Classification of Partial Discharges in Insulation Materials via Support Vector Machine and Discrete Wavelet Transform

Hazlee Azil Illias
Department of Electrical Engineering
Faculty of Engineering
Universiti Malaya
Kuala Lumpur, Malaysia
h.illias@um.edu.my

Neoh Ying Ting
Department of Electrical Engineering
Faculty of Engineering
Universiti Malaya
Kuala Lumpur, Malaysia
yingting@siswa.um.edu.my

Ong Zhen Yu
Department of Electrical Engineering
Faculty of Engineering
Universiti Malaya
Kuala Lumpur, Malaysia
zhenyu.95@siswa.um.edu.my

Masaaki Kando
Department of Electrical Engineering
Tokai University
Tokyo, Japan
mkkando@keyaki.cc.u-tokai.ac.jp

Azrul Mohd Ariffin
Department of Electrical & Electronic Engineering
College of Engineering
Universiti Tenaga Nasional
Kajang, Selangor, Malaysia
azrula@uniten.edu.my

Mohd Fairouz Mohd Yousof
Faculty *of Electrical and Electronic Engineering*
Department of Electrical Engineering
Universiti Tun Hussein Onn
Batu Pahat, Johor, Malaysia
fairouz@uthm.edu.my

***Abstract*—Long term partial discharges (PDs) within an insulation material of high voltage equipment can cause equipment failure. Thus, it is important to detect PDs within the insulation material and classify the PD type with high accuracy so that repair and maintenance can be performed effectively. In this work, three different types of PD, which include internal, surface and corona discharges, are measured from insulation materials. To evaluate the effect of noise on the PD measurement data, different levels of Additive White Gaussian Noise were added to the signals. Then, feature extractions were performed from the PD signals using Discrete Wavelet Transform (DWT). Different types of DWT families were used for feature extraction. The extracted features were then fed into support vector machine (SVM) for training and testing purposes. The classification accuracy of each test was recorded and compared. It was found that classification of PD signals using SVM as a classifier and DWT as a feature extraction yields reasonable classification accuracy results under different noise levels, which is in the range of 90%-99%.**

Keywords— Partial discharge, support vector machine, insulation materials, feature extraction, discrete wavelet transform

I. INTRODUCTION

Partial discharge (PD) is defined as a localized electrical discharge that only partially bridges the insulation between conductors [1, 2]. Although the magnitude of PD is very small, its effect towards power equipment cannot be ignored. Long term PD will lead to deterioration of insulation condition and eventually cause insulation failure. PD can accumulate at a local defect spot and the accumulation will form numerous branching, which is partially conducting discharge channel. Through PD measurement, PDs can be detected and the data can be used to access the insulation condition [3]. With some references to previous PD data, the type and location of the defect can be identified. This can prevent the intensification of PD, thus preventing insulation failure of an equipment.

PD phenomena is described as a sensitive detector to recognize and identify the defect types in power equipment [4]. Commercial PD detectors are widely available in the market nowadays. An experienced expert can identify the defect types in the test object by recognizing the PD patterns from the PD measurement data of the sample [5]. The analysis of PD data can be performed using computer software to visualize the PD data in terms of charge against time. Classification of PDs have been widely performed using artificial intelligence (AI) method to assist in recognizing the defect types. One of the common AI methods used for PD classification is support vector machine (SVM).

Several works have been performed on classification of PDs within insulation materials using SVM since the past. For example, in [6], SVM model was proposed to diagnose the fault in a power transformer. The transformer fault diagnosis model was created by applying binary classification principle of SVM. According to the results shown, SVM has a high potential in practice but there are still some problems need to be solved. For instances, the selection of kernel function and the optimization of parameters, which still require some further study in future research.

Some experimental investigations have shown the effectiveness of SVM in fault prediction of power transformers. Researchers in [7-8] applied SVM in classification of fault using less than 100 training and testing samples and high recognition rates between 80-100% were obtained. SVM was applied in [9] to classify PD sources in transformers. By using 6 moments as the input features, the best accuracy obtained was 98% for PD classification. Another research in [10] reported that by using equivalent time-frequency method, the input features can be extracted from the pulse wave shapes. Different PD patterns were successfully distinguished using SVM, such as corona, internal and surface discharges.

A work in [11] performed PD classification using SVM from artificial defects, including surface discharge in air, internal discharge in oil and corona discharge with remote earth. The proposed system achieved an excellent performance, where only 2 out of 60 samples were misclassified. SVM was used as a classifier for PD detection in GIS partial discharge pattern recognition based on the chaos theory [12]. Statistical features were extracted via chaos theory and used as the input data. The accuracy of PD classification that achieved in the work was 98%.

Many works have been performed on classification of PDs in insulation materials using SVM since the past. However, evaluation of SVM with Discrete Wavelet Transform (DWT) as a feature extraction using different DWT families for PD classification under different noise levels is less likely to be found. Thus, in this work, the classification accuracy of PD types from insulation materials using SVM and DWT was assessed under different levels of signal to noise ratio (SNR). Also, different families of DWT were used to observe the effect on the classification accuracy of PD types.

II. Methodology

Initially, three samples with different defect types were created to generate internal, surface and corona discharges. Fig. 1 shows the samples created. Each sample was stressed with a high voltage and PDs signals were measured using a standard PD measurement setup [13-16]. PD signals were stored in an oscilloscope. After satisfactory PD data were obtained, different levels of Additive White Gaussian Noise were added to the signals to evaluate the effect of noise on the PD measurement data. Then, feature extraction was performed on the PD signals using Discrete Wavelet Transform (DWT). Different types of DWT families were used for feature extraction. After that, the extracted features were fed into SVM classifier for training and testing. 100 data from each sample were used, where 70% were used for training and the remaining 30% for testing. The data were chosen randomly for training and testing for 10 different runs.

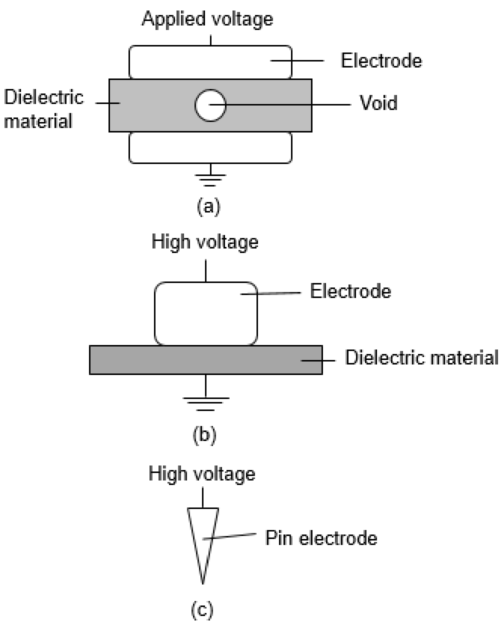


Fig. 1. Samples prepared for (a) internal, (b) surface and (c) corona discharges

SVM is a classifier used widely in statistical learning [17]. SVM has the ability to cope with large feature spaces, making it optimal for large amounts of classification data but able to achieve good accuracy in various fields. SVM works by developing a hyperplane within a high dimensional space by maximizing the margin data points or support vectors as shown in Fig. 2.

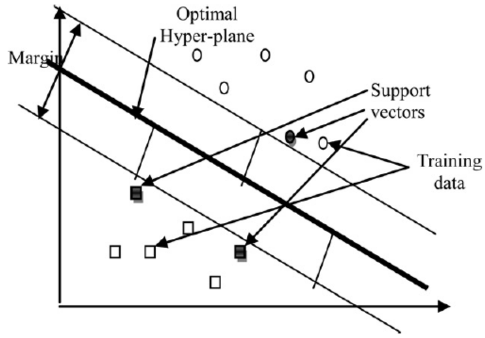


Fig. 2. Separation of two categories using SVM [6]

Sigmoid kernel function and Gaussian radial basis function are the common kernel functions used in SVM. Via supervised learning, training data are used with kernel functions that yield an optimal hyperplane. A radial basis function (RBF) kernel is defined as

$$K(x,x') = \exp\left(- \frac{||x-x'||^2}{2\sigma^2}\right) \quad (1)$$

where x and x' are the support vectors, σ is a RBF kernel parameter and c is a misclassification parameter.

DWT processes a signal by breaking it down into a set of basic functions called wavelets. It transforms a discrete time signal to a discrete wavelet representation to analyze the data. It extracts both frequency and instantaneous information. A signal, $x[n]$ is passed through a string of filters including high pass filter (HPF) and low pass filter (LPF) with an impulse response of $g[n]$ and $h[n]$. Passing the signal through these two filters with impulse response will result the output signal of convolution of the impulse response as

$$LPF: y[n] = (x * g)[n] = \sum_{k=-\infty}^{\infty} x[k]g[n-k] \quad (2)$$
$$HPF: y[n] = (x * h)[n] = \sum_{k=-\infty}^{\infty} x[k]h[n-k] \quad (3)$$

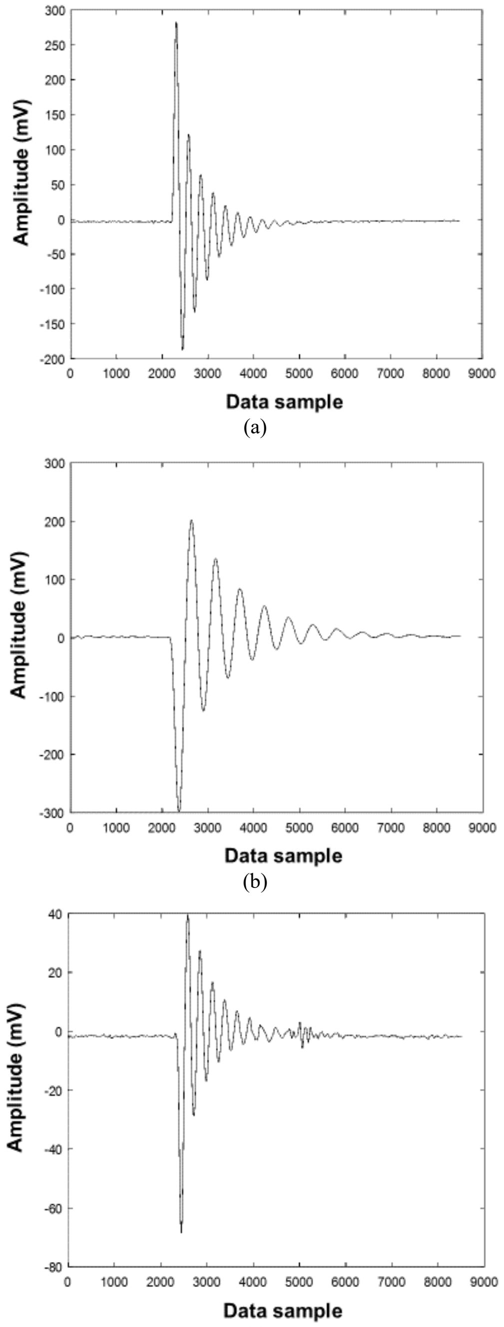
After passing through the filters, the LPF will provide the output with an approximated coefficient while the HPF will provide a detailed coefficient. After that, the signal is then passed to new LPF and HPF with new impulse response of $h[n]$ and $g[n]$. By passing the signal through each level of LPF and HPF, the frequency band of the signal is halved at the output, thus doubling the frequency resolution.

III. Results and Discussion

Fig. 3 shows the PD signals captured from each of the sample. Tables 1 to 3 show the classification accuracy results using SVM with DWT of different wavelet type from PD signals with SNR of 3 dB, 7 dB, 11 dB, 56 dB, 60 dB and 64 dB. Each run was tested with different combinations of training and testing data but by keeping the ratio of 70:30 between the training and testing data.

Referring to Table 1, when Haar DWT wavelet type is used as a feature extraction technique on the PD signals, the average classification accuracy decreases when the SNR is lower. The classification accuracy using SVM-Haar-cA DWT decreases from 98.89% to 90.78% when SNR decreases from 64dB to 3 dB. The classification accuracy using SVM-Haar-cD DWT decreases from 83.33% to 32.89% when SNR decreases from 64dB to 3 dB. Comparison between the approximated coefficient cA and detailed coefficient cD shows that the classification accuracy of SVM is higher when cA is used. The classification accuracy suffers significantly when SNR is lower for cD.

Table 2 show the average classification accuracy when Db1 DWT wavelet type is used as a feature extraction technique on the PD signals. The average classification accuracy decreases when the SNR is lower. The classification accuracy using SVM-Db1-cA DWT decreases from 98.89% to 90.78% when SNR decreases from 64dB to 3 dB. The classification accuracy using SVM-Db1-cD DWT decreases from 83.33% to 33.33% when SNR decreases from 64dB to 3 dB. Again, comparison between the approximated coefficient cA and detailed coefficient cD shows that the classification accuracy of SVM is higher when cA is used.



(a)
(b)
(c)

Fig. 3. PD signals obtained from (a) internal, (b) surface and (c) corona discharges

Referring to Table 3, when Db2 DWT wavelet type is used as a feature extraction technique on the PD signals, the average classification accuracy also decreases when the SNR is lower. The classification accuracy using SVM-Db2-cA DWT decreases from 98.89% to 91.33% when SNR decreases from 64dB to 3 dB. However, the classification accuracy using SVM-Db2-cD DWT on clean PD signals is only 33.33%. The classification accuracy is also 33.33% when SNR is reduced to 3dB. This shows that feature extraction using Db2-cD DWT is not suitable for classification of PD types.

When using approximated coefficient cA of DWT, DWT applies suitable threshold for only using low frequency approximate coefficients for signal reconstruction, excluding high frequency detail coefficients. There is a high chance that noise falls in the high frequency band and it is filtered out together with the detailed coefficients. Thus, SVM-Db1-cA DWT yields better classification accuracy that SVM-Db1-cD DWT.

TABLE I. CLASSIFICATION RESULTS OF SVM USING HAAR DWT WAVELET TYPE

Coeffi-cient Type	Run	Classification accuracy (%)						
		Clean Signal	SNR = 3dB	SNR = 7dB	SNR = 11dB	SNR = 56dB	SNR = 60dB	SNR = 64dB
cA	1	98.89	90.00	93.33	96.67	98.89	98.89	98.89
	2		91.11	92.22	95.56	98.89	98.89	98.89
	3		91.11	95.56	95.56	98.89	98.89	98.89
	4		91.11	93.33	96.67	98.89	98.89	98.89
	5		91.11	93.33	95.56	98.89	98.89	98.89
	6		90.00	93.33	96.67	98.89	98.89	98.89
	7		91.11	93.33	96.67	98.89	98.89	98.89
	8		91.11	93.33	96.67	98.89	98.89	98.89
	9		90.00	94.44	94.44	98.89	98.89	98.89
	10		91.11	92.22	96.67	98.89	98.89	98.89
Average		**98.89**	**90.78**	**93.44**	**96.11**	**98.89**	**98.89**	**98.89**
cD	1	83.33	32.22	33.33	33.33	77.78	81.11	82.22
	2		32.22	33.33	34.44	74.44	82.22	81.11
	3		33.33	34.44	35.56	77.78	83.33	81.11
	4		33.33	34.44	33.33	81.11	82.22	82.22
	5		34.44	32.22	33.33	81.11	82.22	82.22
	6		34.44	35.56	34.44	75.56	82.22	81.11
	7		32.22	34.44	32.22	78.89	83.33	82.22
	8		32.22	32.22	35.56	77.78	82.22	82.22
	9		32.22	32.22	33.33	78.89	82.22	82.22
	10		32.22	33.33	32.22	82.22	82.22	82.22
Average		**83.33**	**32.89**	**33.56**	**33.78**	**78.56**	**82.33**	**81.89**

TABLE II. CLASSIFICATION RESULTS OF SVM USING DB1 DWT WAVELET TYPE

Coeffi-cient Type	Run	Classification accuracy (%)						
		Clean Signal	SNR = 3dB	SNR = 7dB	SNR = 11dB	SNR = 56dB	SNR = 60dB	SNR = 64dB
cA	1	98.89	90.00	91.11	96.67	98.89	98.89	98.89
	2		90.00	93.33	96.67	98.89	98.89	98.89
	3		91.11	93.33	95.56	98.89	98.89	98.89
	4		91.11	91.11	96.67	98.89	98.89	98.89
	5		91.11	93.33	96.67	98.89	98.89	98.89
	6		91.11	93.33	96.67	98.89	98.89	98.89
	7		91.11	94.44	95.56	98.89	98.89	98.89
	8		91.11	93.33	96.67	98.89	98.89	98.89
	9		90.00	93.33	96.67	98.89	98.89	98.89
	10		91.11	92.22	96.67	98.89	98.89	98.89
Average		**98.89**	**90.78**	**92.89**	**96.44**	**98.89**	**98.89**	**98.89**
cD	1	83.33	32.22	33.33	35.56	77.78	81.11	83.33
	2		32.22	33.33	34.44	77.78	82.22	82.22
	3		33.33	34.44	33.33	80.00	83.33	82.22
	4		34.44	35.56	33.33	76.67	81.11	82.22
	5		34.44	33.33	32.22	78.89	82.22	82.22

	6		32.22	33.33	33.33	78.89	83.33	83.33
	7		35.56	34.44	34.44	80.00	83.33	81.11
	8		34.44	32.22	35.56	76.67	81.11	82.22
	9		32.22	35.56	33.33	80.00	82.22	82.22
	10		32.22	33.33	31.11	77.78	83.33	81.11
Average		**83.33**	**33.33**	**33.89**	**33.67**	**78.44**	**82.33**	**82.22**

TABLE III. CLASSIFICATION RESULTS OF SVM USING DB2 DWT WAVELET TYPE

Coeffi-cient Type	**Run**	**Classification accuracy (%)**						
		Clean Signal	**SNR = 3dB**	**SNR = 7dB**	**SNR = 11dB**	**SNR = 56dB**	**SNR = 60dB**	**SNR = 64dB**
cA	1	98.89	91.11	93.33	96.67	98.89	98.89	98.89
	2		92.22	92.22	96.67	98.89	98.89	98.89
	3		91.11	93.33	96.67	98.89	98.89	98.89
	4		90.00	94.44	95.56	98.89	98.89	98.89
	5		91.11	93.33	96.67	98.89	98.89	98.89
	6		91.11	95.56	95.56	98.89	98.89	98.89
	7		91.11	92.22	96.67	98.89	98.89	98.89
	8		91.11	93.33	95.56	98.89	98.89	98.89
	9		93.33	93.33	96.67	98.89	98.89	98.89
	10		91.11	95.56	95.56	98.89	98.89	98.89
Average		**98.89**	**91.33**	**93.67**	**96.22**	**98.89**	**98.89**	**98.89**
cD	1	33.33	33.33	34.44	32.22	34.44	32.22	33.33
	2		33.33	33.33	33.33	34.44	37.78	34.44
	3		33.33	34.44	34.44	34.44	30.00	34.44
	4		34.44	33.33	34.44	33.33	33.33	33.33
	5		34.44	33.33	34.44	32.22	31.11	28.89
	6		32.22	33.33	34.44	32.22	32.22	35.56
	7		34.44	34.44	32.22	32.22	28.89	30.00
	8		32.22	33.33	33.33	32.22	30.00	32.22
	9		32.22	33.33	34.44	32.22	31.11	32.22
	10		33.33	34.44	33.33	33.33	34.44	28.89
Average		**33.33**	**33.33**	**33.78**	**33.67**	**33.11**	**32.11**	**32.33**

IV. CONCLUSIONS

In this work, classification of partial discharge (PD) types within insulation materials has been successfully performed using support vector machine (SVM) and discrete wavelet transform (DWT). Three different types of PD, which include internal, surface and corona discharges, were successfully measured from the test objects. The extracted features from the PD signals using DWT were used as the input to SVM for training and testing purposes. From the results obtained, the average classification accuracy decreases when the signal to noise ratio (SNR) of PD signals is lower. Comparison between using the approximated coefficient cA and detailed coefficient cD of DWT as a feature extraction shows that the classification accuracy of SVM is higher when cA is used. Haar and Db1 DWT wavelet types are suitable to be used as a feature extraction for classification of PD types using SVM.

ACKNOWLEDGMENT

The authors thank the Malaysian Ministry of Education and Universiti Malaya for supporting this work through research grant of HIR (H-16001-D00048), Faculty Research Grant (GPF077A-2018), Impact-Oriented Interdisciplinary Research Grant (IIRG001-2020IISS) and UM Living Labs Grant (LL046-2019).

REFERENCES

[1] S. Boggs and J. Densley, “Fundamentals of Partial Discharge in the Context of Field Cable Testing,” IEEE Electr. Insul. Mag., vol. 16, no. 5, pp. 13–18, 2000.

[2] P. Discharge and L. Nitrogen, “Partial Discharge,” Cutler-Hanmer Predict. Diagnostics, vol. 96, no. 4, pp. 1–7, 1976.

[3] Z. Hao et al., “Case Analysis on Partial Discharge Signal of XLPE Cable T-Joint by Using High-Frequency Pulse Current Method,” Energy Procedia, vol. 141, pp. 545–550, 2017.

[4] E. Gulski et al., “On-site testing and PD diagnosis of high voltage power cables,” IEEE Trans. Dielectr. Electr. Insul., vol. 15, no. 6, pp. 1691–1700, 2008.

[5] F. C. Gu, H. C. Chang, F. H. Chen, and C. C. Kuo, “Partial discharge pattern recognition of power cable joints using extension method with fractal feature enhancement,” Expert Syst. Appl., vol. 39, no. 3, pp. 2804–2812, 2012.

[6] J. E. Fathima, and A. Venkatasami, “Transformer fault classification using support vector machine method,” International journal of advanced information and communication technology, vol. 1, no. 1, pp. 168-172, 2014.

[7] K. Bacha, S. Souahlia, and M. Gossa, “Power transformer fault diagnosis based on dissolved gas analysis by support vector machine,” Electric Power Systems Research, vol. 83, no. 1, pp. 73-79, 2012.

[8] D. H. Liu, J. P. Bian, and X. Y. Sun, X.-Y, “The study of fault diagnosis model of DGA for oil-immersed transformer based on fuzzy means Kernel clustering and SVM multi-class object simplified structure,” International Conference on Machine Learning and Cybernetics, 2008.

[9] R. Sharkawy, K. Ibrahim, M. Salama, and R. Bartnikas, “Particle swarm optimization feature selection for the classification of conducting particles in transformer oil,” IEEE Transactions on Dielectrics and Electrical Insulation, vol. 18, no. 6, pp. 1897-1907, 2011.

[10] S. Wenrong, L. Junhao, Y. Peng, and L. Yanming, L, “Digital detection, grouping and classification of partial discharge signals at DC voltage,” IEEE Transactions on Dielectrics and Electrical Insulation, vol. 15, no. 6, pp.1663-1674, 2008.

[11] L. Hao, and P. Lewin, “Partial discharge source discrimination using a support vector machine,” IEEE Transactions on Dielectrics and Electrical Insulation, vol. 17, no. 1, pp. 189-197, 2010.

[12] X. Zhang, S. Xiao, N. Shu, J. Tang, and W. Li, “GIS partial discharge pattern recognition based on the chaos theory,” IEEE Transactions on Dielectrics and Electrical Insulation, vol. 21, no. 2, pp. 783-790, 2014.

[13] M.Z.H. Makmud, H.A. Illias, C.Y. Ching, S.Z.A. Dabbak, "Partial Discharge in Nanofluid Insulation Material with Conductive and Semiconductive Nanoparticles," Materials, vol. 12, no. 5, pp. 816, 2019.

[14] J.K. Wong, H.A. Illias and A.H.A. Bakar, "Classification of Partial Discharge Measured Under Different Levels of Noise Contamination," Plos One, vol.12, no.1, pp.1-20, 2017.

[15] H.A. Illias, G. Altamimi, N. Mokhtar and H. Arof, "Classification of Multiple Partial Discharge Sources in Dielectric Insulation Material using Cepstrum Analysis-Artificial Neural Network," IEEJ Transactions on Electrical and Electronic Engineering, vol.12, no.3, 2017.

[16] H. A. Illias, M. A. Tunio, A. H. A. Bakar, H. Mokhlis and G. Chen, "Partial Discharge Phenomena within an Artificial Void in Cable Insulation Geometry: Experimental Validation and Simulation," IEEE Transactions on Dielectrics and Electrical Insulation, vol. 23, no. 1, pp. 451-459, February 2016.

[17] H.A. Illias, K.C. Chan, H. Mokhlis, "Hybrid feature selection–artificial intelligence–gravitational search algorithm technique for automated transformer fault determination based on dissolved gas analysis," IET Generation, Transmission & Distribution, Vol. 14, No. 8, pp. 1575-1582, 2020.