SDEMPED 2003
Symposium on Diagnostics for Electric Machines, Power Electronics and Drives
Atlanta, Georgia, GA USA 24-26, August 2003

Statistical Pattern Analysis of Partial Discharge Measurements for Quality Assessment of Insulation Systems in High Voltage Electrical Machinery

Birsen Yazici
Rensselaer Polytechnic Institute
Electrical, Computer and Systems Engineering
Troy, NY 12180
yazici@ecse.rpi.edu

Abstract - **In this paper, we present a new statistical analysis method of phase resolved partial discharge (PD) measurements for the quality assessment of the electrical insulation in high voltage machinery. The method is based on a supervised classification approach which utilizes histogram similarity analysis. The motivation for choosing histogram similarity analysis is twofold: First the phase resolved PD measurement itself is, in fact, a two dimensional histogram. Therefore a histogram matching based approach suits the very nature of the data. Second histogram similarity analysis combines the typical statistical parameters, such as mean, variance and higher order moments, used in recognition of PD measurements in a statistically powerful and rigorous way. In our study, we utilize various histogram types and similarity analysis, including correlation, Chi-Square and Kolmogorov Simirnov tests. Our experimental study on the laboratory samples demonstrates that the method shows strong potential in detection and identification of insulation defects. The results from our study suggest that the proposed method provides a powerful, general and mathematically simple approach to the analysis of phase resolved PD measurements.**

I. INTRODUCTION

Partial discharge (PD) analysis has been established as a reliable diagnostic tool to asses the insulation systems for their integrity and design deficiencies. Interpretation of the PD patterns can reveal the source and the reason for its occurrence, and therefore has been used as a condition monitoring and quality control tool by the manufacturing industry [1]. For many years, the interpretation was performed by human experts. In recent years, advancement of computer hardware and pattern recognition techniques has provided automation and improvement of the PD interpretation process. As a result of the computer aided processing, massive amounts of PD measurements can be interpreted efficiently and reliably. Among the well-known pattern recognition methods applied to PD analysis are expert systems, neural networks, fuzzy classifiers, fractal models, and statistical methods, and so on [2]-[5].

In this paper, a statistical analysis method for the intelligent characterization and interpretation of phase resolved PD patterns are proposed. Phase resolved PD measurement is viewed either as a two dimensional histogram in which the two axes consist of phase windows and PD magnitude or as a collection of one dimensional histograms in which the axis consists of PD magnitude. For quality assessment of high voltage insulation equipment, a set of PD training measurements are acquired from good and defective insulation equipment. A representative PD measurement and a threshold value are determined by histogram similarity analysis of the training data for each insulation quality class. The quality class representatives and thresholds are then stored in a data base to be used as baselines for normal bars and defective bars during the testing stage. In the testing stage, PD measurement is acquired from an unknown insulation condition and its distance is measured to the representative PD measurement using histogram similarity measures. If the distance is within the limits of the normal insulation threshold, the insulation is cleared. Otherwise, it is tagged as a faulty insulation.

II. HISTOGRAM TYPES AND SIMILARITY MEASURES

In a phase resolved PD measurement, PD observation takes place over several voltage cycles and the PD pulses are grouped by their phase angle with respect to the 60 Hz sine wave. Therefore, a phase resolved PD measurement is a two dimensional histogram of PD events with respect to discharge magnitude and phase angle. Hence the very nature of the data motivates a histogram based statistical analysis. In our study, phase resolved PD measurement is viewed as three different kinds of histogram:

- The first histogram is two dimensional consisting of the PD measurement itself. The phase windows and the PD magnitudes form the two dimensions of the histogram. The joint probability mass function of phase windows and PD magnitudes can be computed by normalizing the histogram to a unit area. In our discussion, we shall refer to this histogram as 2D histogram.
- For a given phase window, the number of occurrences of individual PD pulses provides a histogram of PD magnitudes. Hence, phase resolved PD measurements can be viewed as a collection of one dimensional histograms in which the dimension is the PD magnitude. In our discussion, we shall refer to this collection of histograms as 1D phase histograms.
- To simplify the analysis, one can consider only the total number of occurrences for each phase window or PD magnitude, i.e., the sum of the individual PD events for each phase window or for each PD magnitude. We shall refer to the histogram obtained by the sum of the individual PD events for each phase window as the phase marginal histogram or X-marginal histogram, and the histogram obtained by the sum of the individual PD events for each PD magnitude as the magnitude marginal or Y marginal histogram.

In our study, we used various histogram similarity measures to asses the discrimination value of each one. These measures are sample correlation, Kolmogorov-Simirnov distances and Chi-square tests. To keep the notation simple, we shall describe these measures for one dimensional histogram. The two dimensional description is a straightforward extension of the one dimensional version.

To facilitate our discussion, let us introduce some notation. Probability Mass Function or Normalized Histogram, $P(i)$, $i=1,\dots,N$. The definition of cross Correlation value is standard. Kolmogorov-Simirnov distance is defined as follows [8]: Cumulative Distribution Function:

$$H(k)=\sum_{i=1}^{k}P(i), \quad k=1,\dots,N.$$

iia) Kolmogorov Simirnov Distance 1:

$$KS1=\max_{k}\left|H_1(k)-H_2(k)\right|$$

iib) Kolmogorov Simirnov Distance 2:

$$KS2=\sum_{k=1}^{N}\left|H_1(k)-H_2(k)\right|$$

iii) Chi-Square Test [8]: Since we really do not know the "true" probability mass functions of the PD events, what we implement is a modified chi-square test which checks whether the two measurements $h_1(i)$, $i=1,\dots,N$, and $h_2(i)$, $i=1,\dots,N$ are drawn from the same population by some unknown probability distribution function. It can be shown in this case that the chi-square statistics is given by:

$$\chi^2=\sum_{i=1}^{N}\frac{\left(\sqrt{A_2/A_1}h_1(i)-\sqrt{A_1/A_2}h_2(i)\right)}{h_1(i)+h_2(i)}$$

where $\sum_{i=1}^{N}h_1(i)=A_1$ and $\sum_{i=1}^{N}h_2(i)=A_2$.

In our implementation, $h_1(i)$ is the first histogram, and $h_2(i)$ is the second histogram of PD events.

III. QUALITY ASSESSMENT OF THE INSULATION SYSTEM

The proposed quality assessment method is a "supervised" approach, that is, the type and the number of insulation defects, for which a given insulation is tested, is predetermined. The method can be utilized for detection, as well as identification of insulation faults. The detection problem involves an automatic decision making to determine whether a given test insulation is good or not. The identification problem, on the other hand, involves both detection, as well as, classification of the insulation defect to one of the known defect classes. These defect classes may be i) suspended metal particulate in insulation, ii) armor degradation, iii) conductor/insulator delamination, iv) insulator/insulator delamination, v) insulator/armor delamination, among others. The outcome of such a quality assessment system is one of the followings: i) Good insulation, ii) defective insulation with the defect type being one of the classes in the data base iii) defective insulation with unknown defect class.

The method consists of four stages: preprocessing, training, testing, and postprocessing. In the preprocessing stage, the PD data is subjected to some signal conditioning methods, such as gain normalization, phase synchronization and excitation noise suppression. In the training stage, the features of defective and good insulation are learned from a set of so called training data. The learning process involves feature extraction from raw data and statistical analysis. The statistical analysis of the features leads to a set of representatives for each defect class and an optimal decision making rule to obtain a statistically optimal threshold. Finally, the representatives of each class along with the respective thresholds are stored in a data base to be used as baselines during the testing stage. In the testing stage, features are extracted from a test PD data coming from an unknown insulation quality and a statistical distance between the representatives of each class and the features is computed. The resulting distances are compared with the thresholds in the data base to check whether the test measurement belongs to one of the classes in the data base. In the post processing stage, the decisions from multiple PD measurements are combined to increase the confidence level and a probability is associated with the decisions to quantify the accuracy of the process.

III.A Preprocessing

In this stage, the raw PD measurements are processed i) to convert from bipolar mode to a unipolar mode, ii) to synchronize with the 60 Hz supply voltage, and iii) to normalize with respect to gain factor. Also, digitally filtering can be applied to suppress the excitation noise in the measurements, if needed.

III.B Training

It is well-known empirically that the low count high magnitude PD events contain more discrimination information than the low magnitude high count PD events. During the feature extraction process, in order to highlight the relative importance of the low PD counts, we take the logarithm of the PD event counts and suppress the effect of high count low magnitude events by thresholding.

After feature extraction, a representative is computed for each class by averaging the PD data. Let R_i denote the representative of the class **i**, $j=1,\dots,N_i$ and P_i^j, $j=1,\dots,N_i$ be the jth member of the class i where N is the total number of classes and N_i is the number of samples available for class i. Next, an intra class distance between the class representative, R_i, and each of its member, P_i^j, is calculated by using the histogram similarity measures introduced in Section II. $D_i^j = d(P_i^j, R_i)$ where d stands for one of the histogram similarity measures. Note that in the case of 1D phase histograms, the distance function is a vector in which each entry is the distance between the member and reference 1D phase resolved histograms of the respective phase window. To clarify this difference, we introduce vector notation for the 1D phase histograms : $D_i^j(\omega) = d(P_i^j(\omega), R_i(\omega))$, $\omega = 1,...,256$. ω corresponds to the phase window. In order to obtain an optimal radius for each class, the sample mean and standard deviation of the distances within each class are calculated. For the 2D histograms, the intra class mean radius and standard deviation are given as follows:

$$\bar{D}_i = \frac{1}{N_i}\sum_{j=1}^{N_i} D_i^j \qquad \sigma_i = \sqrt{\frac{1}{N_i}\sum_{j=1}^{N_i}\left(D_i^j - \bar{D}_i\right)^2}.$$

Next, an α unit standard deviation tolerance is allowed for each class radius. Note that in the case of Gaussian distribution of the intra class distances, α is typically chosen to be 2 to provide 99% confidence interval. However, in our scheme, it is kept as an input parameter to allow the user to utilize her/his engineering judgment in setting the tolerance. The class radius, for the respective 2D and 1D histograms are given as follows: $\tau_i = \bar{D}_i + \alpha\sigma_i$ Finally, for each class $\{R_i, \tau_i\}$ are stored in the data base as the representative and the threshold of each class. Note that in both 2D and 1D histogram s, the class radius is given by a scalar value.

III.C Testing

Given a test data, we first subject the raw measurements to the operations discussed in the preprocessing stage prior to extracting the relevant features. These features are discussed in the training stage. Next, we compute the distance between the test feature, T, and the references of each class R_i using one of the histogram similarity measures described in Section II. For the 2D histogram, the distance is given by a scalar quantity:

$$\Delta_i = d(T, R_i) \qquad i = 1,...,L.$$

where L is the number of classes. Next, we check if the distance between test feature and the references of each class, Δ_i, is less than the class threshold, τ_i. If it is greater than all the thresholds, the test measurement is tagged as defective. However, the defect is not one of the types for which the algorithm is trained to recognize. If it is less than one or more thresholds, the test measurement is assigned to the class to which the distance between the associated reference and test feature is minimum. If the assignment is the class of good insulation, the measurement is tagged as nondefective.

To improve the accuracy of our decision making, we repeat the testing scheme described above for multiple PD measurements for each insulation equipment. Decisions obtained from each measurement and the distances of each test feature to the class representatives are fed to the postprocessor to finalize the decision on the insulation equipment.

III.D Postprocessing

In the postprocessing stage, there are two major tasks: i) combine the decisions based on each test PD measurement to classify the insulation equipment. ii) Associate a probability to each decision made to reflect the accuracy of the classification process. We shall start with the second task since making a final decision on the insulation equipment depends on the accuracy of each decision made on individual PD measurements. Recall that during the testing stage, a distance between the test feature and the representative of each class is computed: $\Delta_i = d(T, R_i)$, $i = 1,...,L$

IV. EXPERIMENTAL RESULTS

In the training stage, we used 9 generator bars with various insulation defects and 2 generator bars with good insulation quality. From each bar, 5 PD measurements were collected. There are 9 classes, of which 8 are defect classes and 1 is the good insulation class. Totally 52 measurements are used in the experiment.

For the testing stage, we computed the "defect detection", "false positive rate" and "class identification" power of each histogram similarity measure using 2D, 1D and marginal histograms. Probability of defect detection is defined as the likelihood of tagging a test measurement coming from a defective bar as defective. Probability of false positive is the likelihood of tagging a test measurement as defective while it is coming from a nondefective insulation and identification power is the likelihood of identifying the class type of test measurements correctly. In all of these experiments the α tolerance level is chosen to be 2 unit standard deviation. Note that it is desirable to have high probability of detection and class identification power and low probability of false positive. Among the type of histograms used, the performance of the 1D phase resolved histogram appears to be the best in all categories. However both X and Y marginal histograms perform poorly as compared to the first two histogram types, particularly in terms of probability of false positive and class identification power.

With respect to the histogram similarity measures, Kolmogorov-Simirnov test2 performs the best in the case of 1D phase resolved histogram. Out of 42 measurements, collected from the defective bars, all of them correctly tagged as defective yielding 100% defect detection capability. Out of 10 measurements collected from non defective bars all are tagged as "good bars" yielding 0% probability of false positive. Out of 52 measurements collected from 9 classes all except 1 of them are correctly classified yielding 98% class identification power. Table 1 summarizes the results of Kolmogorov-Simirnov test 2. The detailed experimental results for each histogram similarity measure and histogram type are summarized in the confusion matrices shown in Table 2. Note that in a confusion matrix, the number in the ith row and jth column of this table shows how many samples from class i are classified as class j. In the case of perfect classification, the confusion matrix is diagonal. Extensive experimental results show that the performance of the correlation, Kolmogorov-Simirnov1 and 2 tests using 1D phase histograms is excellent.

We performed extensive experiments to show the feasibility of the postprocessing stage in terms of improving the results of the testing stage and providing additional information on the accuracy of the classification performed in the testing stage. Our results confirm that the proposed postprocessing method is very effective in improving the accuracy of the final classification.

Table 3 summarizes the likelihood probabilities of the bars for each class. Although all the test bars have the highest class likelihood probability for the true class, some bars exhibit significant similarity to more than one class. For example in the case of bar with metal wire, it was found that the test bar belongs to the class of "metal wires" with 93% probability while it may also belong to the class of nondefective bars with 88% probability. The class likelihood probabilities may help user to decide if additional measurements and testing is needed before the final decision on the insulation is called.

V. CONCLUSION

In this paper, we presented a new statistical analysis method of PD measurements for the quality assessment of electrical insulation systems in high voltage electrical machinery. The method is based on a supervised classification approach which utilizes histogram similarity analysis. The motivation for choosing histogram similarity measures is twofold: First, the phase resolved PD measurements themselves are in fact two dimensional histograms. Therefore, a histogram matching based approach suits to the very nature of the data. Second, histogram similarity measures combine the typical statistical parameters, such as mean, variance and higher order moments used in recognition of PD measurements in a statistically powerful and rigorous manner. Our experimental study on laboratory samples demonstrates that the method shows strong potential in detection and identification of insulation defects.

REFERENCES

[1] CIGRE WG 21.03, "Recognition of Discharges," Electra, Vol. 11, pp. 61-98, 1969.

[2] E. Gulski and A. Krivda, "Neural Networks as a Tool for Recognition of Partial Discharges," IEEE Transactions on Electrical Insulation, Vol. 28 No. 6, pp. 984-1001, 1993.

[3] E. Gulski and L. Niemeyer, "The importance of Statistical Characteristics of Partial Discharge Data", IEEE Transactions on Electrical Insulation, Vol. 27, pp. 60-69, 1992.

[4] E. Gulski, Computer Aided Recognition of Discharges Using Statistical Tools, Delft University Press, 1991.

[5] L. Niemeyer, L. Pietronero and H.J. Wiesman, "Fractal Dimension of Dielectric Breakdown," Physical Review Letters, Vol. 52, no. 12, pp. 1033-1036, 1984.

[6] B. Yazıcı, J. R. Krahn and K. Rao, "Statistical Pattern Recognition Techniques for Defect Analysis of Large AC Machine Insulation Systems," presented at CIGRE 1997 Conference. April, 1997.

[7] B. Yazıcı and J. R. Krahn, "Statistical Analysis Methods of Partial Discharge Measurements in HV Electrical Machinery," General Electric patent filed in April, 1997. Docket No. RD-25,171.

[8] *Goodness-of-fit techniques*, Edited by R. B. D'Agostino and M. A. Stephens, III Series: Statistics, textbooks and monographs; 1986.

TABLE 1.
PERFORMANCE OF THE TESTING STAGE USING 1D PHASE HISTOGRAMS

Test Type	Defect Detection	False Positive	Defect ID
Chi-squares	100%	70%	78%
Correlation	100%	0%	96%
Kolmogorov-Simirnov	100%	0%	96%
Kolmogorov-Simirnov	100%	0%	98%

TABLE 2.
CONFUSION MATRIX FOR THE KOLMOGOROV-SIMIRNOV2 TEST USING 1D PHASE HISTOGRAMS

	Supp	Meta	NoG	CDe	Puff	Thrn	SDsc	Arm	GBa	Unkn
Supprs	5									
MetalW		5								
NoGrade			9	1						
CDelam				5						
Puffed					5					
ThrmVE						4				
SDischg							4			
ArmCnt								4		
GBars									10	

TABLE 3.
CLASS LIKELIHOOD OF EACH BAR FOR THE KOL_SIM2 TEST USING 1D PHASE HISTOGRAMS

	Suppr	Metal	NoGr	CuDel	Puffed	Therm	SDisc	ArmC	GBars
Grade	0.94	0.15	0.43	0.34	0.79	0.72	0.61	0.48	0.14
TT54	0.72	0.93	0.00	0.00	0.82	0.13	0.00	0.56	0.88
NoGr	0.24	0.00	0.93	0.84	0.00	0.00	0.00	0.00	0.00
NoGr	0.47	0.00	0.93	0.60	0.00	0.00	0.00	0.00	0.00
TT55	0.70	0.00	0.84	0.93	0.00	0.00	0.00	0.00	0.00
TT59	0.88	0.79	0.32	0.00	0.93	0.75	0.11	0.76	0.57
2E18	0.90	0.32	0.10	0.14	0.82	0.93	0.64	0.38	0.14
2G17	0.87	0.00	0.46	0.40	0.56	0.56	0.93	0.09	0.00
2G18	0.81	0.66	0.44	0.00	0.82	0.00	0.00	0.93	0.27
GBar	0.70	0.90	0.00	0.00	0.78	0.00	0.00	0.32	0.93
GBar	0.61	0.89	0.00	0.00	0.72	0.00	0.00	0.08	0.94