Diagnosis of Partial Discharge Signals using Neural Networks and Minimum Distance Classification

Hans-Gerd Kranz

Bergische Universität, Gesamthochschule Wuppertal,
Germany

ABSTRACT

This paper describes two different methods to classify partial discharge (PD) phenomena by an automated personal computer-aided system. The first is concerned with common minimum distance classification using statistical data on pulse quantities such as apparent charge, energy and phase. Applying the correct algorithms and features, such a system is able to discriminate between unknown defects using conventional discharge patterns. The classification with neural networks discussed here offers the new possibility of classifying the shape of the PD pulses without using statistical tools for data reduction. Examples of diagnostic decisions are shown using a gas insulated switchgear system with several artificially introduced defects. The reliability of a diagnosis must be considered as a main precondition for those who are involved in detection and evaluation of partial discharges. This paper gives new information to estimate the reliability of such a diagnosis for both ways of time-resolved detection evaluated by neural networks and classic phase-resolved PD evaluation. A two step strategy of time-resolved preclassification and automated phase-resolved evaluation is introduced in this report. No measuring evidence to prove the described methods is presented in this paper, but to clarify the descriptions, a few examples are shown.

1. INTRODUCTION

THIS approach should be of interest to those involved in detection and evaluation of partial discharges (PD) in modern insulation systems. A new concept for automated PD diagnosis adapts methods of pattern analysis and machine intelligence for this special application. It has indeed been confirmed by experience [10] that PD is a random physical process, characterized by a random scatter of location and magnitude. Additionally each PD sequence contains random events with a large variation in duration and in the phase-charge distribution profile at several applied voltages. In todays environment of powerful microcomputers, it is possible to provide methods of pattern recognition and artificial intelligence for noise reduction and defect identification using computer-driven instrumentation. Many microcomputer-based PD measurement systems are capable of recording and storing the complex time behavior of several PD parameters during a test cycle. To relieve the operator of the mass of informa-

0018-9367/93/$3.00 © 1993 IEEE

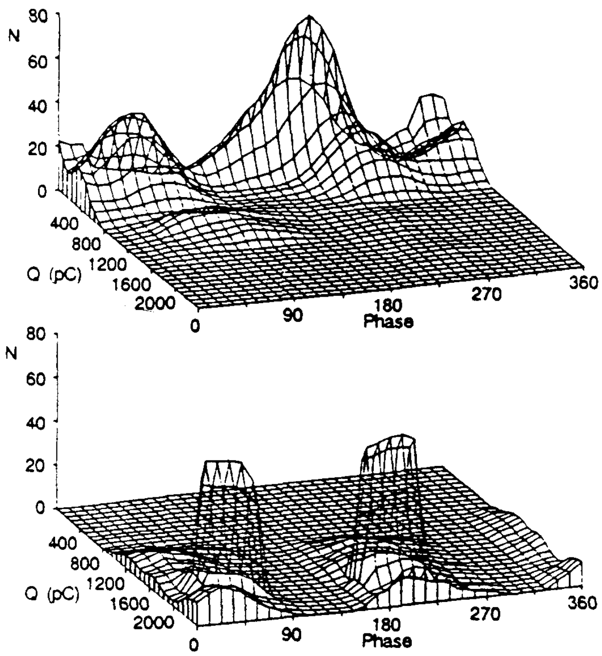

Figure 1.

Conventional Pattern recognition, Patterns of PD activity in a 735 kV transformer winding measured by Vaillancourt and Malewski [7].

tion, the measured PD data must be purified from noise and evaluated to give a reliable and compromised result from the test. It is assumed that the PD evaluation and diagnosis can be discussed separately from the hardware and the test setup.

Multichannel pulse analyzing systems can be driven by any amplifier which produces an amplified PD signal. Usually for diagnosis the number of discharges and/or the measured charge is plotted against the phase-of-appearance. The most common diagnostic tools are two or three dimensional distributions or histograms which exhibit typical patterns of the PD activity (Figure 1). In tests, such a conventional diagnosis is usually obtained at a constant voltage stress, because any variation of the test voltage would add a fourth dimension to the discharge patterns requiring multiple plots. Powerful microcomputers are able to handle n-dimensional discharge patterns, and it is possible to summarize all significant information in computer aided diagnosis. Such a system, published in 1989 by Kranz and Krump [9], additionally has a self-learning characteristic. Some other systems utilize a knowledge based expert system to perform diagnosis by identification of the type of the discharge and its probable location, but the experience in industrial PD testing does not favor using keyboard queries to select prefabricated decisions from a knowledge base. The most common components of an automated (on line) diagnosis system are well known from commercial pattern recognition systems:

1. data measurement, wave shaping and A/D conversion,
2. feature extraction and data reduction by mathematical algorithms, and
3. classification and formation of an identification library.

PD measurement devices have been developed from the requirements of test standards and procedures [13]. Therefore, detection bandwidth is an example of a fundamental parameter for classification among detectors, but the main subject of all commercial detectors is the PD detection sensitivity as measured in charge (pC). Diagnostic systems require the measurement of the discharge current with the correct polarity. Only ultra broadband or broadband systems can fulfill this basic need to acquire sufficient information about the individual PD pulse. It can be shown that diagnosis reliability deteriorates if the polarity is not taken into account in recording the data [1]. To ensure reliable diagnosis of low-level PD signals, it is also necessary to 'win the war against external noise'. Therefore all techniques which are known to increase the sensitivity or to reduce noise improve the diagnostic potential of any PC aided diagnosis.

A diagnosis is useful in providing additional information about the type of the discharge source. An increasing number of power apparatus should be free of partial discharges during a test cycle, and future standards will take this into account. Therefore, the question has changed from: "How much PD is tolerable?" to: "What is the reason for the measured PD activity?". Polymer components, which are especially subject to PD, must be replaced.

2. DATA PROCESSING FOR CLASSIC PHASE-RESOLVED EVALUATION

To carry out a PD diagnosis with a personal computer, all the data flow must be reduced to a manageable size. Therefore data reduction by statistical parameters (in the system published in [9], 64 KB for one measurement) is an essential feature for this type of diagnosis. The main aim is the formation of PD identification (fingerprint), which satisfactorily describes the PD characteristics associated with a particular defect. There are two common methods of data acquisition: distributions of one parameter, and histograms in the time and voltage phase domain.

It has to be taken into account that the PD phase angle, as related to other parameters, is one of the most reliable parameters for PD diagnosis. It has been proven that these histograms have the best diagnostic potential [1,3] with the exception of pulse repetition rate distributions or pulse time interval distributions. Additionally, the well known parameter, apparent charge Q or the equivalent pulse height has been used mostly to aid PD fault diagnosis, therefore it is not surprising that this is also a very important quantity in the entire identification data set [11,14].

From the author's experience, the diagnosis reliability can be increased using further PD quantities. Therefore PD energy (proportional to Q^2) and system energy (proportional to $QV(t_i)$) are significant additional parameters for evaluation [11,14]. In summary, our investigations have shown [1] that the histograms of PD impulse energy have the best diagnostic potential, which means that these histograms when used as a fingerprint in a pattern recognition scheme are most able to discriminate between different PD defects. In comparison to this, when using the system energy distributions, the diagnostic potential for all defects investigated by the author is very poor; but this parameter must be considered as one of the features (of an automated diagnosis), because this is the most useful tool to identify noise.

The parameters Q, Q^2 and QV are highly dependent on the supplied test voltage, which describes the PD hysteresis effect in a test cycle [14]. Therefore, an automated PD diagnosis must be carried out using the test voltage as an additional parameter. Taking this into account, the following features describe the discharge behavior inside the diagnostic system from which the further decisions are taken: charge or pulse height distribution, pulse energy distribution, system energy distribution, pulse interval distribution, charge or pulse height voltage phase histogram, pulse energy voltage phase histogram, and system energy voltage phase histogram.

In this investigation only, the parameters mentioned above are used to define an identification data set. From the authors point of view, all these histograms and distributions are necessary to form a fingerprint (identification data set) with sufficient information content for reliable identification [9].

The formation of this identification data set has been pointed out in [1]. The distributions and histograms are described by 18 step functions at 10 different test voltage steps, which form a N-dimensional description vector with a total length of 64 kB. Every interval (byte) of the step functions may or may not have a contribution. Every contribution can be interpreted as a part of a vector and is representing the i-th description operator of the entire PD measurement. In summary, these functions describe the characteristic fingerprint or pattern of a PD defect, which is then classified by the discussed methods. An automated diagnosis has to attach the reduced data set of the actual measurement to a fault class. Therefore, the formation of essential features like these mentioned above is a supposition using a distance classifier, but this needs much calculational effort. With a common personal computer, such a pattern recognition system cannot be used for real time analysis.

In the case of classification by an expert, further data reduction is necessary. Gulski [2] performs a data reduction using only 15 statistical features. The automated diagnosis system described in this paper needs 64000 parameters, but no expert intervention is required to achieve sufficient discriminating power. Off-line fault classification can be performed also by new methods of artificial intelligence and pattern recognition: neural networks (NN) and hidden Markov models (HMM). The on-line ability of these methods to classify PD defects, which are described by phase-resolved features, will be investigated by the author. Suzuki and Endoh [8] used a combination of phase magnitude pulse count patterns as input information for a back error propagation network (BEP) at constant test voltage. The time duration to discriminate unknown PD signals was ~ 30 s. After learning 30 typical input patterns this NN had 90% correct responses. This is significantly lower than the discrimination potential of the optimized distance classification algorithms used in the system described here, which has more than 99% correct responses with a similar calculation time.

3. MINIMUM DISTANCE CLASSIFICATION OF PD DEFECTS

The discriminating power of minimum distance classifications depends strongly on the classification method. Deterministic algorithms (see Equations (1) to (5)) to recognize a pattern are well known, while nondeterministic methods (NN) are under investigation. The degree of pattern matching in a distance classification is described by the correspondence between the actual identification data set and a stored reference data set [3]. The evaluation of the reliability is a very important component of an automated diagnosis system, but only a few algorithms are able to satisfy this basic need in PD evaluation.

The optimization of the classification algorithms depends also on the needs of the reference library, a few significant or all measurements, and influences the quality of the result. In PD testing, the test engineer gets a basic knowledge of the reliability of the automated decision by a distance classifier only! In this state of research such information can be generated particularly from a modified L_M classification. One of the major problems in using NN or nondeterministic methods is that there is no more any physical relationship between the diagnostic decision and the input data. In this field, it is impossible to calculate any reliability.

First, different methods of distance classification and reference library formation should be discussed under the aspect of PD evaluation. The first system with an automated PD classification was published by R. Krump [3]. The identification data set formed by the features mentioned above was classified by the L_1 or the Manhattan distance. The distance value for classification is determined by

$$d = \frac{1}{N}\sum_{i=1}^{N}|x_i - m_i| \quad (1)$$

where N is the number of description operators (here 64000), x_i is the i-th description operator of the actual measurement and m_i is the i-th description operator of the reference measurement.

The actual measurement is classified by the identification data set (m_i), which has the minimum distance to the measured values. The diagnostic decision indicates the best matching of the pattern related to a stored reference as in Figure 2. This Figure describes the degree of pattern matching as a normalized result from the calculation of the L_1 distance between the actual description vector and, for example, 8 stored references, which describe typical defects. The upper columns show a correct response with 40.65% correspondence to the correct reference, the lower one with 41.33% stands for the successful recognition of internal voids inside a spacer of a GIS system [1]. To achieve good classification results, the reference library should be formed by learning in the test field during routine tests [9]. It should be clear that the problem to discriminate different defects inside a GIS is much more difficult, compared to the yes or no decisions on power cables described in [4, 8].

The L_1 distance classification yields good results in the area of industrial application. It is rugged modification of the common L_M distance classifier

$$L_M = \sqrt[M]{\frac{1}{N}\sum_{i=1}^{N} w_i|x_i - m_i|^M} \quad (2)$$

By choosing $w_i = 1$, all features become rated in the same way. Therefore, by letting $M = 1$ this classification is not able to form a large distance between defects

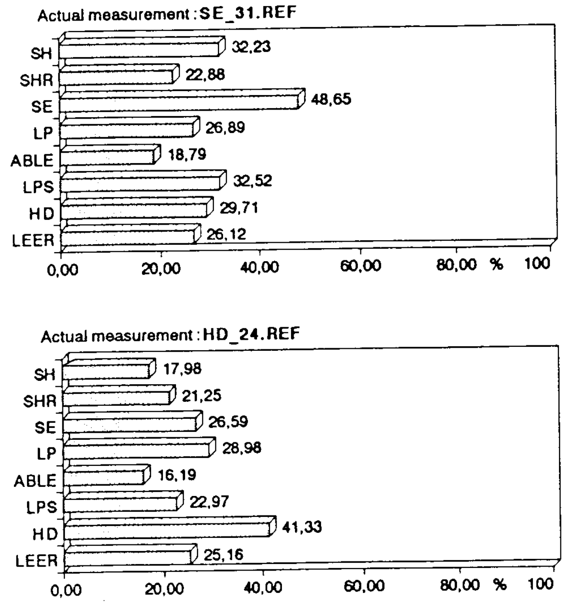

Figure 2.

Degree of pattern matching and classification with L_1 distance classifier [1]. Correspondence to the following 8 selected GIS faults: needle on high potential SH, needle (dull) on high potential SRH, needle on the ground SE, free conducting particles LP, internal voids HD, detached electrode ABLE, fixed particle on dielectric LPS, noise LEER.

of a similar characteristic, so that there might be a risk that scatter may cause a false diagnosis. But this type of classification has a certain degree of traceability, because usually the second place correspondences yield some additional physical information. It should be pointed out that the application to these 8 selected references is only an example for the classification of the two critical defects: needle on ground (SE) and internal voids (HD) (Figure 2 and 3). The realized system is able to handle more than 200 references, which can be selected in the test field. The measuring experience of the author shows that, once measured and stored in the proper library, any special defect will no longer have a misclassification.

The unmodified L_M classification usually is not able to describe a lack of resemblance to a certain defect. But an unknown discharge source can be detected if the identification factor IDAR defined in [1] will be $< 25\%$. That means, that two or more columns in Figure 2 are in the same range. The most important advantage of using L_1 classification in practice is that the result of only one measurement is needed in the reference library in order to obtain a traceable diagnostic decision, because Equation (1) can be solved with only one measurement (vector

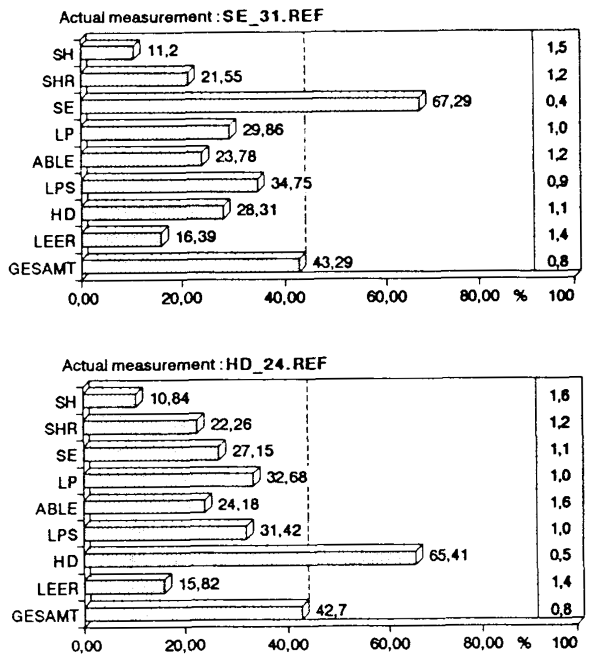

Figure 3.

Degree of pattern matching and L_σ distance classification of the same discharge sources described in Figure 2. The class GESAMT gives the information about classification reliability, it is calculated and represents no physical defect.

x) and one reference (vector m).

With an increasing number of tests and references, the distance should be calculated by Equation (2) with $w_i = 1/\sigma_i^2$ as the parameter. The statistical scatter σ should be determined by taking into account all measurements. The distance value for classification will then be

$$d = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left[\frac{x_i - m_i}{\sigma_i}\right]^2} \tag{3}$$

Figure 3 shows an improved result with a classification using the same defects and references as in Figure 2, with distance values calculated with Equation (3). The values on the right hand side yield the information about the normalized minimum distances to the next failure. Using the variance of all the measurements that are performed, it is possible to generate a nonexisting class with the name GESAMT, which describes the 'acquaintance' of the actually measured defect to the system. Only recognition rates of the actual measurement, which are (including an optional safe distance) higher than the GESAMT class, are reliable diagnostic decisions!

Usually a distance classification gives only a measure of the probability of membership within a particular class, but this for the first time defined group GESAMT enables the operator to reflect the acquaintance of the actual fingerprint and from this the information about the reliability of the automated decision is derived.

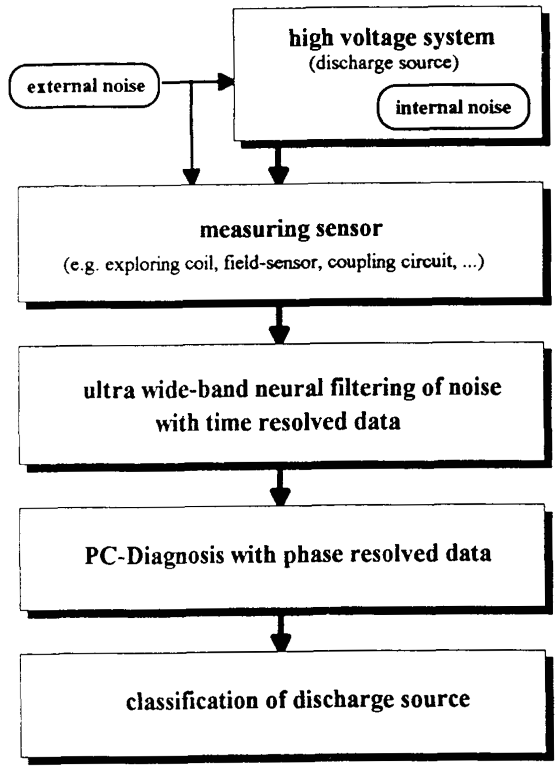

Figure 4.

Noise resistant partial discharge diagnosis, with on-line NN preclassification and off-line automated classification

A very interesting modification concerning the usage of the distance classification for partial discharge diagnosis, was published first by A. Schnettler [4]. The distance is calculated in terms of the Mahalanobis distance criterion, which is primarily used in the common bias classification

$$d = \sqrt{[\vec{x} - \vec{m}]^T C^{-1} [\vec{x} - \vec{m}]} \tag{4}$$

where $\vec{x} = [x_1, x_2, \ldots, x_N]$, C is the covariance matrix, $\vec{m} = [m_1, m_2, \ldots, m_N]$ and $[\,]^T$ denotes a transposed vector. Using this classification the result of Equation (4) can be used to calculate a degree of correspondence P between the actual measurement and a reference defect [12]. The probability for membership to a particular PD defect is always determined by

$$P = \frac{1}{\pi} \int_{-\infty}^{-d} \exp[-\frac{x^2}{2}]\, dx \qquad (5)$$

A pattern recognition corresponds with a correspondence $P > 50\%$. Usually, the measurement is characterized by a $P < 5\%$ if the evaluation describes another defect [4].

Using the minimum distance algorithms, only the L_M classification is able to determine traceable intermediate values, whereas because of the nonlinearity, the Mahalanobis-classification is only able to predict a yes or no, which is generally achievable with all nondeterministic classification methods.

From the engineering points of view, the aim is not to calculate the maximum parameter distances between two different defects; rather the aim is to ensure that the diagnostic decision remains reliable with those defects which may appear once or twice a year in a fabrication. Some classification methods, used in the field of commercial pattern recognition, are able to describe a lack of resemblance to a certain defect, but these methods need many reference vectors, usually not available in PD testing. Alternatively the evaluation with the GESAMT class only needs one reference, with the result that well known defects are classified with high probability and unknown defects could be classified with lower probability.

4. DATA PROCESSING FOR TIME-RESOLVED EVALUATION

In this area of research, there is only a little doubt that the use of NN offers the possibility of classifying the measured PD signals in real time, because a NN concept is able to handle the on-line PD data flow. NN has the ability to learn from examples. A learning algorithm performs the input to output mapping by calculating weight connections of artificial neurons. All nets adjust the parameterized weights during the learning phase, which represent the entire knowledge base.

Some NN are able to self organize the defect classification [5], but the ideal parameters for PD classification are still unknown to permit optimization of the NN learning process. The learning is usually separated from the classification. Neural information or artificial intelligence is formed by the repeated input of known and classified references. Some algorithms are able to form a reference data base by themselves during unsupervised learning. Additionally, it is possible to change the weights by adaptive learning during the classification.

NN perform an associative classification, which do not extract special features. Because of the generated weights, there is only a 'conditional traceability' between the physical process and the diagnostic decision [5]. Much effort has been expended over the last few years to provide NN for technical applications. Today the perceptron net, the Madaline algorithm, the BEP network, the Kohonen feature map (KFM), the Hopfield network and the adaptive resonance theory are enlarged on technical problems [5, 6]. All these networks have been improved in the field of pattern recognition, but they are still far removed from the ability of an expert brain. For PD applications can be summarized at this time that all the learning procedures are critically slow, the diagnostical performance is very fast, but with a narrow reliability and in practice with no traceability for the user.

One of the first applications of NN to PD fault detection has implemented a multilayered feed-forward network, using the generalized delta rule with back propagation for training, items are explained in [5]. However, it was found quite impossible to deal with real time problems, because of the training time, which was found to increase exponentially with the number of neurons and weights [6]. Phase-resolved pattern recognition requires a large number of neurons. It is shown in [8] that even the use of $320\times20\times3$ neurons (BEP) gives only 90% correct responses. Therefore, it can be concluded that the BEP-NN constitutes only a 'promising approach in developing an intelligent alarm processor for discrimination of partial discharge signals from noise', [8]. All of the mentioned nets have been investigated by the author for classification of phase-resolved data. The aim of this investigation is to obtain more understanding about the number of neurons and layers required, as well as to optimize the number of classes and the traceability of the decisions.

In this paper the first results are presented using the BEP network and the KFM for time-resolved discharge classification. From the authors' point of view, the artificial NN designed only for noise identification represents the best approach to reduce the number of classes and neurons to achieve real time recognition ability. To face this on-line problem, only an ultra broadband measuring system is able to create the needed information. However it is impossible to reduce the data flow by statistical tools. Therefore, the first step is to optimize the number of input neurons.

In this investigation the discharge current was measured by a Tektronix DSA 602 with 2 G samples/s in the time domain. The PD data had to be stored with a high speed A/D converter forming a n = 1000 dimensional vector with nearly continuous values. An optional fast feature extraction (*e.g.* correlation filter) may reduce this vector to m dimensions ($m < n$). A taught NN fed by such a m-dimensional input vector is able to determine an output vector, which can be related to a particular discharge source.

In this investigation the diagnostic decision remains reliable, only if every 50th contribution ($m = 20$) of the input vector was processed. Therefore the BEP network was finally implemented with $20 \times 10 \times 5$ neurons, and the KFM with $20 \times 12 \times 12$ neurons. The BEP network performs diagnosis by the so-called maximum activity (evaluated in percentage) of one of the output neurons, which represents the references or fault classes. The KFM performs the diagnosis with a minimum distance area in the 12×12 output field (marked area in the two dimensional feature map on Figure 5, lowest row).

In using the shape of the discharge pulse as a characteristic parameter of a neural filter, it is not necessary to know the true pulse shape. The bandwidth of the detection system should be in the range of 100 MHz, particularly if one wishes to discriminate the pulse shape of a discharge event as for example that which may characterize the PD in a GIS system. The observed plots of the PD signals measured in this investigation are shown in the upper row of Figure 5.

In this work, PD fault detection is performed by a two step strategy shown in Figure 4:

1. Time-resolved online classification of PD signals and noise using the sampling data of the measured signal processed by a fast neural filter with an unavoidable but tolerable misclassification.
2. Phase-resolved off-line diagnosis with high reliability and traceability, which should be robust against that remaining influence of noise [3].

5. CLASSIFICATION BY NN

The scanned values of the measured PD signals in Figure 5(a) to (c) also can be interpreted as an input pattern, which of course can be classified off line better by any expert or distance classification. Therefore, off-line neural PD classification at this stage of our knowledge has no reasonable advantage. But due to the possibility of developing an efficient NN in hardware there is the feasible option for a real time preclassification. This preclassification by example of two NN is described in Figure 5. The optimized neural filter for PD measurements should preferably have only two output classes, one for the PD signal, the other for stochastic impulse noise. This is a basic need in saving computer time, but a slight increase of the number of output classes offers the possibility of discriminating well known local disturbances, *e.g.* classes 4 and 5 in the BEP output on Figure 5.

The diagnostic potential of all investigated NN was proved by learning typical PD signal shapes. After this, these measured signals were manipulated by noise and time delays. Figure 5(c) and (d) show the results from this investigation for only one example: the PD defect needle on high potential SH. This Figure also shows an off-line diagnosis of the following discharge impulses from a gas insulated switchgear: LEER, stochastic noise taught *e.g.* to BEP neuron 2 with class #2 (Figure 5(a)), LPS conducting particle fixed on a dielectric taught to BEP neuron 3 in Figure 5(b), SH needle on high potential (Figure 5(c)).

Using the shape of the discharge pulse as a characteristic, it is not necessary to process the time-resolved data, with the resolving power shown in the oscillograms of Figure 5. Every column (a) to (e) of this Figure has three components. The upper one shows the NN input signal and under that the diagnostic decisions (output) of two different NN are plotted. The signals (a), (b) and (c) are actually measured. In comparison with this, the data of the signals (d) and (e) are manipulated to investigate the influence of noise and trigger delays on signal (c) and the reliability of the algorithms.

It can be demonstrated clearly (Figure 5), that both neural networks (BEP and KFM) are able to determine an uniqueness of the off-line diagnostic decision with a minimum of 20 input neurons. It is also well known that white noise and trigger delays are influencing diagnosis reliability of time resolved data generally, but as has also been confirmed by this PD investigation NN are influenced in a tolerable range, because for example in Figure 5(d) a PD impulse from a needle on high potential superimposed on 20% white noise, will be recognized by both networks unaffected by this noise. Compared to this, the trigger delay creates more difficulty (Figure 5(e)), but also without any misclassification. If the difference between two evaluation columns or normalized neuron activities is $> 25\%$ *e.g.* Figure 5(e), BEP output, the diagnostic decision can be estimated to be reliable, by using the NN as an intelligent filter.

6. CONCLUSIONS

A combination of minimum distance and NN classification for fault and noise identification in PD diagnosis is proposed. Such methods complement standardized measuring methods and conventional PD evaluation

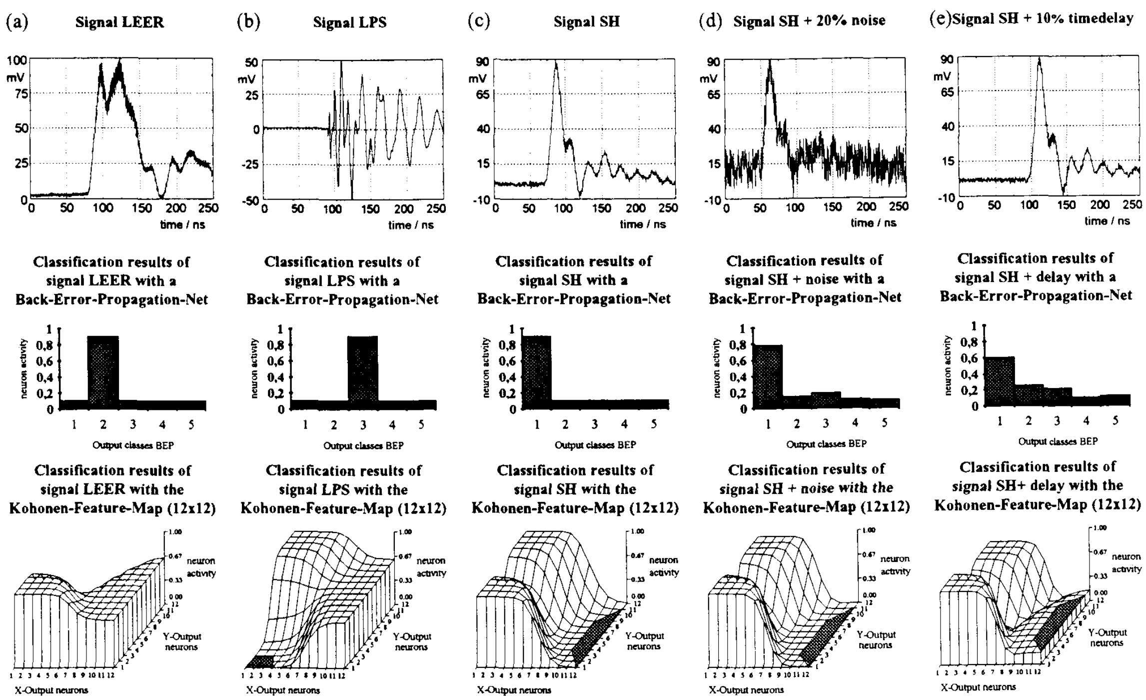

Figure 5.

Off-line PD classification with NN. (a) LEER stochastic noise impulse, (b) LPS-PD signal from a conducting particle fixed on a dielectric, (c) to (e) SH-PD signal from needle on high potential, Back Error Propagation Net: 20×10×5 BEP network, Kohonen Feature Map: 20×(12×12) KFM. Upper row: time-resolved NN input signals, below NN diagnosis, Middle row: Diagnostic decision of the BEP output, respectively activated output neuron, percentage of neuron or class activity. Lower row: Diagnostic decision of the feature map, respectively activated output area.

schemes. The reliability of the distance classification and the neural filter is currently being investigated in field tests. The following conclusions can be made from the experimental results.

1. All minimum distance classifications need an essential data reduction to form the characteristic fingerprint of a PD defect during a more complex calculation.
2. Only the L_M distance is able to determine reasonable intermediate values and quantities for the examination of the diagnostic decision or the evaluation of the diagnostic reliability.
3. The bias- and NN classifications still have no or only a conditional traceability.
4. Using broadband measuring systems, NN provide the potential for real time identification of pulse shaped noise and PD impulses.
5. Simple but fast NN are able to discriminate significantly different measuring signals with sufficient redundance against signal deformation.
6. The transfer of the experiences with time-resolved laboratory measurements evaluated by NN to practical on-site applications has to be proved (present research) systematically by extensive investigations.

The new method for calculating the acquaintance of a PD measurement, using a modified L_M classification is proposed as the best parameter to obtain sufficient information about the diagnosis reliability, which can be additionally improved if the input data are evaluated by a neural filter. All evaluation methods must take into account that most currently manufactured power apparatus has a low PD fault rate; therefore, the PD diagnosis should be able to handle at least one defect with only one reference but with high reliability. This requires a special interpretation of pattern analysis and machine intelligence in partial discharge diagnosis.

ACKNOWLEDGMENT

The author wishes to express his gratitude to Dipl.-Ing. T. Hücker and Dipl.-Ing. A. Groß for their contributions to this approach. The financial support of the Deutsche Forschungsgemeinschaft DFG is greatly acknowledged.

REFERENCES

[1] H.-G. Kranz, R. Krump, "Partial Discharge Diagnosis using Statistical Parameter Optimization on a PC Based System", IEEE Transactions on Electrical Insulation, Vol. 27, pp. 93–98, 1992.

[2] E. Gulski, *Computer-aided Recognition of Partial Discharges using Statistical Tools*, The Hague, Delft University Press, pp. 150–158, 1991.

[3] R. Krump, *Ein Störsignalresistentes Verfahren Zur Computergestützten Teilentladungs Diagnostik in SF_6-isolierten Schaltanlagen*, Thesis, Bergische Universität Wuppertal, 1989.

[4] A. Schnettler, D. Peier, "Rechnergestützte Teilentladungsdiagnose an Mittelspannungs VPE-Kabeln", ETG Fachbericht, VDE Verlag, Nr. 40, pp. 327–332, 1992.

[5] S. Grossberg, *Neural Networks and Natural Intelligence*, MIT Press, Cambridge, London 1988.

[6] H. Ogi, H. Tanaka, "Module Neural Network Application for Power System/Equipment Diagnosis", Neuro Conference, Nime France, 1991.

[7] G. Vaillancourt, R. Malewski, "Digital Acquisation and Processing of Partial Discharges During Acceptance Tests of HV-Transformers", Paper 88WM049-9, IEEE/PES Winter Meeting 1988.

[8] H. Suzuki, T. Endoh, "Pattern Recognition of Partial Discharges in XLPE Cables Using a Neural Network", IEEE Transactions on Electrical Insulation, Vol. 27, pp. 543–549, 1992.

[9] H.-G. Kranz, R. Krump, "The Abilities of Self Operating Expert Systems for Statistical Partial Discharge Analysis of GIS-Test Signals", 6th Intern. Symp. HV Engineering New Orleans LA, 22.13, 1989.

[10] R. J. Van Brunt, S. V. Kulkarni, "Method for Measuring the Stochastic Properties of Corona and Partial Discharge Pulses", Rev. Sci. Instrum., Vol. 60, No 9, pp. 3012–3023, 1989.

[11] S. A. Boggs, "Partial Discharges - Part III: Cavity-induced PD in Solid Dielectrics", IEEE Elec. Insulation Magazine, Vol. 6, pp. 11–20, 1990.

[12] H. C. Andrews, *Introduction to Mathematical Techniques in Pattern Recognition*, Wiley-Interscience, p. 113, New York 1972.

[13] IEC Publication 270, *Partial Discharge Measurements*, second edition, 1981.

[14] S. A. Boggs, G. C. Stone, *Fundamental Limitations in the Measurement of Corona and Partial Discharges*, IEEE Transaction on Electrical Insulation, Vol. 19, No 2, pp. 143–150, 1982.

This paper is based on a presentation given at the 1992 Volta Colloquium on Partial Discharge Measurements, Como, Italy, 26–28 August 1992.

Manuscript was received on 16 February 1993, in final form 23 September 1993.