Ruijin LIAO, Ke WANG, Lijun YANG, Lian DUAN, Jian LI

State Key Laboratory of Power Transmission Equipment & System Security and New Technology, Chongqing University

A New Hybrid Feature Extraction Method for Partial Discharge Signals Classification

Abstract. *In this paper, a new hybrid feature extraction method combining adaptive optimal radially Gaussian kernel (AORGK) time-frequency representation with two dimensional nonnegative matrix factorization (2DNMF) is proposed for partial discharge (PD) classification. Firstly, AORGK is applied to obtain the time-frequency matrices of PD ultra-high-frequency (UHF) signals. Then 2DNMF is employed to compress the AORGK amplitude (AORGKA) matrices to extract various feature vectors with different (d_1, d_2) combinations, i.e. (5, 5), (5, 10), (10, 5) and (10, 10). Finally, the extracted features are classified by fuzzy k nearest neighbor (FkNN) classifier and back propagation neural network (BPNN). 600 samples sampled from four typical artificial defect models in Laboratory are adopting for testing of the proposed feature extraction algorithm. It is shown that the successful rate by FkNN and BPNN are all higher than 80%, and FkNN has superior classification accuracies than BPNN under four circumstances of (d_1, d_2) combinations. In addition, FkNN achieves the highest classification accuracy 93.73% with (10, 5) combination. The results demonstrate that it is feasible to apply the proposed algorithm to PD signal classification.*

Streszczenie: *W artykule przedstawiono nową hybrydową metodę klasyfikacji wyładowań niezupełnych (ang. Partial Discharge), wykorzystującą algorytm AORGK (ang. Adaptive Optimal Radially-Gaussian Kernel) o nieujemnej, matrycowej faktoryzacji dwuwymiarowej (ang. 2-Dimensional Nonnegative Matrix Factorization). W metodzie wykorzystano także algorytm k najbliższych sąsiadów oparty na teorii zbiorów rozmytych (ang. Fuzzy k Nearest Neighbour Classifier) oraz sieci neuronowe (ang. Back Propagation Neural Network). (**Hybrydowa metoda analizy obrazu do klasyfikacji wyładowań niezupełnych**)*

Keywords: partial discharge; feature extraction; adaptive optimal radially Gaussian kernel (AORGK); fuzzy k nearest neighbor classifier
Słowa kluczowe: Wyładowanie niezupełne, wybór cech obrazu, AORGK, 2DNMF, FkNNC

1. Introduction

High voltage electrical apparatus such as generators, power transformers, gas insulated switchgear (GIS), and XLPE cable etc. which are responsible for the generation and transmission of power energy are key elements of the power grid. The equipment failure would lead to a widespread blackout and great economic losses. It has been proved that partial discharge (PD) is a symptom and one of the main causes of high-voltage insulation failure [1-2]. It is of practical value to identify the underlying defects as earlier as possible by the diagnosis of partial discharges.

Since the defect types can be identified by the measured PD signals, much attention has been paid to the PD classification and recognition [3-5]. Among the existing PD detectors, Ultra high frequency (UHF) antenna has many advantages such as wide frequency band, high sensitivity, and strong anti-interference capability as well, compared with traditional pulse current method. In the cited articles, independent component analysis (ICA) [6], wavelet packet transformation (WPT) [7], and envelope comparison technology [8] have been applied for partial discharge UHF signals classification in electrical apparatus, showing that the UHF approaches can be effectively applied for PD classification. As been reported, PD pulse is a typically transient and nonstationary signal [9]. Joint time-frequency representation (TFR) can provide more comprehensive information of PD than time or frequency description alone. However, the most time-frequency analysis (TFA) methods have some intrinsic shortages. On the one hand, linear TFA methods such as short time Fourier transform (STFT), Gabor expansion (GS), and wavelet transform (WT), etc. cannot describe the instantaneous power spectral density (PSD) very well. Moreover, STFT cannot satisfy requirements to both high time and frequency resolution due to its constant window function [10]; the single wavelet basis makes WT difficultly suitable for all PD types [11]. On the other hand, quadratic TFA methods, represented by Wigner-Ville distribution (WVD), have serious cross-term interference when the signal contains multiple components in spite of the nice time-frequency resolution [12]. Consequently, it is needed to carry out deeper research on the time-frequency representation technology of PD signals and relevant feature extraction algorithm, to provide robust features for PD diagnosis in high voltage equipment.

In this paper, a new hybrid feature extraction method based on adaptive optimal radially Gaussian kernel (AORGK) time-frequency representation (TFR) combined with two dimensional nonnegative matrix factorization (2DNMF) technology is proposed for partial discharge UHF signals classification. Firstly, the UHF signals of four typical artificial defect models are measured and registered in Laboratory. Then AORGK is applied to acquire the time-frequency information of the measured UHF signals, and 2DNMF is further employed to compress the time-frequency amplitude matrices to extract various feature matrices. Finally, the extracted features are examined by fuzzy k-nearest neighbor (FkNN) classifier and back propagation neural network (BPNN).

2. Feature Extraction based on AOGRK and 2DNMF

2.1 AORGK-TFR

AORGK is a time-frequency analysis method originally proposed by Jones et.al has good resolutions in both time domain and frequency domain [13]. AORGK is developed and improved based on WVD, which could separate the independent term in signal effectively by restraining the influence of cross-terms and reflect the time-frequency information of original signals accurately.

AORGK introduces a short time ambiguity function (STAF) into the optimal radially Gaussian kernel (ORGK) [14] in which only one kernel is designed for the whole signal, to solve a problem that ORGK is not suitable for long time nonstationary signal processing. The AORGK-TFR of a signal $s(t)$ is given by

$$(1)\quad P_{AORGK}(t,w)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}A(t;\theta,\tau)\phi_{opt}(t;r,\tau)e^{-j\theta t-j\tau w}d\theta d\tau$$

where: t – time, τ – time delay, θ – frequency offset, j – prefix of imaginary number, $\phi_{opt}(t;\ r,\ \tau)$ – adaptive optimal kernel function (AOKF), $A(t;\ \theta,\ \tau)$ – STAF. STAF is defined as

$$(2)\quad A(t;\theta,\tau)=\int_{-\infty}^{\infty}s^{*}(u+\frac{\tau}{2})w^{*}(u-t-\frac{\tau}{2})s(u+\frac{\tau}{2})w(u-t-\frac{\tau}{2})e^{j\theta u}du$$

where: $w(t)$ – a symmetrical window function, "*" – complex conjugate. The variables τ and θ are usual ambiguity plane

parameters. Due to $w(u) = 0$ while $|u|>T$ (T is the length of time window), the ambiguity function is calculated only in the interval of $[t\text{-}T, t+T]$.

A high-quality time-frequency representation result is obtained when the kernel is well matched to the components of a given signal. Accordingly, AOKF $\phi_{opt}(t; r, \tau)$ can be obtained by solving the following optimization problem:

$$\max_{\phi} \int_0^{2\pi} \int_0^{\infty} \left|A(t;r,\psi)\phi(t;r,\psi)\right|^2 r dr d\psi \tag{3}$$

subject to

$$\left.\begin{aligned} \phi(t;r,\psi) &= \exp(-\frac{r^2}{2\sigma^2(\psi)}) \\ \frac{1}{4\pi^2}\int_0^{2\pi}\int_0^{\infty}\left|\phi(t;r,\psi)\right|^2 r dr d\psi &= \frac{1}{4\pi^2}\int_0^{2\pi}\sigma^2(\psi)d\psi \le \alpha \\ \alpha &\ge 0 \end{aligned}\right\} \tag{4}$$

where: $\sigma(\psi)$ – spread function, ψ – radial angle, $r = \sqrt{\theta^2 + \tau^2}$ – the radius variable. Spread function $\sigma(\psi)$ controls the spread of the Gaussian at radial angle ψ. The angle $\psi = \arctan(\tau/\theta)$ is measured between the radial line through point (θ, τ) and the θ axis. The AORGK-TFR for discrete signals can also be obtained by the similar procedure, more details is acquired in [13].

In this study, the AORGK-amplitude (AORGKA) matrix which is used to represent the original PD pulse is obtained as

$$A_{AORGK}(t,w) = \left|P_{AORGK}(t,w)\right| \tag{5}$$

2.2 2DNMF

2DNMF is an extension and improvement of NMF. Given a data matrix $\boldsymbol{V}_{m\times n}$ with $V_{ij} \ge 0$ and a prespecified positive integer $r < \min(m, n)$, NMF finds two non-negative matrixes $\boldsymbol{W}_{m\times r}$ and $\boldsymbol{H}_{r\times n}$ such that

$$\boldsymbol{V}_{m\times n} \approx \boldsymbol{W}_{m\times r} \times \boldsymbol{H}_{r\times n} \tag{6}$$

where: $\boldsymbol{W}_{m\times r}$ – the base vector matrix, $\boldsymbol{H}_{r\times n}$ – the coefficient matrix. So far many algorithms have been proposed for NMF computation such as multiplicative updates (MU) [15], projected gradients (PG) [16], and alternating least squares (ALS) [17], etc.

While NMF is applied to extract features from 2D AORGKA matrices, the matrices must be transformed into 1D vectors to achieve a great dimension reduction effect, which may lead to a quite high-dimensional vector space and huge computation burden. Moreover, the structural information hidden inside the AORGKA matrices are also destroyed by NMF. With an extension of NMF, 2DNMF perform a compression on original AORGKA matrices $\{\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_N\}$ in the horizontal and vertical direction simultaneously, where N is the sampling number of PD signals. 2DNMF is described briefly as follows [18]:

(1) The m training AORGKA matrices with a dimension of $p\times q$ are aligned into a $p\times qm$ matrix $\boldsymbol{X} = [\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_m]$. By NMF computation, $\boldsymbol{X}$ can be expressed as

$$\boldsymbol{X}_{p\times mq} \approx \boldsymbol{L}_{p\times d_1}\boldsymbol{H}_{d_1\times mq} \tag{7}$$

(2) The similar $q\times pm$ matrix $\boldsymbol{X}' = [\boldsymbol{A}_1', \boldsymbol{A}_2', \cdots, \boldsymbol{A}_m']$ is constructed by the same m training AORGKA matrices as well. $\boldsymbol{X}'$ is written as

$$\boldsymbol{X}_{q\times mp}' \approx \boldsymbol{R}_{q\times d_2}\boldsymbol{C}_{d_2\times mp} \tag{8}$$

(3) The $d_1\times d_2$ feature matrices $\boldsymbol{D}_k$ extracted from the kth AORGKA matrices $\boldsymbol{A}_k$ can be obtained as

$$\boldsymbol{D}_k \approx \boldsymbol{L}^{\mathrm{T}}\boldsymbol{A}_k\boldsymbol{R} \tag{9}$$

The NMF algorithm used in 2DNMF in the current study is PG. The theory and computation steps can be found at [16].

3. FkNN Classifier

In this work, a well-known FkNN classifier originally proposed by James et al. [19] is employed to perform PD classification. The k nearest neighbor (kNN) rule represents one of the most widely used classifiers in pattern recognition, which offers many advantages over other classifiers, including its simplicity, ease of parallel implementation, adaptability, and online learning. For an unknown sample x, the kNN classifier searches the k training samples adjacent to x and assigns x to the class that appears most frequently in the neighborhood of k samples. The similarity between two samples x and y in the kNN classifier is normally measured by the Euclidean distance d_{xy}, which can be expressed as

$$d_{xy} = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \tag{10}$$

where: x_k – the kth feature of sample x, y_k – the kth feature of sample y.

The FkNN classifier is an improvement over the kNN classifier, which uses the concepts of fuzzy logic to assign degree of membership of the samples to different classes while considering the distance of its kNNs. In Keller's method, the membership value of the unknown sample x in class i is given by

$$\mu_i(x) = \sum_{j=1}^{k}\frac{\mu_i(x_j)}{d_j^{\frac{2}{m-1}}} \Bigg/ \sum_{j=1}^{k}\frac{1}{d_j^{\frac{2}{m-1}}} \tag{11}$$

where: $d_j = \left\|x - x_j\right\|$ – Euclidean distance between x and x_j, $\mu_i(x_j)$ – the membership value of sample x_j for class i. This method calculates the membership of the unknown sample using the kNN labeled samples' memberships and similarity with the unknown sample. The calculated type of similarity in Keller's algorithm [19] is fixed on $1\Big/ d_j^{\frac{2}{m-1}}$. In the present study, $m = 2$ is chosen to represent Euclidean distance-based similarity measurement. The class label of the unknown sample x, $i(x)$ is determined as follows

$$i(x) = \arg\max_i(\mu_i(x)) \tag{12}$$

4. Experimental Arrangement

Artificial defect models are designed to produce PD sample data in the experiments, which are grouped into four categories [7, 20]: (1) Internal cavity discharge; (2) Surface discharge in oil; (3) Floating electrode discharge in oil; (4) Corona discharge in oil. The above PD types are described as model G, S, F, C for short.

The electrode structures corresponding to each type of PD are shown in Figure 1. In Figure 1(a), an epoxy board is supported by a circular insulating paper and placed on a plane electrode to form a gas cavity with a diameter of 38 mm. The three parts are subjected to firm pressure by two organic-glass boards to prevent leaking of surrounding insulation oil into the cavity during the experiments. The thicknesses of the epoxy board and the circular insulation paper are 1 and 0.15 mm, respectively [20]. In Figure 1(b) and (c), the epoxy board placed on the plane electrode is 1 mm thick. A metallic particle with a diameter of 0.3 mm is placed at the edge of the epoxy board for the P3 model. In Figure 1(d), the distance between the metal needle and 1mm thick epoxy board is 1mm.

In Laboratory test, the defect models are all put into a test tank filled with insulating oil. The PD pulses are detected by an UHF antenna, which has three resonant frequencies

equal to 261, 516, and 869MHz, respectively. The pass band around all resonant frequencies are approximately 150 ~ 320MHz, 430 ~ 620MHz and 740 ~ 1000MHz while the voltage standing wave ratio (VSWR) in all frequency bands do not exceed 5. The UHF antenna is connected to an amplifier with an overall bandwidth from 300 to 1000 MHz. A LeCroy Wavepro 7100 digital oscilloscope is used to display and store the PD data. The sampling frequency for recording the UHF signals is 5 GS/s.

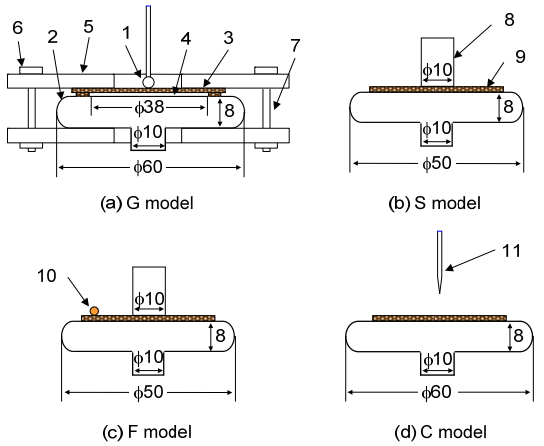

Fig. 1. Artificial defect models of PD
1-HV sphere electrode 2-LV plane electrode 3, 9-epoxy board 4-air gap 5-organic glass plane 6-insulating nut 7-insulating bolt 8-cylinder electrode 10-metallic particle

5. Results and Discussion

5.1 Data pre-processing

For each type of PD, 150 samples are measured under three applied voltages which is approximately 1.3~1.8 times of the inception voltage (50 samples per voltage, and overall 600 samples of 4 PD types). The UHF signals are originally measured in 1000ns range to cover the complete waveforms. Statistical analysis on the measured waveforms shows that all the signals can be expressed by 1000 samples. Then a segment of 1000 significant samples which is from the 2,451th to the 3,450th point is selected as the pre-selection process to facilitate the subsequent feature extraction. The pre-selection process is shown in figure 2.

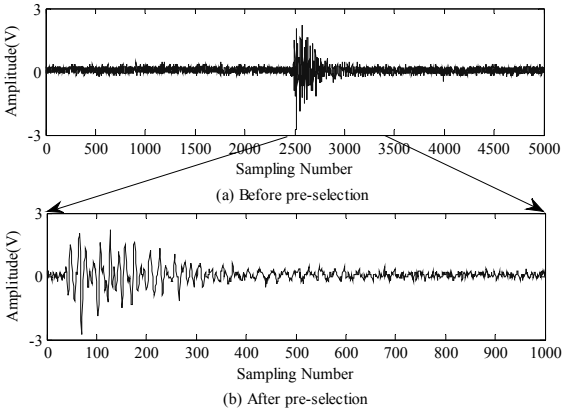

Fig. 2. Pre-selection of measured UHF signals

Due to existing randomness in PD pulses, there are available dispersions in the amplitude of the recorded UHF signals. In order to eliminate the influence of signal amplitude on AORGK-TFR, a normalization method is employed, shown as

$$x_n(t) = \frac{x(t)}{\max(|x(t)|)} \tag{12}$$

where: $x(t)$ – original measured UHF signal, $\max(|x(t)|)$ – maximum absolute value of $x(t)$, $x_n(t)$ – normalized signal. Figure 3 shows typical UHF signals corresponding to various PD types after data pre-processing.

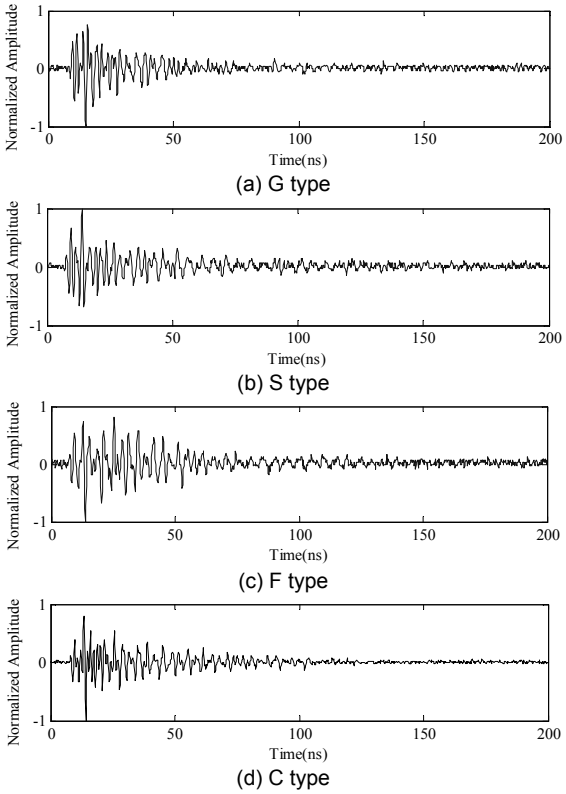

Fig. 3. Typical recorded UHF signals of various PD type

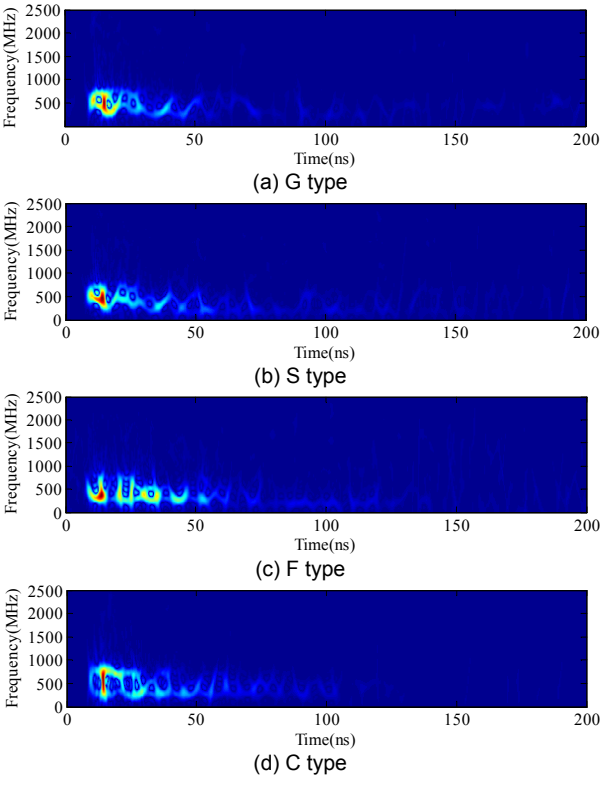

Fig. 4. Typical recorded UHF signals of various PD categories

5.2 Classification results

All the 600 samples are then pre-processed by the aforementioned approach. By employing AORGK, 600 AORGKA matrices are obtained as the time-frequency representations of PD recorded signals. The AORGKA matrices related to the PD signals in figure 3 are shown in Figure 4, revealing that AORGKA matrices associated with various PD categories are visibly different. However, the dimension of AORGKA matrix in this work is 512×1000 = 512000 (512 is the sampling number of FFT in discrete AORGK computation), which is too large and impossible for classification directly. It is necessary to reduce the data dimension to an acceptable scale for automated classification. In the meantime, the information of AORGKA matrix should be preserved as much as possible to obtain a brilliant classification performance.

As described in Section 2.2, 24 samples (6 samples per PD type) are randomly selected as the trained matrices, i.e. m = 24. Then each 512×1000 dimensional AORGKA matrix is compressed into a $d_1 \times d_2$ dimensional feature matrix by 2DNMF algorithm. In this study, (d_1, d_2) are assigned as (5, 5), (5, 10), (10, 5) and (10, 10) respectively, to explore the influence of different (d_1, d_2) combinations on the classification performance. Figure 5 gives the extracted feature matrices of (5, 5) combinations.

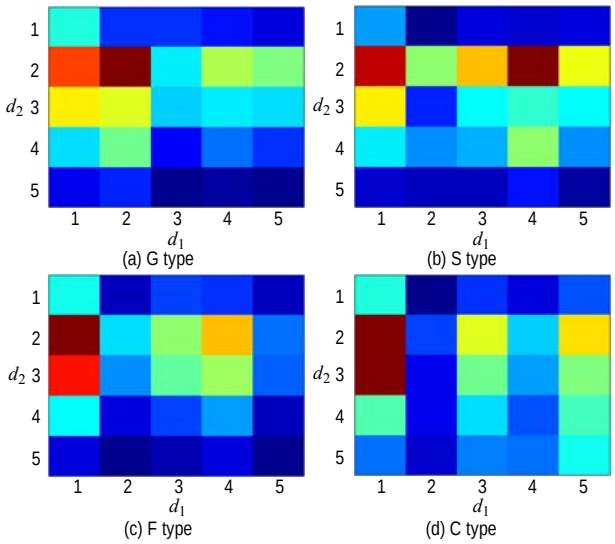

Fig. 5. Typical recorded UHF signals of various PD categories

The feature matrices of 600 samples are firstly transformed into vectors before classification. FkNN classifier is then responsible for the classification task. The experimental 600 samples are randomly partitioned into 300 training samples and 300 testing samples for 10 times, and the average classification rate is served as the final diagnosis result. In order to carry out comparison, a widely used PD classification classifier back-propagation neural network (BPNN) is implemented to the same experimental data. In this study, we apply a single hidden-layer network with the transfer function of sigmoid to train the BPNN. The number of neurons in input layer is equal to the feature dimension, and the number of neurons in output layer is 4. The number of neurons in hidden layer is determined by an empirical equation [13], shown as:

$$n_2 = 2 \times n_1 + 1 \tag{13}$$

where: n_2 – neuron number in hidden layer, n_1 – neuron number in input layer. Levenberg-Marquardt algorithm [22] is employed to train the BPNN classifier. More details on BPNN training can be found in [23].

The classification results of PD signals obtained by the above two approaches are reported in Table 1. It is clear that FkNN classifier achieves higher classification rates than BPNN under four circumstances of (d_1, d_2) combinations. In addition, FkNN classifier has an obvious advantage over partial discharge BPNN, which is its scalability. While a new discharge type is introduced into the classification issue, FkNN only needs to add a new group of training templates to the original training samples. In contrast with BPNN, the addition of a new PD type to the trained BPNN implies the retraining of the entire network, which is a very time-consuming task with huge sample numbers and large feature dimension. Thus, FkNN classifier has better generalization characteristics over BPNN.

Table 1. Classification results by FkNN classifier and BPNN

(d_1, d_2)	(5, 5)	(5, 10)	(10, 5)	(10, 10)
FkNN	87.50%	91.47%	93.73%	90.73%
BPNN	83.57%	84.70%	89.37%	88.67%

It can also be identified from Table 1 that, no matter FkNN or BPNN are used, the optimal classification accuracy is achieved with (10, 5) combination and (10, 10), (5, 10) combinations come to the second place. Besides, (5, 5) combination gives the minimum classification accuracy. As a result the lower feature dimension does not mean the better PD classification accuracy in the proposed feature extraction algorithm. Further analysis reveals that the feature dimension of (5, 5) combination has been decreased dramatically, whereas more information of the original AORGKA matrix may lose simultaneously, which generates comparatively lower classification accuracy. Though (10, 10) combination reserves relatively more information of the original AORGKA matrix, it probably introduces some redundant information and noises that causes a limitation on the classification accuracy in the meantime. The dimensions of (5, 10) and (10, 5) combination are the same, but the compression process in horizontal and vertical direction is not equivalent, which leads to differences in the classification accuracies of these two combinations.

Based on the above observations, it is evident the proposed hybrid feature extraction algorithm can be effectively applied to partial discharge classification by using a superior FkNN classifier compared with traditional BPNN. The proposed algorithm would provide a new partial discharge analysis tool for online condition monitoring of electrical apparatus.

6. Conclusion

The important conclusions arrived at based on the present investigations are for the followings:

(1) The time-frequency representations obtained by AORGK have visibly differences between different PD types, which imply the possibilities of AORGK to recognize different defect types.

(2) The feature dimension can be significantly reduced by applying 2DNMF which performs projections on the high dimensional matrices in the horizontal and vertical direction simultaneously. Moreover, the structural information of original AORGKA matrix is also preserved at the same time.

(3) The classification accuracies of FkNN classifier are all higher than that by BPNN under four circumstances of (d_1, d_2) combinations, i.e. (5, 5), (5, 10), (10, 5) and (10, 10). The highest classification accuracy 93.73% is obtained with (10, 5) combination. In addition, FkNN classifier has better generalization characteristic than BPNN due to its easy expansion capability.

The authors would like to thank the Fund for innovative research groups of China (51021005) and scientific research foundation of SKL of power transmission and

system security (2007DA10512708103) for the financial support provided. The authors also wish to thank Dr. T. Jiang for his kind support in the experiment.

REFERENCES

[1] Stone G.C., Partial discharge diagnostics and electrical equipment insulation condition assessment, *IEEE Trans. Diel. Electr. Insul.*,12 (2005), No. 5, 891-904

[2] Wang M. H., Partial discharge pattern recognition of current transformers using an ENN, *IEEE Trans. Power Del. 20* (2005), No. 3, 1984-1990

[3] Mazzetti C., Mascioli F.M.F., Baldini M., Panella M., Risica R., Bartnikas R., Partial discharge pattern recognition by neuro-fuzzy networks in heat-shrinkable joints and terminations of XLPE insulated distribution cables, *IEEE Trans. Power Del. 13* (2006), No. 4, 1035-1044

[4] Gu F.C., Chang H.C., Chen F.H., Kuo C.C. Partial discharge pattern recognition of power cable joints using extension method with fractal feature enhancement, *Expert Syst. Appl.*, 39 (2012), No. 3, 2804-2812

[5] Abdel-Galil T.K., Hegazy Y.G., Salama M.M.A., Bartnikas R., Fast match-based vector quantization parial discharge pulse pattern recognition, *IEEE Trans. Instru. Measur.*, 54 (2005), No. 1, 3-9

[6] Chang C.S., Jin J., Chang C.,Hoshino T., Hanai M., Kobayashi N., Online source recognition of partial discharge for gas insulated substations using independent component analysis, *IEEE Trans. Diel. Electr. Insul.*, 13 (2005), No. 4, 892-902

[7] Jiang T.Y., Li J., Zheng Y.B., Sun C.X., Improved bagging algorithm for pattern recognition in UHF signals of partial discharges, *Energies*, 4 (2011), No. 7, 1087-1101

[8] Pinpart T., Judd M.D., Differentiating between partial discharge sources using envelope comparison, *IET Scie. Meas.Technol.*, 4 (2010), No. 5, 256-267

[9] Abdel-Galil T.K., El-Hag A.H., Gaouda A.M., Salama M.M.A., Bartnikas R., De-noising of partial discharge signal using eigen-decomposition technique, *IEEE Trans. Diel. Electr. Insul.*, 15 (2008), No. 6, 1657-1662

[10] Cifrek M., Medved V., Tonkovic S., Ostojic S., Surface EMG based muscle fatigue evaluation in biomechanics, *Clin. Biomech.*, 24 (2009), No. 4, 327-340

[11] Shi G.M., Chen X.Y., Song X.X., Qi F., Ding A.L., Signal matching wavelet for ultrasonic flaw detection in high background noise, *IEEE Trans. Ultrason. FERR.*, 58 (2011), No. 4, 776-787

[12] Roshan-Ghias A., Shamsollahi M.B., Mobed M., Behzad M., Estimation of modal parameters using bilinear joint time-frequency distributions, *Mech. Syst. Signal Proces.*, 21 (2007), No.5, 2125-2136

[13] Jones D.L., Baraniuk R.G., An adaptive optimal-kernel time-frequency representation, *IEEE Trans. Signal Proces.*, 43 (1995), No. 10, 2361-2371

[14] Baraniuk R.G., Jones D.L., A radially-Gaussian, signal-dependent time-frequency representation, *ICASSP-91*, 5 (1991), 3181-3184

[15] Lee D.D., Seung H.S., Learning the parts of objects by non-negative matrix factorization, *Nature*, 401 (1999), No. 6755, 788-791

[16] Lin C.J., Projected gradient methods for nonnegative matrix factorization, *Neural Comput.*, 19 (2007), No. 10, 2756-2779

[17] Paatero P., Tapper U., Positive matrix factorization: a non-negative factor model with optimal utilization of error estimates of data values, *Envirometrics*, 5 (1994), No. 2, 111-126

[18] Li B., Zhang P.L., Liu D.S., Mi S.S., Ren G.Q., Tian H., Feature extraction for rolling element bearing fault diagnosis utilizing generalized S transform and two-dimensional non-negative matrix factorization, *J. Sound Vib.*, 330 (2011), No. 10, 2388-2399

[19] Keller J.M., Gray M.R., Givens J.A., A fuzzy k-nearest neighbor algorithm, *IEEE Trans. Syst. Man Cyber.*, 15 (1985), No. 4, 580-585

[20] Li J., Sun C.X., Gryzbowski S., Taylor C.D., Partial discharge image recognition using a new group of features, *IEEE Trans. Diel. Electr. Insul.*, 13 (2006), No. 6, 1245-1253

[21] Guo Z.H., Wu J., Lu H.Y., Wang J.Z., A case study on a hybrid wind speed forecasting method using BP neural network, *Knowl.-based Syst.*, 24 (2011), No. 7, 1048-1056

[22] Demuth H., Beale M., Neural network toolbox for use with MATLAB, *The MathWorks*, *Inc*. 2001

[23] Feng C.X.J., Gowrisankar A.C., Smith A.E., Yu Z.G.S., Practical guidelines for developing BP neural network models of measurement uncertainty data, *J. Manuf. Syst.*, 25 (2006), No. 4, 239-250

***Authors**: prof. Ruijin Liao, Room 227, High Voltage Lab, District A, Chongqing University, Shapingba District, Chongqing, P. R. China, 400030, E-mail: rjliao@cqu.edu.cn; Ke Wang, Room 205, High Voltage Lab, District A, Chongqing University, Shapingba District, Chongqing, P. R. China, 400030, E-mail: cquwangke@cqu.edu.cn*