NOVEL FREQUENCY DOMAIN FEATURES FOR THE PATTERN CLASSIFICATION OF PARTIAL DISCHARGE SIGNALS

R. Ambikairajah[1*], B. T. Phung[1], J. Ravishankar[1], T. R. Blackburn[1], Z. Liu[1]

[1]School of Electrical Engineering & Telecommunications, University of New South Wales, Sydney, Australia
*Email: r.ambikairajah@student.unsw.edu.au

Abstract: The detection of partial discharge (PD) signals and identifying its patterns has been an area of interest in the analysis of insulation defects in high voltage cables. The features for classifying PD patterns are generally acquired from time-frequency domain analysis. This paper investigates the use of sub-bands to extract novel frequency domain features for PD pattern classification. The discrete Fourier Transform maps the sampled PD signal into equivalent discrete frequency bins. The frequency bins are grouped such that they form *N* sub-bands. The log energy that corresponds to each sub-band is obtained and these energies are smoothed using the discrete cosine transform, thus providing *N* coefficients as features in the classification of PD signals via the sparse representation classifier. Three types of artificial PDs, namely surface, corona and internal discharges, are measured in a high voltage laboratory environment at different applied voltages. The data obtained is then divided into training and testing sets for classification purposes. Results show that the features obtained using sub-bands are robust under noisy conditions, and in combination with the sparse representation classifier, it provides a better classification rate of PD signals compared to features obtained using wavelet analysis.

1 INTRODUCTION

There are increasing demands to manage the physical assets of the electricity supply industry and these calls for the implementation of predictive diagnosis and condition monitoring of assets, such as high voltage cables. Defects that generate partial discharge (PD) signals in cables, particularly in the cable joints and termination points, need to be identified, located and understood as part of the condition monitoring process. Understanding the type of partial discharge provides useful information in estimating the level of damage caused by PD. In general, there are 3 key types of PD including corona discharge, which occurs at sharp points that protrude from electrodes in gases and liquid, surface discharge that can occur in gases or in oil and internal discharge which occurs in both gas or oil filled cavities.

Attributes of a PD signal such as rise time, amplitude and recurrence rate provide useful information on the PD pattern [1]. PD signals can be analysed in both the time and the frequency domains. For a single PD type, it is sufficient to use time domain based methods to analyse the discharge. However, when there are multiple PD types, it is more effective to use frequency domain based analysis.

One of the challenges in condition monitoring of PD signals is that they are often corrupted by noise from the environment in a substation and surrounding areas resulting in the recorded PD signal being completely buried in noise. De-noising algorithms have been developed in both the time [2] and frequency domains [3] for use prior to analysing the PD signal.

Once a PD signal has been de-noised, a number of discriminatory features of the signal need to be identified and extracted. Methods of extracting these features is an ongoing area of interest. The discriminatory quality of the features that are chosen determine how many features are required for successful classification.

The wavelet transform is suitable for identifying transients in a signal and is used to obtain time-frequency representations of different PD types. Wavelet packet decomposition has been used to approximate PD signals and as a result, features that characterise various types of PD can be obtained from this decomposition [4]. The wavelet coefficients are used to calculate features such as the mean, kurtosis, normalised skewness and standard deviation at each level of wavelet decomposition. However, this paper presents novel features that characterise PD patterns in the frequency domain. These features are extracted with the use of a filter-bank, which consists of many sub-bands.

Commentary on numerous classifiers for pattern recognition is available in existing literature. The most common classifiers that have been used to date to classify the type of PD are the probabilistic neural network (PNN) [5] and the support vector machine (SVM), [6] and their performances have recently been compared against each other, demonstrating that the SVM classifier produced a higher accuracy rate that the PNN in testing 4 types of PD [7]. This paper introduces the sparse representation classifier for classifying PD, which has been used for face recognition [8]. A PD based classification system is shown in Figure 1 and there are two distinct stages. The training phase in this system has 3 steps: de-noising the signal, extracting features from the de-noised signal and then using these features to train a classifier. The testing phase for the system also has 3 steps: the signal is de-noised, features are extracted and based on the classifier input, a decision is made.

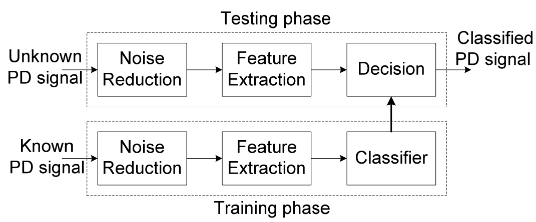

Figure 1: PD classification system, highlighting both the training and testing phases.

2 FEATURE EXTRACTION

2.1 Frequency domain based features

Before the features are extracted from the PD signal, the signal is de-noised using the Short-Time Fourier Transform (STFT) based signal boosting technique [3]. The PD signal is recorded over an entire AC cycle of 20ms with a sampling frequency of 100MHz, which results in $2x10^6$ time samples.

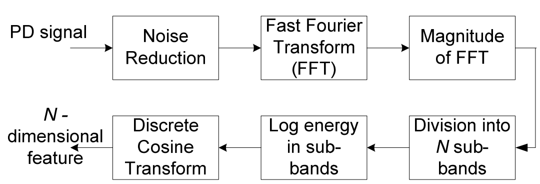

Figure 2: Feature extraction steps after de-noising

Figure 3 highlights the steps involved in the feature extraction process. The Fourier Transform maps the PD signal into 2 million discrete frequency bins. These bins are then grouped to form N sub-bands that are unequal in length. The log average energy corresponding to each sub-band is calculated. It is possible to reduce the dimensionality of the N-dimensional energy vector ($E_1, E_2, \ldots, E_N$) using Principal Component Analysis (PCA), but this is computationally inefficient. It has been shown that the energies in the sub-bands, as in Figure 3(a), can effectively be smoothed using the Discrete Cosine Transform (DCT) [9] to obtain transformed coefficients ($C_0, C_1, C_2, \ldots C_{N-1}$) as given in (1) and these are used in this paper as features for the classifier.

$$C_i = \beta \sum_{k=1}^{N} E_k \cos\left[i \frac{k\pi}{N} \right] \tag{1}$$

where: N is the number of sub-bands,
$i = 0, 2, \ldots, L-1$
L is the number of transformed coefficients $L \leq N$
β is a scaling factor.

The coefficient C_0 is the arithmetic mean of the log energies and C_1 describes the average spectral tilt. The spectral energy affected by any noise will be reflected in the higher order coefficients, C_{N-1}, C_{N-2}, etc. Due to the unique packing properties of the DCT, as in Figure 3(b), the dimensionality of the energy vector can be reduced by discarding these higher order coefficients. This provides a smoother representation of the log energy spectrum, when reconstructed using the inverse DCT.

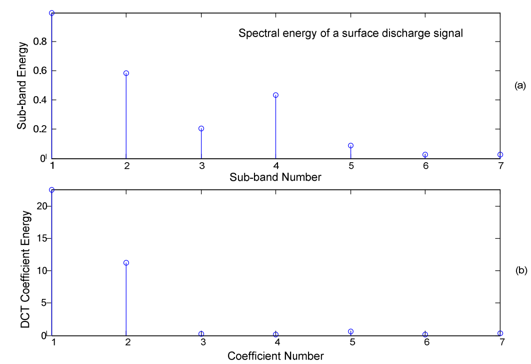

Figure 3: **(a)** Sub-band spectral energy of a surface discharge signal, **(b)** Corresponding DCT coefficient energy of a surface discharge signal

Figure 3(a) shows that the spectral energy is spread across 7 sub-bands. When the log energy of these sub-bands is transformed via the DCT, only the first few of the 7 coefficients (C_0 to C_6) carry most of the energy of the signal, which are then used as features for classification.

2.2 Time-frequency domain features

In order to establish the robustness of the frequency domain features, another set of features were obtained using wavelet decomposition. The PD signal is de-noised using the technique as outlined in [10] The Discrete Wavelet Transform (DWT) used to obtain these features has a 6-level decomposition (7 frequency sub-bands as shown in Figure 4) and uses the Daubechies wavelet, 'db8'.

The 7 sub-bands in Figure 4 match the 7 frequency sub-bands used in the frequency domain features. Using the wavelet coefficients in each band, the energy for each band is calculated and these energies are smoothed using the DCT shown in (1) and the resulting coefficients are used as features.

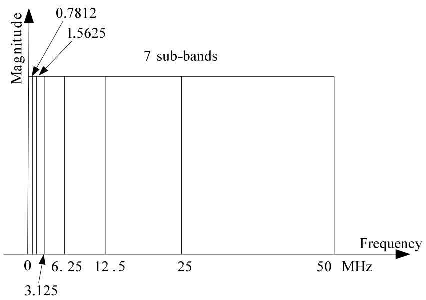

Figure 4: Frequency spectrum of the input signal split into 7 sub-bands due to 6-level decomposition

3 SPARSE REPRESENTATION CLASSIFIER

3.1 Classifiers

A new classifier based on sparse representation [8] has become popular in recent years. The sparse representation only selects basis vectors that compactly represents the signal and rejects the other vectors. In this paper, the discriminative nature of sparse representation is exploited to classify the PD signal.

If the feature vector is represented by U_{ij}, where *i* is the class index and *j* is the training data index. The training data from an i^{th} class is represented by a matrix $A_i = [U_{1i}, U_{2i}, \ldots, U_{ni}]$, where *n* is the total number of training data. Each *U* will contain a feature vector of *N*-dimension for a particular PD signal, thus A_i forms an $N \times n$ matrix. Now a global dictionary matrix, *W*, for all classes is formed for the Sparse Representation Classifier (SRC) by concatenating all the A_i matrices for each class where $W = [A_1, A_2, A_3, \ldots, A_i, \ldots A_k]_{Nx(nxk)}$. *k* is the total number of PD classes, *n* is the number of training vectors for each class and *N* is the number of feature vectors.

A test PD pattern ***y*** [N x 1] is now represented as a linear combination of all training vectors ***A*** [N x n] and ***x*** [(*n x k*) x 1] are the vector coefficients that will be sparsely represented in ***W***, as shown in (2)

$$y = W\ x \qquad (2)$$

Note that only the entries of ***x*** that are non-zero correspond to the class of ***y***. In solving (2) for ***x***, the class of the test pattern ***y*** can be found. The dimensionality of the feature vector N (rows) in the case of this paper, and as such it is smaller than the number of training data (columns) for all classes. This means that (2) is under-determined. Wright et. al [8] indicate that if ***x*** is sparse, (2) can be solved using l_1 norm minimisation as in (3).

$$\tilde{x} = \arg\min \|x\|_1 \quad subject\ to\ y = Wx \qquad (3)$$

y is assigned to the index that has the largest value in the sparse vector x and this index is located to identify the class.

4 RESULTS

4.1 Experimental Set-up

Experiments were set up in a high voltage laboratory to create a 5 class problem on the 3 types of PD: surface, corona and internal. A High Frequency Current Transformer (HFCT) that was created in-house was used to measure the PD signal as in Figure 5.

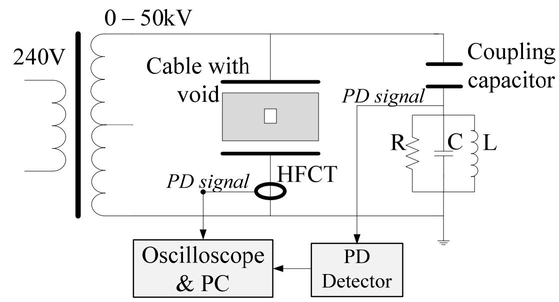

Figure 5: PD measurement circuit used in high voltage laboratory

For each type of PD, the applied voltage was steadily increased; this data recorded was used as training data for the SRC. The applied voltage was also steadily decreased for each PD type and this recorded data was used as test data. Variations on this included moving the HFCT to different positions in the circuit, namely next to the coupling capacitor, in order to determine differences in sensitivity and the polarity of the circuit was also reversed.

Class type	Surface discharge (Class 1)	Corona discharge (Class 2)	Internal discharge (Class 3)	Surface discharge - HFCT moved (Class 4)	Surface discharge - polarity reversed (Class 5)
No. of Training Samples	15	15	15	15	15
No. of Test Samples	10	10	8	6	10

Table 1: Details of test and training data for surface, corona and internal discharge for 5 classes

Spectral analysis of all 5 classes shows that there are considerable spectral magnitude differences in all sub-bands amongst the data.

4.2 Evaluation

The 7 dimensional features were extracted as per the process in Figure 2 for all the training and test data. These features were extracted using both the frequency domain technique and wavelet based techniques after undergoing the de-noising paper. The dictionary W was created in each case using the respective features and then $\tilde{x}$ was calculated from (3) using the test data. Results show that the frequency domain based features produced a classification accuracy of 100% for the 5-class problem compared to a classification accuracy of 89.47% for the wavelet based features.

In the case of the frequency domain features, the compactness of the transformed coefficients C_0 to C_6 was tested for classification accuracy, by creating a feature vector of 7 dimensions to 1 dimension. The accuracies are provided in Table 2.

Class type	Feature C0 to C7	Feature C0 to C6	Feature C0 to C5	Feature C0 to C4	Feature C0 to C3	Feature C0 to C2	Feature C0 to C1
Accuracy (%)	100	100	94.7	87.7	68.1	31.6	16.7

Table 2: Classification accuracies for selected transformed coefficients

Table 2 indicates that 6 transformed coefficients are sufficient for providing a 100% accuracy rate for frequency domain features, however to match the classification accuracy of wavelet based features, only 4 coefficients are required.

5 CONCLUSION

This paper presents the steps of feature extraction using the sparse representation classifier and shows that the frequency domain based features obtained using sub-bands are robust under noisy conditions. Results indicate that these features provide a better classification rate of PD signals compared to features obtained using wavelet analysis.

6 REFERENCES

[1] N.C. Sahoo, M.M.A. Salama, R. Bartnikas: "Trends in Partial Discharge Pattern Classification: A Survey," IEEE Transactions on Dielectrics and Electrical Insulation, Vol.12, No.2, pp. 248- 264, 2005

[2] D. Evagorou, A. Kyprianou, P. L. Lewin, A. Stavrou, V. Efthymiou, G. E. Georghiou: "Evaluation of Partial Discharge Denoising using the Wavelet Packets Transform as a Preprocessing Step for Classification," Annual Report Conference on Electrical Insulation and Dielectric Phenomena, pp.387-390, 2008

[3] R. Ambikairajah, B. T. Phung, J. Ravishankar, T. R. Blackburn: "A Comparison of Noise Reduction Techniques for Online Monitoring of Partial Discharge in High Voltage Power Cables," International Conference on Condition Monitoring and Diagnosis (CMD), pp. 267-270, 2010

[4] T. Yuming, Z. D. Wang, P. A. Crossley: "Partial Discharge Pattern Recognition based on 2-D Wavelet Transform and Neural Network Techniques," IEEE Power Engineering Society Summer Meeting, Vol.1, pp.411-416, 2002

[5] D. Evagorou, A. Kyprianou, P. L. Lewin, A. Stavrou, V. Efthymiou, A. C. Metaxas, G. E. Georghiou: "Feature extraction of partial discharge signals using the wavelet packet transform and classification with a probabilistic neural network," IET Science, Measurement & Technology, Vol.4, No.3, pp.177-192, 2010

[6] L. Hao, P. L. Lewin, S. J. Dodd: "Comparison of support vector machine based partial discharge identification parameters," IEEE International Symposium on Electrical Insulation, pp.110-113, 2006

[7] J.A. Hunter, L. Hao, P.L. Lewin, D. Evagorou, A. Kyprianou, G. E. Georghiou: "Comparison of two partial discharge classification methods," IEEE International Symposium on Electrical Insulation (ISEI), pp.1-5, 2010

[8] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry and Y. Ma: "Robust face recognition via sparse representation," IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 210-227, 2008

[9] A. K. Jain: "Fundamentals of Digital Image Processing," Prentice Hall: Englewood Cliffs, NJ, 1989

[10] X. Ma, C. Zhou, I. J. Kemp" "Interpretation of wavelet analysis and its application in partial discharge detection", IEEE Transactions on Dielectrics and Electrical Insulation, Vol. 9, pp. 446-457, 2002