

Research on Optimization of Feature Extracting Based on PD Fingerprints in Pattern Recognition

Wei LI¹, Lei SHI¹, Hong-Jing LIU², Huan XIE², Yuan GUI²,
Dian-Chun ZHENG^{3*} and Pei-jing HOU³

¹ State Grid Beijing Electric Power Research Institute, Beijing, China, 100075

² Beijing Electric Power Research Institute, Beijing, China, 100075

³ Harbin University of Science and Technology, Key Laboratory of Engineering Dielectrics and Its Application, Ministry of Education, Harbin, China, 150080

*Corresponding author: zhengdianchun@126.com

Keywords: PD, Fingerprints, Pattern recognition, BP neural network.

Abstract. In order to investigate the effect of feature extraction on pattern classification for partial discharge (PD) signals appearing potential insulating failures in high voltage apparatus while operation, the PD-fingerprints acquired by Hipotronics DDX-7000 digital PD detector are taken as database and method on the feature extracting from the database is carried out, and its practicability is demonstrated by the mathematics and experiment, respectively. The BP neural network is made of three layers and the transfer function of hidden layer and output layer are tansig, then the influence of the structure of neural network on recognition results is studied at the same time. As a result, the optimal characteristic vector with obvious separability and the number of hidden layer are obtained, and achievements of research show that the network convergence is not only quickly, but also the recognition rate very high so much as up to 100%.

Introduction

Partial discharge is caused by defects in insulation equipment, which is one of the important parameter to evaluate the insulation properties of online high-voltage electrical equipment. Insulation defects are closely related to partial discharge. Insulation defects, discharge development and residual life of high voltage electrical equipment can be predicted by monitoring the status of high voltage apparatus on-line, gathering, analyzing and classifying partial discharge signals at real time. Therefore, the partial discharge detection plays an important role in preventing accidents of high voltage equipment. Partial discharge pattern recognition is essential to diagnose and locate the defects in electrical devices as well, which can reduce the maintenance cost of high voltage electrical equipment.

The extraction of characteristic quantity has great influence on the result of partial discharge pattern recognition. Experiment and theoretical analysis have proved that partial discharge signal which can characterize the insulation fault has statistical properties. The use of statistical methods to extract partial discharge characteristics is beneficial to classifier design, algorithm implementation and pattern classification[1]. PD-fingerprints can reflect the statistical features intuitively, widely used in pattern classification and have a ideal classification consequence of defects [2-3].

PD Model and Fingerprint Characteristics

Experimental Models and Circuits

Partial discharge depends on the applied electric field type and electric field distribution in the insulation structure. In this paper, the distribution of electric field is emphasised in the experiments. Therefore, five types of experimental models simulating PD in air are designed according to the in homogeneity of electric field in this paper, which are needle-needle, needle-plate, needle-sphere, sphere-plate and sphere-sphere electrode systems. In which the gaps of needle-needle, needle-plate, needle-sphere are 12mm, while sphere-plate and sphere-sphere are 15mm. The needle electrodes are all copper cylindrical with 4mm diameter and the tip angle is 30° . The plate electrodes are made of copper with 100mm diameter and surface polishing, edge chamfering. The sphere electrodes are all solid copper ball with 20mm diameter and surface polishing. The needle-plate electrode is shown in Fig. 1.



Fig.1 Electrode of needle-plane

Partial discharge in experiments were measured with pulse current method, and the PD-signals were collected by Hipotronics DDX-7000 digital partial discharge detector.

PD Signal and Fingerprints

The experimental procedure was carried out under 50Hz ac voltage, the acquired partial discharge waveform of sphere-plate is shown in Fig.2.

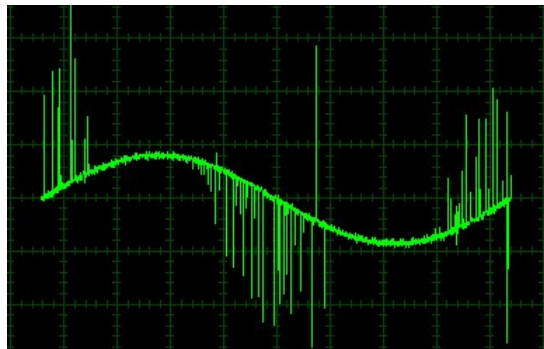


Fig.2 The measured sphere-plate discharge waveform

Fingerprints obtained in experiments are histograms which composed of PD statistical characteristic parameters. The shape of histogram can reflect the nature of partial discharge signals. The features of a histogram are statistical characteristics, where the histogram is used as a model of the probability distribution of a pattern. These statistical features provide us with the characteristics of a PD pattern. In this work, these statistical parameters including skewness(S_k), kurtosis(k_u), cross correlation(cc), peaks(P), asymmetry(A), and phase factor(Pf) are all calculated based on the PD

signals[4]. 250 sets of data were collected from PD-signals in experiments, and 50 sets for each type of discharge.

Feature Extraction

Rule of Distance Separability

All types of samples can be separated because they are located in different regions of the feature space, obviously, the greater the distance the greater the class separability between these regions. The purpose of separability measure is to choose characteristics large between classes distance and small within class variance in the feature vector space. That is to say, a far distance between characteristics of different classes, while a dense aggregation of characteristics in the same class[5]. Criterion as follows:

Within class scatter matrix :

$$S_w = \sum_{i=1}^M P_i S_i \quad (1)$$

Where S_i is covariance matrix of w_i .

$$S_i = E[(x - \mu_i)(x - \mu_i)^T] \quad (2)$$

P_i is the prior probability of class w_i , namely, $P_i \approx n_i / N$. Where n_i is the sample number of w_i in all samples N .

Between class scatter matrix

$$S_b = \sum_{i=1}^M P_i (\mu_i - \mu_0)(\mu_i - \mu_0)^T \quad (3)$$

Where μ_0 is a global average vector, $\mu_0 = \sum_{i=1}^M P_i \mu_i$.

Mixed scatter matrix: $S_m = S_w + S_b$. Generally, J is used to represent the distribution of samples, $J = \text{tr}\{S_w^{-1} S_m\}$. With the numerical of J increase, the separability of characteristic vector becomes more obvious.

Feature Extraction

As can be seen from Fig.4, there are 29 characteristic quantities in each fingerprints. As we all know, excessive amount of the characteristic quantities for recognition will make the structure of the neural network and the training process become more complex. Therefore, several groups of characteristic quantities are selected for recognition, and the rationality is analyzed.

The 9 characteristic quantities can be obtained from the distribution function $H_{qn}(\varphi)$ in fingerprints, which are S_k^+ , S_k^- , K_u^+ , K_u^- , P^+ , P^- , A , cc , Pf . $H_{qn}(\varphi)$ is the average magnitude of the discharge impulses as a function of the phase angle. The nine characteristics can be divided into 5 groups respectively, defined as: X1(A , cc), X2(A , cc , Pf), X3(S_k^+ , S_k^- , K_u^+ , K_u^-), X4(S_k^+ , S_k^- , K_u^+ , K_u^- , A), X5(S_k^+ , S_k^- , K_u^+ , K_u^- , P^+ , P^- , A , cc , Pf). By distance separability rule, then we get $J_{X1}=177.3$, $J_{X2}=41.9$, $J_{X3}=1243.8$, $J_{X4}=1905.6$, and $J_{X5}=1589.9$. The synthesis of J values (generally when $J > 10$, the characteristic quantity can be separated in the

classification problem[6]) show that 5 groups of feature combinations can satisfy the requirements of classification and have obvious separability, especially X4.

NN Classifier

In this article BP neural network is used as classifier. For numerous studies show that BPNN has good performance in the presence of memory, exists a rich learning algorithm, has good effect in partial discharge pattern recognition as well[7]. BPNN is a multilayer feedforward network propagates in one-way.

As can be seen from the figure, The BPNN is a hierarchical feed-forward artificial neural network consisting of three or more fully interconnected layers of neurons.

The basic idea of BP neural network algorithm is sending back the error signal through connect channel, and modify the neuron weight of every layer to make smallest error. Usually the gradient descent algorithm is applied in BP neural network algorithm, therefore, the output function must be differentiable and the Sigmoid function as the output function is normally used.

Though the back propagation algorithm can solve the weights modified problem of the hidden layer, there are several problems should pay attention to in the process of learning and training. For example, the number of hidden layer, the initial weight settings, the setting of the learning step and inertial coefficient , and trapped in local minimum , etc. These issues will affect the recognition performance of BP network, moreover , the network may not converge in learning and training process. This article will focus on the impact of the hidden layer on neural network recognition results.

In practice, the number of hidden layer has a great influence on the solution of the problem. Though the Kolmogorov theorem points out that any continuous function can be achieved by a three layer BP neural network accurately, the Kolmogorov theorem only puts forward the three layer BP neural network can simulate any continuous function, but not necessarily optimal. Therefore, how to choose the number of neural network hidden layer neurons, determine the structure and size of the artificial neural network are still of great significance.

Recognition Results and Error Analysis

The five sets characteristic quantities selected above are recognized by a three layer BP neural network in MATLAB neural network toolbox. The input data for the recognition system is the PD-fingerprints, the hidden layer with 20 neurons and the output layer depends on the defect types to be identified, which is 5 in this work. The transfer function of hidden layer and output layer are tansig:

$$f(n) = \frac{e^n - e^{-n}}{e^n + e^{-n}} \quad (4)$$

30 sets of PD-fingerprints of each kind of discharge are used to train the recognition system, and the rest 20 sets as testing data. The training function is 'traingdx', and the training error curves of characteristic quantities X3 and X4 are shown in Fig.3. While the recognition results are presented in table I.

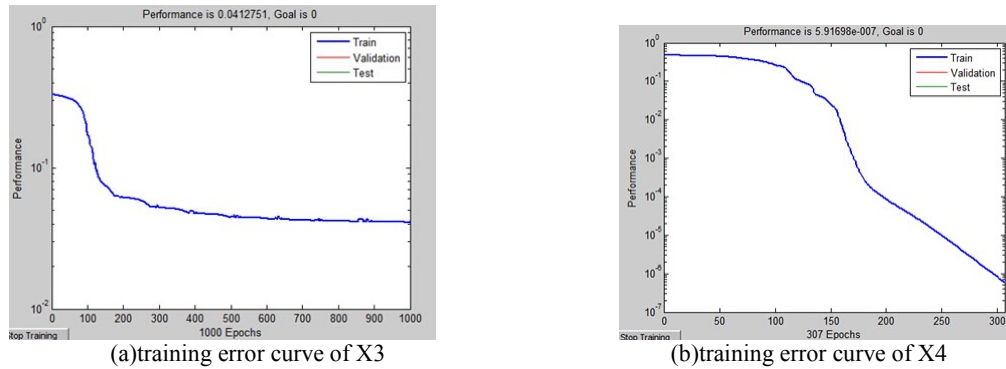


Fig.3 The training error curve

As can be seen from the figure (a), the convergence of the X3 is very slow in the training process, and the performance cannot reach to the goal level after 1000 epochs training. Therefore, the final recognition rate is relatively low. While the network convergence quickly and has a high training accuracy in finger (b), a satisfactory recognition result can be reached as well.

TABLE I Recognition results

Recognition rate Test model	Character - istics	X1	X2	X3	X4	X5
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%

TABLE I shows that fingerprints can be a good characteristic to identify discharge types, but it does not mean the more characteristic quantities the more higher recognition results. Actually, identification rate of characteristic quantity X4 is the highest, which correspond with the distance separability rule.

Hidden Layer Effect on Recognition Results

The number of hidden layer neurons selecting is a very complex issue, often need to be determined by the designer's experience and numerous experiments. Thus, there does not exist an ideal analytical formulas. At present, the hidden layer neurons can be selected by the Kolmogorov theorem roughly[8].

To the obtained feature vector X4 which has the optimal classification properties, the best number of hidden layer neurons has been obtained according to the experiment. Recognition rates by experiments are shown in TABLE II.

TABLE II Recognition results of X4

hidden layer recognition rate 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%

TABLE II shows that the hidden layer structure has obvious effect on 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 the recognition rate of 100%. However, too many hidden layer neurons can lead to a long learning time and the error

is not certain to be a minimum. Therefore, the optimal structure of the BPNN is obtained, and the corresponding hidden layer neurons are 6.

Conclusions

In this paper, the partial discharge fingerprints which have statistical property were extracted from the partial discharge waveforms in five kinds of electrode model, and for simplifying the structure of neural network and the recognition process, 5 sets of characteristic quantity were selected from the experimental data, and the rationality of feature extraction were verified by the distance separability rule, and they were simultaneously used to train the BP neural network selected. The recognition results show that the PD in 5 kinds of electrode model can be identified well with fingerprints, and the highest recognition rate can reach up to 100%. Besides, different combination of the characteristics have a significant impact on the identification results instead of the more characteristic quantities the better recognition results. In addition, the facts are demonstrated that the characteristic vectors whose separability are obvious could get better recognition results, however, BPNN structure has great effect on the recognition rate, so the optimal structure of hidden layers is achieved, and the recognition rate of five types of discharge can reach high enough with the optimal BPNN.

References

- [1] Yazici, Birsen, Statistical pattern analysis of partial discharge measurements for quality assessment of insulation systems in high-voltage electrical machinery[J], IEEE Transactions on Industry Applications, Nov./Dec. 2004, Vol. 40, No. 6 1579-1594
- [2] Gulski, E., Discharge pattern recognition in high voltage equipment, Science, Measurement and Technology, IEE Proceedings, Jan 1995, Vol.142, No.1, 51-61,
- [3] Kreuger, F.H.; Gulski, E.; Krivda, A., Classification of partial discharges, Electrical Insulation[J], IEEE Transactions on. Dec 1993, Vol.28, No.6, 917-931,
- [4] Gulski, E.; Kreuger, F.H., "Computer-aided recognition of discharge sources," Electrical Insulation, IEEE Transactions on., Feb 1992, Vol.27, No.1, 82-92
- [5] ZhaoQi Bian, XueGong Zhang. Pattern Recognition[M]. second edition. Tsinghua University Press, 2000
- [6] Sergios Theodoridis, Konstantinos Koutroumbas. Pattern Recognition[M]. Academic Press, 2008
- [7] Kai Gao; Kexiong Tan; Fuqi Li; Chengqi Wu, PD pattern recognition for stator bar models with six kinds of characteristic vectors using BP network, Dielectrics and Electrical Insulation, IEEE Transactions on., Jun 2002, Vol.9, No.3, 381-389
- [8] ChaoChun Yin, Hao Yin. Artificial Intelligence and Expert System[M]Beijing, China WaterPower Press, 2002