Determination of Voltage Level from Audible DC Corona Noise by Using Neural Network

Suna Bolat Sert[1], Özcan Kalenderli[2]

[1, 2] Electrical Engineering Department, Istanbul Technical University, 34469 Maslak-Istanbul, Turkey
bolats@elk.itu.edu.tr, ozcan@elk.itu.edu.tr

Abstract

In this study, a signal recognition application is presented for classification of positive DC corona voltage level using recorded sound data of corona (electrical discharge) and utilizing probabilistic neural network (NN). The recorded positive DC corona sound data are acquired experimentally from a test set-up intentionally producing corona sound by applying different levels of positive DC high-voltage. Recordings have been used in training and test sets of the neural network. The main objective of this study is developing a model to determine source voltage level by only analyzing the recorded corona sound data. During the application of algorithmic method, linear prediction coefficients are used to pre-process the sound data for feature extraction. It is shown that reasonable results can be obtained by following the proposed method.

1. Introduction

Corona is a partial electrical discharge occurring at the points of highly concentrated electric fields. It is a self-sustained electrical gas discharge which consists of a high-field active electrode surrounded by an ionization region where the free charges are produced. Corona discharge can be characterized by electric current, energy loss, radio interference, mechanical vibrations, chemical reactions, visible light, and audible noise [1-4]. This study, mainly concerns about audible noise of a positive DC corona.

Acoustic methods can be used for discharge detection and location, fault detection, diagnostics and long term system monitoring and evaluation studies. Using sound of a discharge instead of voltage of the discharge is a very important progress for measurement techniques especially for high voltages which are very difficult to measure directly [5].

The electrical discharge depends on many variables such as, medium temperature, pressure, humidity, material type, dimensions, geometry, homogeneity, duration and type of the applied voltage [1-2]. For that reason, using audible noise of DC corona in order to measure the voltage level which generates the corona is usually not considered as a dependable way. However, it is now considered as an interesting measurement technique by the developments of computer skills and improvements on the signal processing, measurement and evaluation techniques.

The measurement of voltage is essential to power system control, protection and monitoring. Traditional measurement techniques require contact with the point of measurement. By the developments of technology, many new methods are being developed in order to measure line voltage. This study presents a new approach for voltage measurements without any contact with the line. This method is cost effective and less complex in terms of installation maintenance than the traditional measuring methods [6].

In this study, by means of probabilistic neural networks (PNN) a new approach has is presented that the sound recordings of positive DC corona (electrical discharge) are used to determine the voltage level, which causes the electrical discharge [7-8]. The data that are used to classify the voltage from audible DC corona noise using the neural networks have been collected from a high-voltage line model installed in a high-voltage laboratory. In order to apply the sound recordings to the neural network, a signal processing routine should be applied to the sound data. This procedure called as feature extraction, which is preprocessing the original data. This routine is used for simplify the data when the problem has much data but not much information. In order to pre-process the sound data, it has been analyzed by linear predictive coding (LPC) technique.

2. Feature Extraction

Feature extraction involves simplifying the amount of resources required to describe a large set of data accurately. When performing analysis of complex data one of the major problems stems from the number of variables involved. Analysis with a large number of variables generally requires a large amount of memory and computation power or a classification algorithm which overfits the training sample and generalizes poorly to new samples. Feature extraction is a general term for methods of constructing combinations of the variables to get around these problems while still describing the data with sufficient accuracy.

Common denominator of all recognition systems is the signal processing front-end, which converts the acoustic waveform to some of type of parametric representation. This parametric representation is then used for further analysis and processing. In this study, linear predictive coding (LPC) is used for feature extraction of the DC corona sound data.

2.1. Linear Predictive Coding (LPC)

Linear predictive coding (LPC) is a tool used for representing the spectral envelope of a digital signal in compressed form, using the information of a linear predictive model. It is one of the most useful methods for encoding good quality sound at a low bit rate and provides extremely accurate estimates of sound parameters. It is a way to obtain a smooth approximation of the sound spectrum. The objective of this method is to design a filter which resembles the spectrum of the signal that is desired to obtain frequency response [9]. The spectrum is modeled with an all-pole function, which concentrates on spectral peaks. In this study, LPC is used for feature extraction

In the classical forward linear prediction, an estimate for the next sample $\hat{y}(n)$ of a linear discrete-time system is obtained as a linear combination of p previous output samples. The predicted signal value can be expressed as,

$$\hat{y}(n)=\sum_{i=1}^{p} a_i y(n-i) \tag{1}$$

where a_i denotes the linear predictor (LP) coefficients. They are fixed coefficients of a predictor all-pole filter, whose transfer function is

$$H(z)=\frac{1}{A(z)}=1\Big/\left(1-\sum_{i=1}^{p} a_i z^{-i}\right) \tag{2}$$

The goal of the linear prediction is to find the set of the linear predictor coefficients $\{a_1, a_2, ..., a_p\}$ that minimize the short-time mean-squared prediction error.

$$e=E\left\{\left|y(n)-\sum_{i=1}^{p} a_i y(n-i)\right|^2\right\}\approx\sum_{n=-\infty}^{\infty}\left|y(n)-\sum_{i=1}^{p} a_i y(n-i)\right|^2 \tag{3}$$

where $E\{.\}$ is the expected value. By definition, e is also the prediction error power [10]. The most common choice in optimization of parameters a_i is the root mean square (RMS) criterion which is also called autocorrelation. In this method the expected value of the square error is minimized. To solve the minimization problem, the Levinson-Durbin algorithm is used in this study. Detailed explanations can be found in [11-13].

3. Probabilistic Neural Networks (PNN)

Probabilistic neural networks can be used for classification problems. When an input is presented, the first layer computes distances from the input vector to the training input vectors and produces a vector whose elements indicate how close the input is to a training input. The second layer sums these contributions for each class of inputs to produce as its net output a vector of probabilities. Finally, a transfer function on the output of the second layer picks the maximum of these probabilities, and produces a 1 for that class and a 0 for the other classes.

Considering a pattern vector x with m dimensions that belongs to one of two categories K_1 and K_2, let $F_1(x)$ and $F_2(x)$ be the probability density functions (pdf) for the classification categories K_1 and K_2, respectively. From Bayes' decision rule, x belongs to K_1 if (4) is true, or belongs to K_2 if (4) is false;

$$\frac{F_1(x)}{F_2(x)}>\frac{L_1P_2}{L_2P_1} \tag{4}$$

where L_1 is the loss or cost function associated with misclassifying the vector as belonging to the category K_1 while it belongs to the category K_2, L_2 is the loss function associated with misclassifying the vector as belonging to the category K_2 while it belongs to the category K_1, P_1 is the prior probability of occurrence of the category K_1, and P_2 is the prior probability of occurrence of the category K_2. In many situations, the loss functions and the prior probabilities can be considered as equal. Hence the key for using the decision rule given by (4) is to estimate the probability density functions from the training patterns [14].

In the PNN, a nonparametric estimation technique known as Parzen windows [15] is used to construct the class-dependent probability density functions for each classification category required by Bayes' theory. This allows determination of the probability a given vector pattern lies within a given category. Combining this with the relative frequency of each category, the PNN selects the most likely category for the given pattern vector. Both Bayes' theory and Parzen windows, which are theoretically well established, have been in use for decades in many engineering applications, and are treated at length in a variety of statistical textbooks. If the j[th] training pattern for the category K_1 is $\boldsymbol{x}_j$, then the Parzen estimate of the pdf for the category K_1 is

$$F(x)=\frac{1}{(2\pi)^{m/2}\sigma^m n}\sum_{i=1}^{n}\exp\left[-\frac{(\boldsymbol{x}-\boldsymbol{x}_j)^T(\boldsymbol{x}-\boldsymbol{x}_j)}{2\sigma^2}\right] \tag{5}$$

where n is the number of training patterns, m is the input space dimension, j is the pattern number, and σ is an adjustable smoothing parameter [14]. Figure 1 shows the basic architecture of the PNN.

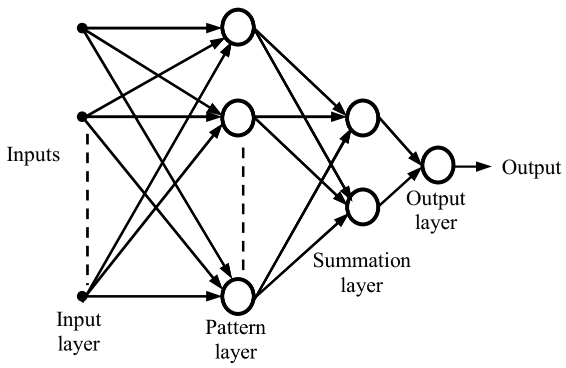

Fig. 1. The basic architecture of the PNN.

The first layer is the input layer, which represents the m input variables ($\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_m$). The input neurons merely distribute all of the variables $\boldsymbol{x}$ to all neurons in the second layer. The pattern layer is fully connected to the input layer, with one neuron for each pattern in the training set. The weight values of the neurons in this layer are set equal to the different training patterns. The summation of the exponential term in equation (5) is carried out by the summation layer neurons. There is one summation layer neuron for each category. The weights on the connections to the summation layer are fixed at unity so that the summation layer simply adds the outputs from the pattern layer neurons. Each neuron in the summation layer sums the output from the pattern layer neurons, which correspond to the category from which the training pattern was selected. The output layer neuron produces a binary output value corresponding to the highest pdf given by equation (5). This indicates the best classification for that pattern [15].

This case is a binary decision problem. Therefore, the output layer has just one neuron and summation layer has two neurons [15-19].

4. Experimental Study

In this study, high voltage transmission line set up is installed in a high voltage laboratory in order to produce corona. In the transmission line model 5 m long having circular cross-section area of 2.5 mm^2, smooth, clean, dry copper wire was laid at a height of 220 cm above from the ground between two support insulators as a part of the experimental setup. The simple diagram of the experimental setup is shown in Figure 2. Seven different positive DC voltages which are 40 kV, 45 kV, 50 kV, 55 kV, 60 kV, 65 kV and 70 kV were applied to the wire, respectively [7-8].

The experiment was repeated for each above given voltage levels and the corona sound was recorded throughout 180 seconds in each test. The microphone was located at 1.5 m above from the ground and at a horizontal distance of 1 m away from the midpoint of the wire.

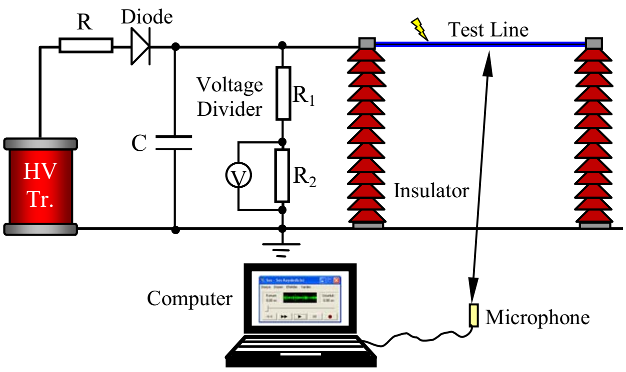

Fig. 2. Experimental Setup

The positive DC corona discharge sound was recorded for each voltage level via a capacitive microphone. The microphone is a part of a sound level meter. The similar studies can be carried out by using sound pressure levels instead of sound recordings. The microphone of the sound level meter used in this study has 1.25 cm diameter.

The experiment was carried out in electromagnetic shielded laboratory at the air pressure of 1019 Pa (765 mmHg) and room temperature of 19 °C at a relatively quiet ambient. Probabilistic neural network was used for the classification of the corona sound in terms of the applied DC voltage [7-8].

Theoretical and experimental studies show that, electric field increases for small curvature radii and it has a highest magnitude for a point. A strong electric field will cause ionization of air molecules to occur. This ionization generates audible noise.

Corona discharges appear at a lower magnitude for negative voltages than for positive voltages. This is because the secondary processes that occur at the cathode for negative corona are delayed for positive corona, as the cathode is isolated from the ionization region by the drift region [6]. For that reason, positive DC voltages are applied to the wire in this study.

5. Application

As explained before, electrical discharge (corona) sound samples at each voltage level are acquired for the duration of 180 seconds. The sound samples which have 22050 Hz sampling frequency and resolution of 16 bits are recorded by using MATLAB packet program.

First 22050 data of the sound sample which represents one second at 50 kV DC voltage is shown in Figure 3. As shown in Figure 3, each recording from seven different DC voltage levels has an impulse noise at the beginning of the signal. In order to clear the sound signal from this noise, the first one second of the sound signal recorded for seven different voltage levels have been removed individually.

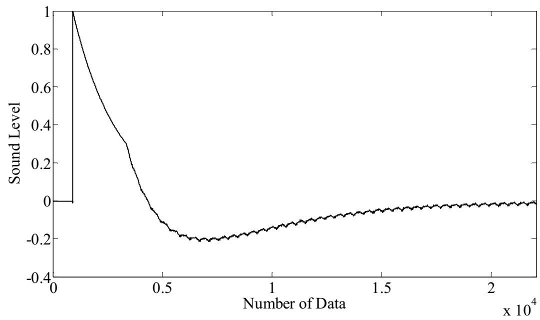

Fig. 3. The first 22050 data of the DC corona sound (1 second) at 50 kV voltage level

The linear prediction coefficients are computed for each frame having the length of 20 ms during a second on the recorded sound samples and those frames are not overlapped. After LP coefficients are computed for each frame, the average value of all LP coefficients is computed. Each one-second sound sample is represented by its average LP coefficient value. 20^{th} degree LP coefficients are computed for all the recordings. Therefore, there are 179 data for each voltage level. Since there are seven different applied voltage levels, per second for 179 seconds produces 1253 sound samples and the LP coefficients. The training and test sets for the PNN are obtained from these 1253 data. The first 840 LP data which are computed from the first 120 seconds of sound data recorded at each voltage level are used as the training set; the remaining data are used as the test set. Following that procedure, the analyzed noisy sound data is applied to a probabilistic neural network and classification percentages are determined. The input of the PNN is the LP coefficients of sound data and the output vector of the PNN is the voltage magnitude at which the sound samples are recorded. Input vectors have 20 neurons which are degree of the LP coefficients, output vectors has only one neuron which is the voltage level.

The PNN calculates the probability of belonging of a vector $\boldsymbol{x}$, corona sound to all classes at the summation layer output, P_i (voltage level_i|corona sound). Classes are the voltage levels. After summation layer, the PNN decides to which voltage level the corona sound sample belongs by using Bayes' decision rule in the output layer.

The only parameter that has been used in the PNN is spread value, given by σ. The optimum spread value of the PNN is determined by using trial-by-error process in the studies. The σ value is the spread of radial basis functions used in the PNN. If spread is near zero, the network acts as a nearest neighbor classifier. As spread becomes larger, the designed network takes into account several nearby design vectors.

In Table 1, change of misclassification with respect to the spread value for the measured sound signal for different seven positive DC voltages is shown. All the recorded data is used in the classification problem in order to determine the DC voltage level of a sound sample.

The training set of the PNN contains 840 patterns of the LP coefficients which consist of 120 sound samples for seven voltage levels (120 x 7 = 840). The test set of the PNN contains 413 patterns of the LP coefficients which consist of 59 sound samples for seven voltage levels (59 x 7 = 413).

The voltage levels representing the classes in the training set and the test set are discrete. All inputs are classified into those defined seven classes. Because it is not continuous, the sound samples can be classified only into the classes that have been defined.

Table 1. The change of misclassification with respect to σ for the measured sound signal

	Misclassification of Training Set		Misclassification of Test Set	
σ	**Number**	**[%]**	**Number**	**[%]**
0.05	65	7.738	17	4.116
0.04	44	5.238	18	4.358
0.03	22	2.619	16	3.874
0.02	0	0	10	2.421
0.0155	0	0	10	2.421
0.015	0	0	10	2.421
0.01	0	0	12	2.905

The results of test set in Table 1 can also be analyzed for each of the voltage levels separately (Table 2). The classification of the test sets has been given in Table 2 separately for each of seven voltage levels where 59 data corresponding to one voltage level.

Table 2. Number and percentage of misclassifications for the test set at each voltage level.

	Number of Misclassifications						
σ	**40 kV**	**45 kV**	**50 kV**	**55 kV**	**60 kV**	**65 kV**	**70 kV**
0.05	2	0	1	3	8	3	0
0.04	2	0	1	3	8	3	1
0.03	1	0	1	2	7	4	1
0.02	1	1	0	1	4	2	1
0.0155	1	1	0	1	5	1	1
0.015	1	1	0	1	5	1	1
0.01	1	1	0	1	6	1	2
σ	**Percentage of Misclassifications [%]**						
0.05	3.390	0.000	1.695	5.085	13.559	5.085	0.000
0.04	3.390	0.000	1.695	5.085	13.559	5.085	1.695
0.03	1.695	0.000	1.695	3.390	11.864	6.780	1.695
0.02	1.695	1.695	0.000	1.695	6.780	3.390	1.695
0.0155	1.695	1.695	0.000	1.695	8.475	1.695	1.695
0.015	1.695	1.695	0.000	1.695	8.475	1.695	1.695
0.01	1.695	1.695	0.000	1.695	10.169	1.695	3.390

The results in Table 1 and Table 2 show that as the spread value decreases, and the number of misclassifications in the training set decreases, too. However, the number of misclassifications in the test set does not show a significant change. The sound data in the training set are classified perfectly without any error, the sound data in the test set are classified with an acceptable error of 2.42%. If the sigma value is smaller than 0.02, the model will over fit the data, because each training point will have too much influence. The accuracy of the measurements would need to improve if the technique was to find commercial usage in substations. Environmental noises would affect the performance of the network. The network might not be able to classify the voltage levels correctly in a noisy environment, therefore in order to improve the results, the sound samples can be cleaned by applying several signal processing procedure.

6. Conclusions

In this study, identification of the positive DC corona sound and classification of the DC voltage by using the sound data produced at that voltage level are presented. For that purpose, the sound data is analyzed by the linear predictive coding. The processed sound information is applied to the probabilistic neural network. Consequently, the magnitude of the DC voltage is classified by using the recorded corona sound signals.

Performance of the identification depends on perception and recording of the sound data. Corona discharges are significantly affected by atmospheric conditions and discharge sound heavily depends on the geometry, dimensions, cleanliness, dryness, and smoothness of the structure in which the discharge appears. For that reasons, as the diameter of the wire is small and the magnitude of voltage is high, electrical stress on the conductor is bigger than breakdown field of air surrounding the wire, which leads to a discharge phenomena.

When the input data to an algorithm is too large to be processed and it is suspected to be notoriously redundant then the input data will be transformed into a reduced representation set of features. Transforming the input data into the set of features is called features extraction. If the features extracted are carefully chosen it is expected that the features set will extract the relevant information from the input data in order to perform the desired task using this reduced representation instead of the full size input. In order to pre-process the sound data, it is possible to use different methods. In this study, linear predictive coding is preferred. It has been seen that the results are appropriate for the data but different methods can be applied to the problem and compared to each other in order to determine the optimum method. It is obvious that algorithmic methods such as fuzzy logic and genetic algorithms can be also used for this study.

Although the classification error of the test set is 2.42 % the accuracy of the measurements would need to improve. In order to improve the results, the sound samples can be cleaned by applying several signal processing procedure. The accuracy of measurements may improve with investments in new, calibrated equipment.

This study is a minor step for measuring the voltage magnitude from the sound data using signal processing techniques. It is an application for future studies of monitoring and evaluation in electrical systems and remote sensing and forecasting voltage magnitude using the corona sound that appears during a line fault without any physical connection to the line especially at high voltage levels.

7. References

[1] M. Khalifa, "High Voltage Engineering", Marcel Dekker, New York, 1990.

[2] Ö. Kalenderli, "Analysis of Partial Discharge Reference Source under High Direct Voltage Conditions", Ph.D. Thesis, Istanbul Tech. Univ., Ins. of Science and Technology, Istanbul, 1991.

[3] M. C. Taplamacioglu, "Electric Potential and Field Modification by Corona in High Voltage Systems", Ph.D. Thesis, University of Wales College of Cardiff, 1993.

[4] A. Akses, Ö. Kalenderli, "Electromagnetic characteristics of high voltage DC corona", *The 2003 IEEE International Symposium on Electromagnetic Compatibility (EMC)*, Istanbul, Turkey, May 11-16, 2003

[5] Y. Zebboudj, "Measurements of the current and electric

field distributions beneath a positive DC wire-to-plane corona using a linear biased probe*", Proc.-Sci. Meas. Technol*., Vol. 147, No.2, pp. 74-80, 2000.

[6] P. J. Moore, D. L. Hickey, M. G. G. Urbaneja, "Remote sensing of voltage using optical assessment of corona", *IEEE Power Engineering Society Summer Meeting*, Vol. 2, pp. 1159-1164, 2000.

[7] S. Bolat Sert, Ö. Kalenderli, "Determination of source voltage from audible corona noise by neural networks", *IEEE Transactions on Dielectrics and Electrical Insulation,* Vol. 16, No. 1, pp. 224-231, February 2009.

[8] S. Bolat Sert, Ö. Kalenderli, "Determination of corona location from audible corona noise by using artificial neural network", *International Symposium on Innovations in Intelligent Systems and Applications* (INISTA 2009), June 29-July 1, 2009 Trabzon, Turkey (Accepted to publication).

[9] A**.** Eronen, "Automatic Musical Instrument Recognition", MS thesis, Tampere University of Technology, 2001.

[10] L. Rabiner, B. Juang, "Fundamentals of Speech Recognition", Prentice Hall, New York, 1993.

[11] J. O. Smith, "Viewpoints on the history of digital synthesis", Proc. ICMC, Helsinki, 1991.

[12] J.A. Moorer, "Signal processing aspects of computer music: A survey", *Proc. IEEE*, Vol. 65, pp. 1108-1137, 1977.

[13] M. Lenz, K. Feser, K., et al, "Wavelet transform application during impulse tests of power transformers", *13th Int. Symp. on High Voltage Engineering*, Netherlands, 2003.

[14] T.C. Goh, "Probabilistic neural network for evaluating seismic liquefaction potential", *IEEE Int. Sympos. Intelligent Systems*, Varna, Bulgaria, pp. 16-20, 2002.

[14] E. Parzen, "On estimation of a probability density function and mode", *Annals of Math. Statistics*, Vol. 33, pp. 1065 – 1076, 1962.

[15] R. Senani, "New Current-Mode Biquad Filter", *Intern. J. Electronics*, Vol. 73, pp. 735-742, 1992.

[16] R. Bozzo, L. Sciutto and R. Zunino, "BP networks for partial discharge analysis and dielectric classification", *Intern. Conf. Neural Networks*, Vol. 1, pp. 37 – 42, 1997.

[17] S. Haykin, "Adaptive Filter Theory", Prentice Hall, NY, 1991.

[18] M.H. Hayes, "Statistical Digital Signal Processing and Modelling", John Wiley & Sons, NY, 1996.

[19] B. Bolat and T. Yıldırım, "Active Learning for Probabilistic Neural Networks", Lecture Notes in Computer Science, Vol. 3610, pp. 110-118, 2005.