is divided into phase windows

representing the phase angle axis (0˚ to 360˚). If the observation takes place over several voltage cycles, four quantities can be determined in each phase window, *i.e.* the sum of the discharge magnitudes, the number of discharges, the average value of discharges and the maximum value of discharges. In the research process of this paper, the following phase-position quantities are processed just as reference [15]: the impulse count distribution $H_n(\varphi)$, which represents the number of observed discharges in each phase window as a function of the phase angle; the mean impulse height distribution $H_{qn}(\varphi)$, which represents the average amplitude in each phase window as a function of the phase angle. $H_{qn}(\varphi)$ is calculated by dividing the total discharge amount in each phase window by the impulse number of discharges in the same phase window. These quantities, observed throughout the whole angle axis, result in distributions of discharge recurrence as a function of phase angle. Based on previous research, we already know that the discharges occur during a voltage cycle in two sequences, and for each half of the voltage cycle separate discharge patterns can be measured. But in the case of similar inception conditions for each half of the voltage cycle, similar discharge patterns can be expected. Therefore, $H_{qn}(\varphi)$ and $H_n(\varphi)$ are characterized by two distributions, *i.e.* $H_{qn}^{+}(\varphi)$ and $H_n^{+}(\varphi)$ for the positive half of the voltage cycle, and $H_{qn}^{-}(\varphi)$ and $H_n^{-}(\varphi)$ for the negative half of the voltage cycle.

2.3. Fingerprint of PD

As mentioned above, the five different PD signal waveforms (shown in **Figure 4**) over a sinusoidal period are respectively divided into N phase windows, as illustrated in **Figure 5**. For each phase window n (pulse count), q (discharge quantity) and φ (phase angle) are calculated. Usually two distributions are determined: 1) $H_n(\varphi)$ gives the number of discharge impulses in each phase window, 2) $H_{qn}(\varphi)$ takes the average magnitude of the discharge pulses as a function of the phase angle. Therefore, the characteristic quantities of PD can be defined as the reference [16].

Fingerprints obtained in the experiments are shown as a set of the histograms composed of PD statistical characteristic parameters, and the shape of each histogram could present the essential nature of PD signals. Then each feature of a histogram has statistical characteristics for specified PD type. Therefore, these statistical features are used as pattern character vectors of PD classification for our research objectives. In this paper, these statistical features include skewness (S_k), kurtosis (k_u), asymmetry (A), cross correlation (cc) and phase factor (P_f) that are all calculated based on the PD signals. These parameters are defined as follows:

$$S_k = \frac{\sum_{i=1}^{w}(x_i - \mu)^3 \times p_i}{\sigma^3} \tag{1}$$

where, x_i is the discrete value of phase window i, p_i is the probability of fre-

quency of appearance for that value x_i in phase window i, μ is the mean value $\sum x_i \cdot p_i$, and σ is the variance $\sigma^2 = \sum (x_i - \mu)^2 \cdot p_i$. The skewness represents the asymmetry of the distribution. If the distribution is symmetric, $S_k = 0$. If it is asymmetric to the left, $S_k > 0$, otherwise $S_k < 0$.

$$k_u = \frac{\sum_{i=1}^{w} (x_i - \mu)^4 \times p_i}{\sigma^4} - 3 \tag{2}$$

The kurtosis represents the sharpness of the distribution. If the distribution has the same sharpness as a normal distribution, $k_u = 0$. If it is sharper than normal, $k_u > 0$, and $k_u < 0$ means it is flatter. Examples of these two statistical parameters are demonstrated in Figure 6.

The discharge factor A describes the difference in the mean discharge level in the negative and positive distributions. It is defined as following:

$$A = \frac{Q_s^- / n^-}{Q_s^+ / n^+} \tag{3}$$

where, Q_s^+ and Q_s^- are the sum of discharges of the mean pulse height distribution in the positive and the negative half of the voltage cycle; n^+ and n^- are the number of discharges of the mean pulse height distribution in the positive and the negative half of the voltage cycle.

The cross correlation factor is calculated as:

$$cc = \frac{\sum q_i^+ q_j^- - \sum q_i^+ q_j^- / N'}{\sqrt{\left(\sum (q_i^+)^2 - \sum (q_j^+)^2 / N'\right)\left(\sum (q_i^-)^2 - \sum (q_j^-)^2 / N'\right)}} \tag{4}$$

where, q_i^+ is the mean discharge magnitude in window i of the positive half cycle and q_i^- the mean discharge magnitude in the corresponding window in

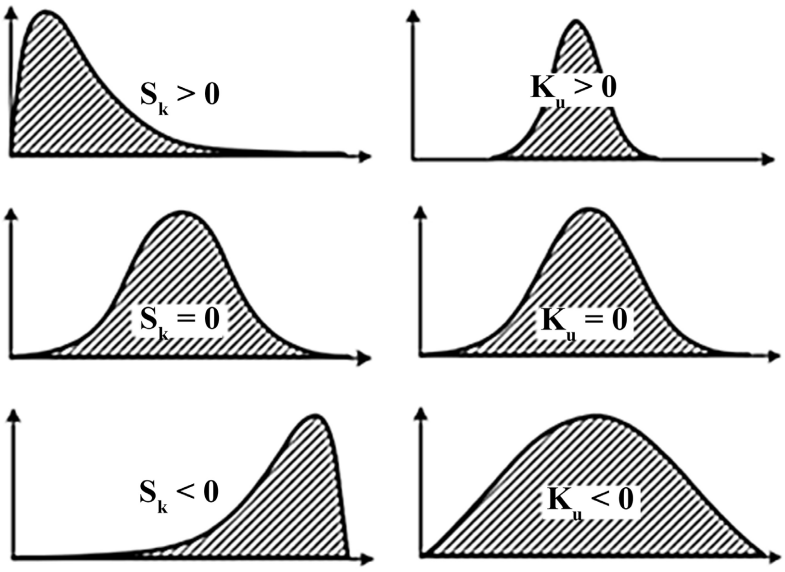


Figure 6. Illustration of different values for skewness S_k and kurtosis k_u of a distribution [4].

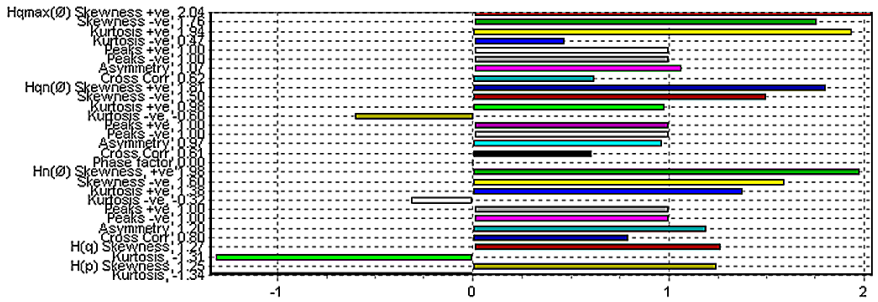
the negative half cycle; N' is the number of phase windows per half cycle. Thus, if the shapes are the same, then $cc=1$; if they differ completely, then $cc=0$.

Phase asymmetry is:

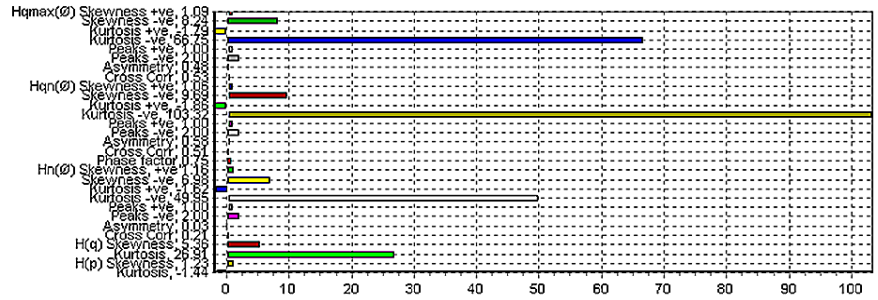
$$P_f = \frac{\varphi_{in}^-}{\varphi_{in}^+} \tag{5}$$

where, φ_{in}^+ and φ_{in}^- are the inception discharge phase of the positive and the negative half of the voltage cycle. The phase asymmetry P_f is used to study the difference in inception voltage of the positive and negative half of the voltage cycle.

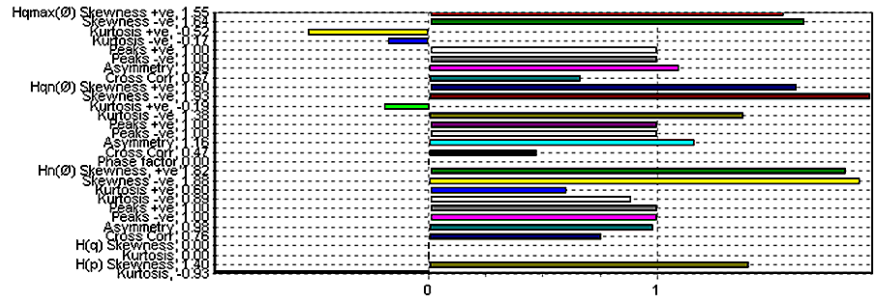
As known from the aforementioned discussion, statistical parameters of the PD pulse sequence are useful to describe the distribution histograms of the several quantities of PD pulses. In general, it seems to be a better choice to use parameters which describe the shape of the distribution, rather than using absolute values which depend on the real data such as the mean value of the discharge magnitudes. **Figure 7** shows 5 types of PD fingerprint acquired respectively, and each fingerprint contains 29 quantitative values, which is a set of unique spatial vectors that correspond to the PD events of the specified electrode configuration under periodic 50 Hz AC voltage. Therefore, these 5 types of PD fingerprint represent their intrinsic characteristics in the nature. In other words, five vector spaces are constructed and each is 29-dimension. Nevertheless, they are not directly utilized to carry out PD pattern classification. The reason is that due to the large number of data vectors, the structure of the classifier may be very complex and hence, the recognition algorithm cannot be completed successfully, then eventually the ideal classification effect may not be achieved.



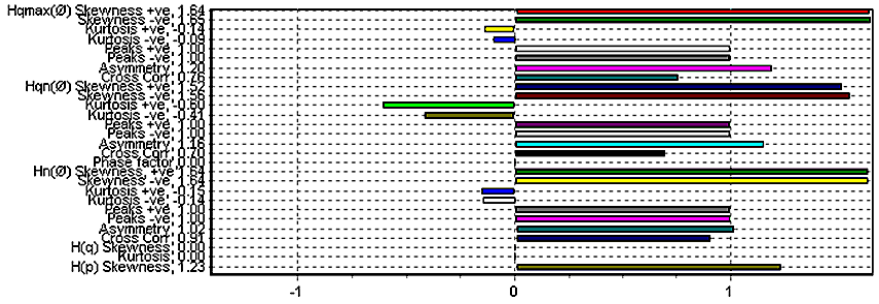
(a)



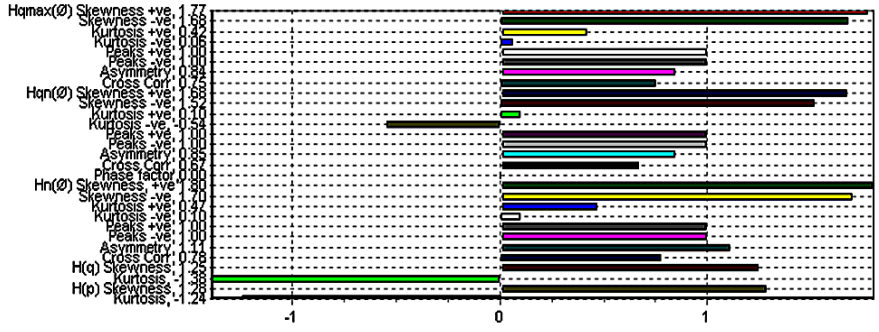
(b)



(c)



(d)



(e)

Figure 7. PD fingerprints of the five models. (a) Fingerprint of the needle-needle PD; (b) Fingerprint of the needle-plate PD; (c) Fingerprint of the needle-sphere PD; (d) Fingerprint of the sphere-plate PD; (e) Fingerprint of the sphere-sphere PD.

As well known, the high-dimensional data space vector can be reduced to a low-dimensional space through related operations. This can not only reduce the data amount, but also convert complex problems into simple and easy-to-handle ones. There have been effective analytical methods applied in different fields of Artificial intelligence recognition for fault detection, diagnosis and prediction. Further, such methods have also been successfully applied in PD pattern recognition and fault classification [17].

3. Feature Extraction of PD Fingerprint

In general, intelligent recognition procedures include signal process, feature extraction, classifier design, classification algorithm and result verification. However, any patterns which can be recognized and classified all have a number of

discriminative features. Thus, the first step in any recognition process is to consider the problem of what discriminatory features to select and how to extract these features from the signal data. It is quite clear that the number of features needed for successful classification depends on the discriminative quality of the chosen features.

3.1. Class Separability Criterion

One of the crucial issues in classification is the curse of dimensionality, hence, a low-dimension feature space is desired for simplifying calculation and reducing costs. Therefore, features should be carefully selected for preserving the maximum class separability. In our study, the criterion based on within and between-class scatter values is adopted as the measure of discrimination at a set of histogram values. The within-class scatter value measures the scatter of feature vectors of different classes around their respective mean values, and the between-class scatter value is defined as the scatter of the conditional mean values around the overall mean value [18] [19]. Therefore, the class separability criteria are expressed as follows:

1) Within-class scatter matrix:

$$S_w = \sum_{i=1}^{M} P_i S_i \tag{6}$$

where, S_i is covariance matrix of w_i, which is calculated as follows:

$$S_i = E\left[\left(x-\mu_i\right)\left(x-\mu_i\right)^{\mathrm{T}}\right] \tag{7}$$

Here, P_i is a prior probability of the class w_i. Namely, $P_i \approx n_i / N$, where n_i is the sample number of w_i in all samples N.

2) Between-class scatter matrix:

$$S_b = \sum_{i=1}^{M} P_i \left(\mu_i - \mu_0\right)\left(\mu_i - \mu_0\right)^{\mathrm{T}} \tag{8}$$

where, μ_0 is a global average vector $\mu_0 = \sum_{i}^{M} P_i \mu_i$.

3) Mixed scatter matrix:

$$S_m = S_w + S_b \tag{9}$$

4) Distance separability criterion:

$$J = tr\left\{S_w^{-1} S_m\right\} \tag{10}$$

All in all, J presents the dispersion of vectors within the class and also the dispersion between the classes, and it is a criterion for judging the structure of vector space within a class and the separability between classes. A large value of J indicates that the space vectors within the class are reasonably constituted. The feature quantity extracted from the fingerprint of the class not only is able to completely characterize its similarity, but also can indicate that the distance between the classes is also large meaning that there is separation [20].

3.2. Feature Selection

It is shown from Figure 7 that 29 characteristic quantities are included in fingerprint for each PD model. As evidenced by well-known theories and practices, too many characteristic quantities for recognition can make the structure of the neural network more complex and the training process more difficult and time-consuming. Therefore, it is a very important task to carefully select different combinations of feature quantities. This task not only requires a certain amount of practical experience and multiple experiments, but also it must be met principle of classification and verified by the class separability criterion. Hence, nine characteristic quantities are picked from the distribution function $H_{qn}(\varphi)$ in each fingerprint of PD models, which are respectively S_k^+, S_k^-, K_u^+, K_u^-, Q_s^+, Q_s^-, A, cc, P_f. At the same time, the nine characteristics can be divided into 5 groups, the first group X_1 contains three quantities (A, cc, P_f), the second group X_2 has four quantities (S_k^+, S_k^-, K_u^+, K_u^-), the third group X_3 has five quantities (S_k^+, S_k^-, K_u^+, K_u^-, A), the fourth group X_4 contains seven quantities (S_k^+, S_k^-, K_u^+, K_u^-, Q_s^+, Q_s^-, A), and last group X_5 Contains all nine characteristic quantities (S_k^+, S_k^-, K_u^+, K_u^-, Q_s^+, Q_s^-, A, cc, P_f).

Based on the distance separability criteria according to Section 3.1, we derive $J_{X_1}=177.3$, $J_{X_2}=41.9$, $J_{X_3}=1243.8$, $J_{X_4}=1905.6$ and $J_{X_5}=1589.9$ respectively. The J values show that 5 groups of feature combinations all satisfy the requirements of classification and have obvious separability, and moreover, X_4 is particularly distinct.

3.3. BPNN Classifier

In this study, Back-Propagation Neural Network (BPNN) is used as classifier. Numerous studies show that BPNN, a rich learning algorithm, achieves good performance in the presence of memory, and is well-preforming in PD pattern recognition as well. The BPNN structure is shown in Figure 8.

The BP algorithm is a form of supervised learning for the feed-forward ANN and it consists of two steps, namely forward and backward learning. Although the back-propagation algorithm can calculate the modified weights of the hidden layer in learning, there are several problems should pay attention to in the process of learning and training of BPNN. For example, the number of neurons each hidden layer, the initial weight settings, the setting of the learning step and inertial coefficient, and the case if the learning is trapped in local minimum, etc. These issues may affect the recognition performance of BPNN, and may even cause BPNN to fail to converge in the training process [21].

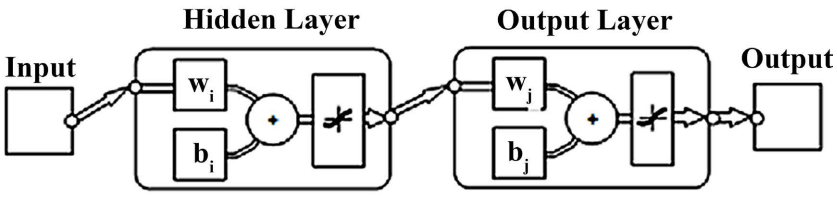


Figure 8. Sketch of BPNN.

4. Result and Discussion

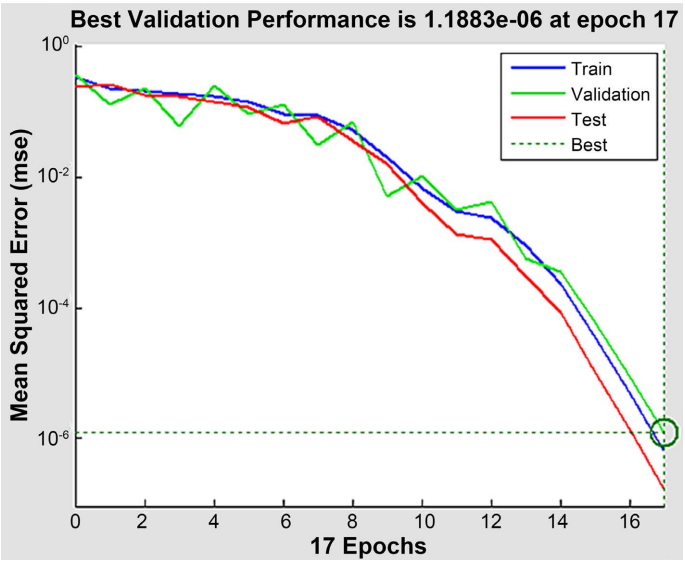
The five sets of characteristic quantities selected are changed into five group input-eigenvectors, and a three-layer neural network is built using the neural network toolbox of the MATLAB. Then, each group input-eigenvector is used to train the specified neural network, and the optimal convergence effect of the neural network is obtained by changing the number of hidden neurons and network parameters.

4.1. Recognition Results

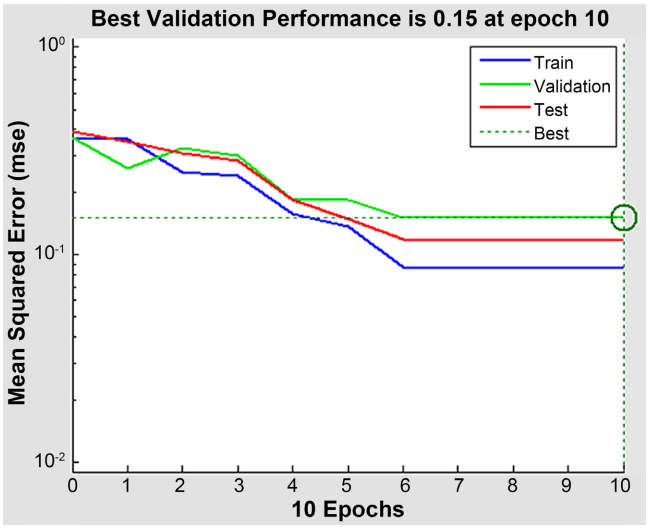
Among all 200 sets of samples for each type of PD fingerprint, 75 sets are as training samples, the other 75 sets take as test samples and the last 50 sets are as verification samples. The training function is "traingdx", and the training error curves of characteristic quantities X_1, X_2, X_3, X_4 and X_5 are shown in Figure 9. Further, the number of hidden layer neurons is 6, 8, 10, 13 and 19 respectively. The recognition results are presented in Table 1.

Table 1. Recognition results.

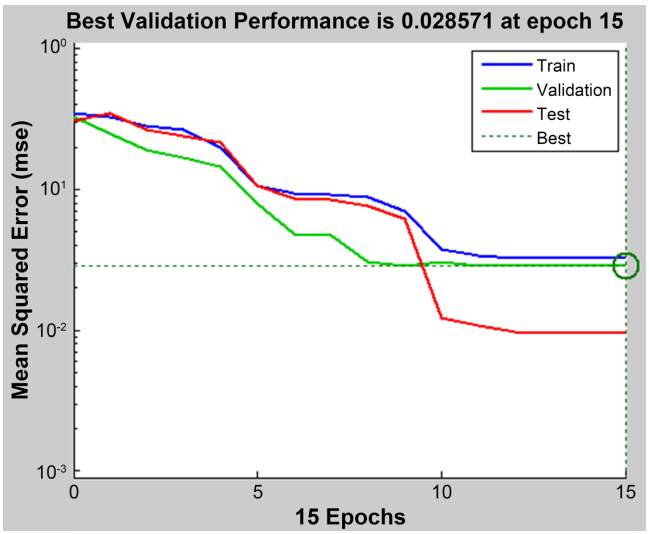
Characteristics / Recognition rate / Test model	X_1	X_2	X_3	X_4	X_5
Needle-needle	100%	95%	100%	100%	100%
Needle-plate	95%	100%	100%	100%	95%
Needle-sphere	100%	100%	100%	100%	100%
Sphere-plate	100%	100%	0%	100%	100%
Sphere-sphere	100%	100%	100%	100%	100%
Total	99%	99%	80%	100%	99%



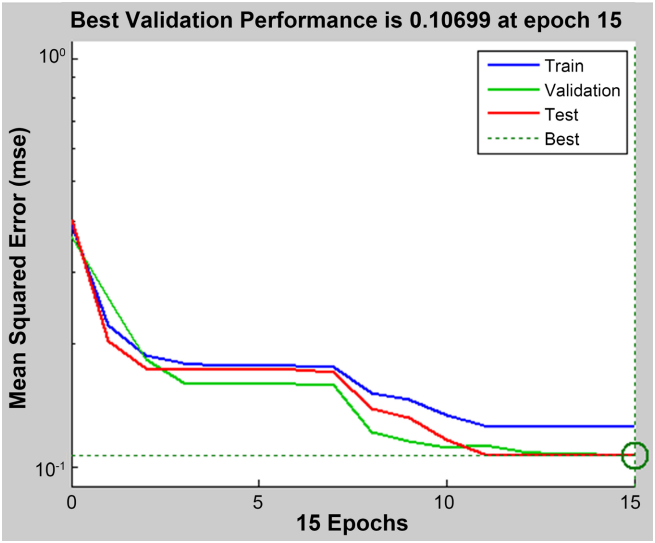
(a)


Best Validation Performance is 0.15 at epoch 10
Train
Validation
Test
Best
Mean Squared Error (mse)
10 Epochs

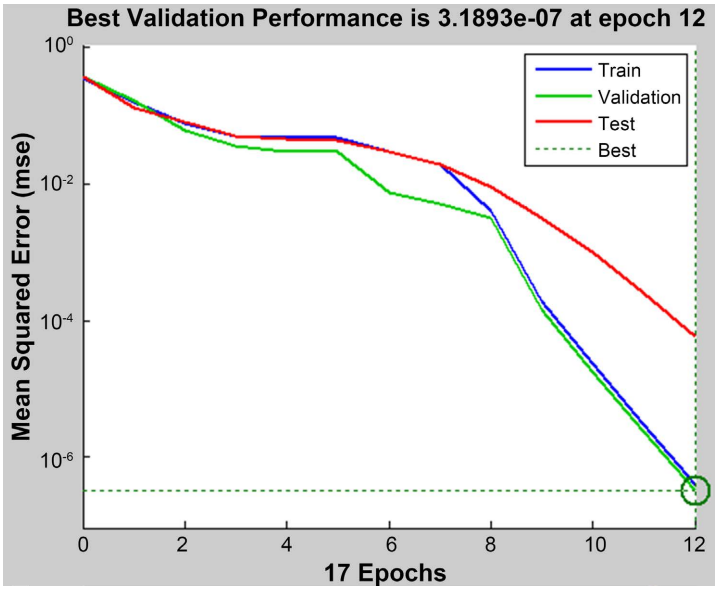
(b)


Best Validation Performance is 0.028571 at epoch 15
Train
Validation
Test
Best
Mean Squared Error (mse)
15 Epochs

(c)


Best Validation Performance is 0.10699 at epoch 15
Train
Validation
Test
Best
Mean Squared Error (mse)
15 Epochs

(d)



(e)

Figure 9. BPNN training curves. (a) Training curve of X_1; (b) Training curve of X_2; (c) Training curve of X_3; (d) Training curve of X_4; (e) Training curve of X_5.

It should be emphasized here that the five groups of curves in **Figure 9** have implied the influence of neurons on the recognition results of the neural network. For X_1, X_2, X_3, X_4 and X_5, using the network structure of **Figure 8**, their input the number of feature quantities is 3, 4, 5, 7, and 9, respectively. Because the number of neurons in the input layer of each neural network should match it, the number of neurons in the input layer should be 3, 4, 5, 7, and 9, respectively. On this basis, the corresponding optimized neural network structure was obtained through repeated experiments. The corresponding number of neurons in the hidden layer of the neural network were respectively 6, 8, 10, 13 and 19, and the output layer of the neural network must be the number of neurons is 5, because this paper is to identify five PD patterns, and according to the machine learning criterion described in the literature [22], the convergence criterion of the neural network learning and training in this paper is given.

4.2. Effect of Hidden Layer Structure

Table 1 shows that selected fingerprint can be a useful and informative characteristic to identify PD modes under the condition that the hidden layer neurons are 19 except x_3, but this does not mean that the more feature quantities, the higher the recognition result. Actually, the identification rate of characteristic quantity X_4 is the highest out of 5 sets of input feature vectors, because the vector x_4 best satisfies the class separability criterion.

4.3. Effect of Hidden Layer Structure

How to select the number of hidden layer neurons is a very complex issue, which often needs the designer's experience and numerous experiments. There are

Table 2. Recognition results of X_5.

Recognition neuron rate / Hidden layer / Test model	1	2	3	4	5	6	10	15	20
Needle-needle	0	100%	100%	100%	100%	100%	100%	100%	100%
Needle-plate	0	100%	100%	100%	100%	100%	100%	100%	100%
Needle-sphere	0	100%	100%	100%	0%	100%	100%	100%	100%
Sphere-plate	0	0%	100%	100%	100%	100%	100%	100%	100%
Sphere-sphere	0	100%	0%	0%	100%	100%	100%	100%	100%
Total	0	80%	80%	80%	80%	100%	100%	100%	100%

currently no ideal analytic formulas and rules to follow. In this paper, the number of neurons in the hidden layer is chosen according to Kolmogorov's theorem [23]. Table 2 shows that the hidden layer structure of BPNN has an obvious effect on the recognition results, especially for a certain type of electrode model. When the number of hidden layer neurons is 6 or above, the neural network can obtain a recognition rate of 100%. However, excessive hidden layer neurons in artificial neural networks may have the following problems:

1) The structure of artificial neural networks is too complex, which not only requires more time to train, but also increases the cost.

2) Because the network structure is too complex, its convergence becomes slow or even fails to converge.

3) Due to the complex network structure, its robustness and fault tolerance rate are greatly reduced.

To sum up, the artificial neural network structure optimization is also a technical problem that cannot be ignored besides the correct acquisition of the original signal of the PD, the determination of the PD mode class and the selection of a reasonable PD feature vector. As shown in Table 2, X_5 is a reasonable set of the PD feature vector in the midst of the PD fingerprint, and when there are 6 neurons in the hidden layer, the recognition rate of the network has reached 100%, which means that the network structure and performance at this time are the best, at the same time, having the best classification effect is a matter of course.

5. Conclusions

The research in this paper deliberately simulates five kinds of electric field intensity distributions. Through the experiments, the original PD signals are collected, and five kinds of PD fingerprint patterns are established. According to the class separability criterion, the input feature vectors of the ANN are carefully selected from the PD fingerprint, they are X_1, X_2, X_3, X_4 and X_5. The J values show that 5 groups of feature combinations satisfy the requirements of classification and have obvious separability. Furthermore, the effect of input feature vectors, the

number of neurons in the hidden layer and the recognition rate of ANN are discussed experimentally. The conclusions obtained are as follows:

1) The PD models structured five kinds of electrode systems truly simulate different electric field distributions, and these PD signals have obviously statistical behaviors. Further, the extracted fingerprint features not only can correctly characterize the intrinsic difference of gas-insulation PD, but also have distinct class separability. These findings are supported by the experimental results shown in Figure 4 and Figure 7.

2) Although the feature selection and extraction from PD fingerprint is a complex process, it can be implemented using effective algorithms. Therefore, a reasonable set of PD characteristics can not only avoid the curse of dimensionality, but also it can simplify the design of the classifier and improve the effect of pattern classification. This conclusion has reference value and guiding significance for further research on PD pattern classification technology, as shown in Table 1.

3) After the PD pattern characteristics and pattern classes are determined, the choice of the hidden layer structure is a key issue, and even it is decisive. As shown in Table 2, 6 neurons in the hidden layer are the best structure of the network. If less than 6 neurons, a low recognition rate is caused. However, if there are more than 6 neurons, the recognition rate can be guaranteed, but the network structure becomes complicated and the learning process may be time-consuming, and can be not even converged.

Although intelligent PD pattern recognition is a complicated task, with the advancement of computing technology and the improvement of artificial intelligence, PD pattern recognition technology is also developing rapidly towards the intelligent direction. However, there are still many problems to be further explored to meet the needs of industrial development.

Funding

This work was supported by National Natural Science Foundation of China (Grant No. 51077032).

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Yazici, B. (2004) Statistical Pattern Analysis of Partial Discharge Measurements for Quality Assessment of Insulation Systems in High-Voltage Electrical Machinery. *IEEE Transactions on Industry Applications*, **40**, 1579-1594. https://doi.org/10.1109/TIA.2004.836174

[2] Fwenando, F.G., Jacinto, T.J., Ruben, J.V., *et al.* (2017) Analysis of Insulating Material of XLPE Cables Considering Innovative Patterns of Partial Discharges. *Mathematical Problems in Engineering*, **2017**, Article ID: 23794118.

https://doi.org/10.1155/2017/2379418

[3] Zheng, D.C. (2019) Dielectric Characteristics and Application of the SF_6(M). Science Press, Beijing.

[4] Firuzi, K., Vakilian, M. and Phung, B.T. (2019) Partial Discharges Pattern Recognition of Transformer Defect Model by LBP & HOG Features. *IEEE Transactions on Power Delivery*, **34**, 542-550. https://doi.org/10.1109/TPWRD.2018.2872820

[5] Gao, K. Tan, K.X., Li, F.Q., *et al.* (2002) PD Pattern Recognition for Stator Bar Models with Six Kinds of Characteristic Vectors Using BP Network. *IEEE Transactions on Dielectrics and Electrical Insulation*, **9**, 381-389. https://doi.org/10.1109/TDEI.2002.1007700

[6] Wong, J.K.R., Hazlee, A., Ab, H.A., *et al.* (2015) Partial Discharge Classifications: Review of Recent Progress. *Measurement*, **68**, 164-181. https://doi.org/10.1016/j.measurement.2015.02.032

[7] Si, W.R., Li, J.H., Yuan, P., *et al.* (2008) Digital Detection, Grouping and Classification of Partial Discharge Signals at DC Voltage. *IEEE Transactions on Dielectrics and Electrical Insulation*, **15**, 1663-1674. https://doi.org/10.1109/TDEI.2008.4712671

[8] Callender, G. (2018) Modelling Partial Discharge in Gaseous Voids. University of Southampton, Southampton.

[9] Lu, C.B. (2002) Dry-Type Power Transformer Theory and Calculation. Liaoning Science and Technology Press (China), Shenyang.

[10] Yan, Z. and Zhu, D.H. (2007) High Voltage Insulation Technology. China Electric Power Press (China), Beijing.

[11] Krivda, A. (1995) Automated Recognition of Partial Discharge. *IEEE Transactions on Dielectrics and Electrical Insulation*, **2**, 796-821. https://doi.org/10.1109/94.469976

[12] Pearson, J.S., Farish, O., Hampton, B.F., *et al.* (2002) Partial Discharge Diagnostics for Gas Insulated Substations. *IEEE Transactions on Dielectrics and Electrical Insulation*, **5**, 893-905. https://doi.org/10.1109/94.469984

[13] Zheng, D.C. (2005) PD Pattern Recognition Based on BP Network. Ph.D. Dissertation, Harbin University of Science and Technology, Harbin.

[14] Mas'ud, A.A., Stewart, B.G. and McMeekin, S.G. (2014) Application of an Ensemble Neural Network for Classifying Partial Discharge Patterns. *Electric Power Systems Research*, **110**, 154-162. https://doi.org/10.1016/j.epsr.2014.01.010

[15] Kreuger, F.H., Gulski, E. and Krivda, A. (1993) Classification of Partial Discharges. *IEEE Transactions on Electrical Insulation*, **28**, 917-931. https://doi.org/10.1109/14.249365

[16] Thummapal, D., Jain, A. and Kushare, B.E. (2014) Partial Discharge Pattern Recognition of HVGIS by Using Artificial Neural Networks. *International Journal of Engineering Research & Technology*, **3**, 449-454.

[17] Sahoo, N.C. and Salama, M.M.A. (2005) Trends in Partial Discharge Pattern Classification: A Survey. *IEEE Transactions on Dielectrics and Electrical Insulation*, **12**, 248-264. https://doi.org/10.1109/TDEI.2005.1430395

[18] Heidari, M. (2016) Combined Diagnosis of PD Based on the Multidimensional Parameters. *Modelling and Simulation in Engineering*, **2016**, Article ID: 5949140. https://doi.org/10.1155/2016/5949140

[19] Contin, A., Cavallini, A., Montanari, G.C., *et al.* (2000) Artificial Intelligence Methodology for Separation and Classification of Partial Discharge Signals. 2000 *An-*

nual Report Conference on Electrical Insulation and Dielectric Phenomena, Victoria, 15-18 October 2000, 522-526.

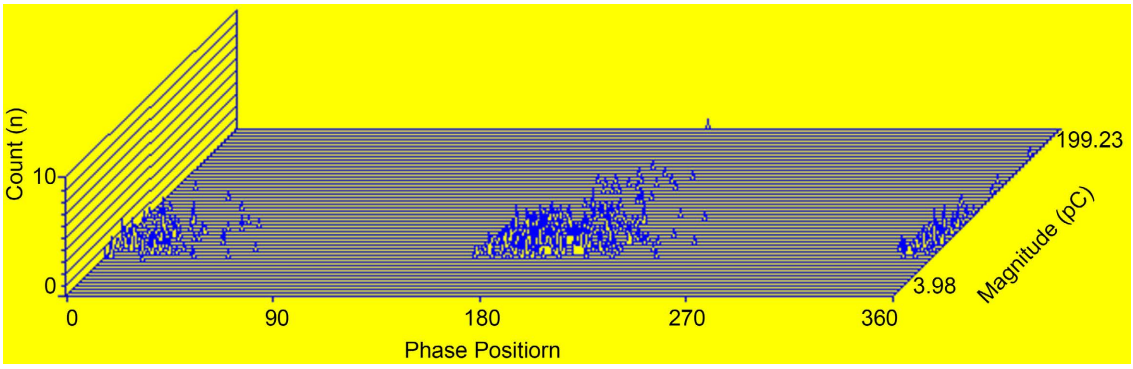
[20] Nguyen, C.H. and Ho, T.B. (2008) Pattern Recognition. *Pattern Recognition*, **41**, 366-372. https://doi.org/10.1016/j.patcog.2008.04.005

[21] Ho, T.K. and Basu, M. (2002) Complexity Measures of Supervised Classification Problems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **24**, 289-300. https://doi.org/10.1109/34.990132

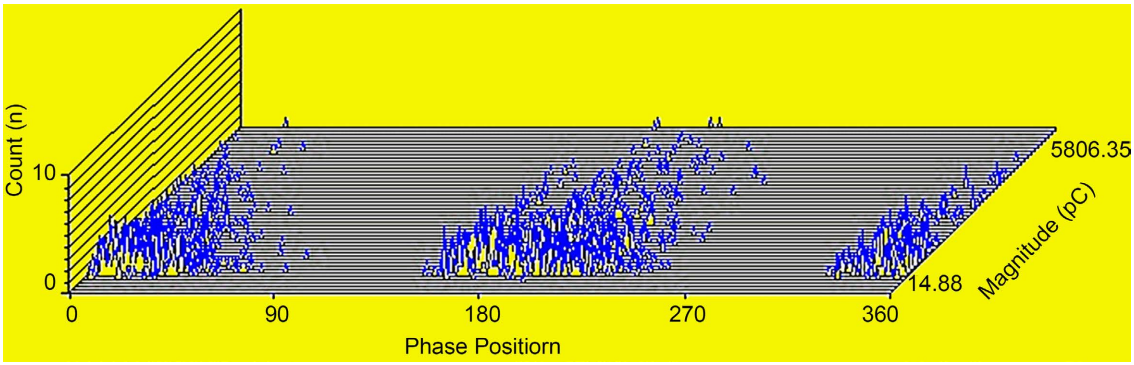
[22] Srivastava, N., Hinton, G.E., Krizhevsky, A., *et al.* (2014) Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*, **15**, 1929-1958.

[23] Sergios, T. and Konstantinos, K. (2004) Pattern Recognition. 2nd Edition, Publishing House of Electronics Industry (China), Beijing, 116-117.

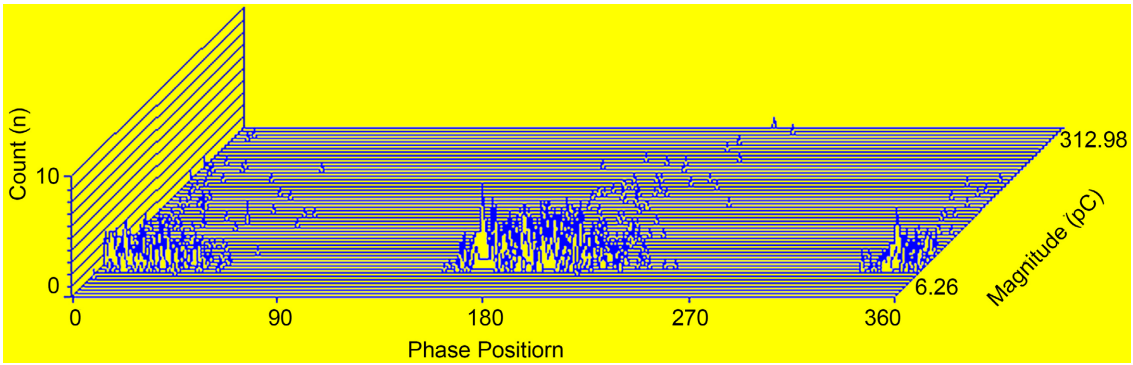
Appendix



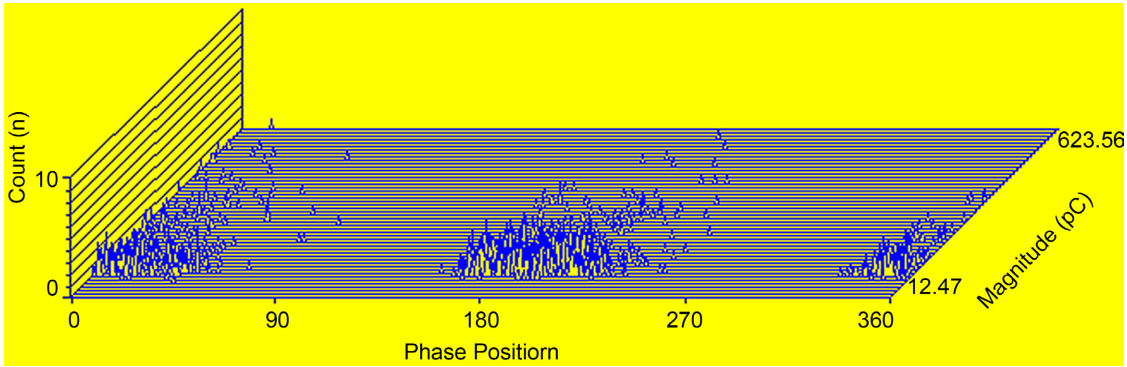
(a)



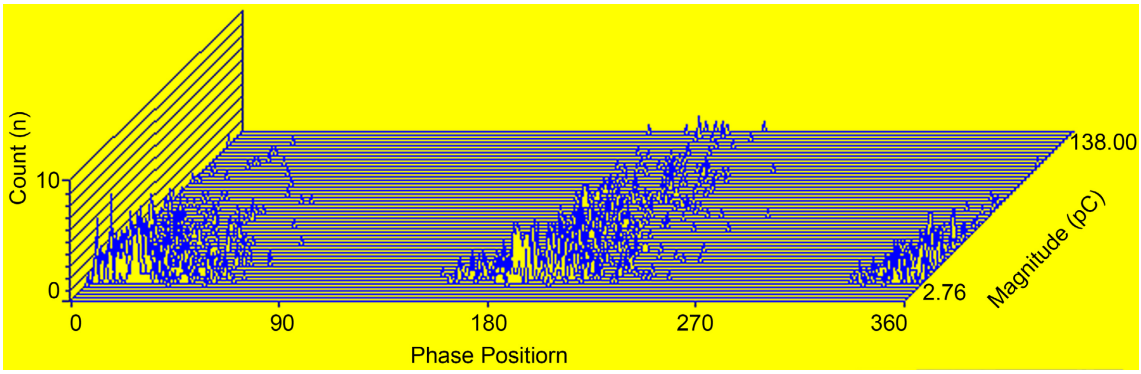
(b)



(c)



(d)



(e)

Figure A1. PD $\varphi - q - n$ maps of five kinds of electrodes. (a) $\varphi - q - n$ map of the needle-needle PD; (b) $\varphi - q - n$ map of the needle-plate PD; (c) $\varphi - q - n$ map of the needle-sphere PD; (d) $\varphi - q - n$ map of the sphere-plate PD; (e) $\varphi - q - n$ map of the sphere-sphere PD.