

## Remote Detection and Recognition of Electrostatic Discharge from HVDC Transmission Lines

Yue Zhang\*, Shanghe Liu, and Xiaofeng Hu

**Abstract**—To remotely detect corona discharge from High-Voltage Direct Current (HVDC) transmission lines, a detecting system combining detecting platform and data progressing system is designed. Detecting platform is developed resorting to the principle of differential noise reduction, which can fulfill narrow-band detection breaking away interference from broadcasting and easily catch the electrostatic discharge signal. To get rid of interference from spark discharge, a data progressing system containing feature extractions, clustering and recognition technologies is developed. Clustering is realized by extracting five discharge features, including peak factor, form factor, skewness, kurtosis and mean square error. The unsupervised clustering Fuzzy C-Means (FCM) method is used to achieve fast separation for electrostatic discharges and provide training set for pattern recognition. Pattern recognition resorts to Support Vector Machine (SVM) method. For comparison, Back Propagation (BP) and Learning Vector Quantization (LVQ) approaches are taken to test the recognition ability. The results show that SVM recognizer with a recognition rate of 97.5% achieves higher performance than BP and LVQ methods. It can be concluded that the detecting system can be an interesting alternative for electrostatic discharge detection.

### 1. INTRODUCTION

As transmission voltages of High-Voltage Alternating Current (HVAC) as well as HVDC transmission lines increase, corona discharge becomes a design factor, not only because of corona losses and their impact on efficiency of power transmission, but also due to the electromagnetic interference (EMI) caused to broadcast [1–5]. Moreover, the corona-generated ions produced by high-voltage transmission lines may affect health by enhancing exposure to aerosol pollutants [6]. So real time detection technology is desired. For real time detection, a stationary detecting device is needed. However, HVDC transmission lines have the characteristics of large tower size, long span, complex geographic locations and weather conditions. Close range detection will be a cost of heavy work and difficult to be operated in practice, especially in high-altitude and serious icing regions. Therefore, the necessity of researching electromagnetic radiation characteristics of corona discharge and related technologies, particularly remote detecting and recognition methods becomes pressing.

A great number of EMI detections have been made over the past 50 years, primarily in the AM broadcast [1, 7, 8]. Some measurements have also been made in the TV band. In 1986, Chartier et al. detected corona interference up to 800 ~ 1000 MHz by using a high gain parabolic antenna and low-noise preamplifier [9]. Recently, large number of researches have been employed to the transmission lines detecting applications. Detecting methods of ion-current [10] and electric field underneath the transmission lines were researched [11]. And measurements of corona effects were carried out based on the discharge phenomenon, including corona loss (CL) [12], radio interference (RI) [13], audible noise

---

*Received 29 July 2014, Accepted 7 October 2014, Scheduled 4 November 2014*

\* Corresponding author: Yue Zhang (L\_forty@163.com).

The authors are with the Institute of Electrostatic and Electromagnetic Protection, Mechanical Engineering College, Shijiazhuang 050003, China.

(AN) and ion current distribution [14]. But little published work on remote detection was done for electrostatic discharge from HVDC transmission lines. CISPR TR 18-2 gives on-site measurements on HV overhead power lines in the frequency range from 0 to 300 MHz to determine limits for the radio noise produced by HV power lines. The suggested distance is up to 100 m, beyond which the noise level is normally so low that reliable measurements are not practicable [4]. In addition, there are some researches on non-electricity detection, such as acoustic measurements, ultraviolet and visible light (UV) imaging devices [15], but they are not suitable for remote detection owing to the low sensitivity.

For the discharge signal processing, there is much knowledge about the analysis of the partial discharge (PD) source and measurement systems, especially high-voltage equipments. Most of the research was done by Fromm [16–18], Morshuis et al. [19] and Gulski et al. [20]. Fromm has built a model with parameters including discharge magnitude  $q_i$ , time of discharge interval  $t_i$ , and instantaneous voltage across the test object, which describes the stochastic discharge process at dc voltage. Based on the Fromm's work, Morshuis and his collaborators presented the detection and analysis of PD with experimental proofs. In Gulski's research, the main parameters are phase angle  $\phi$ , discharge magnitude  $q$  and frequency  $n$  for PD at ac voltage. According to the two classic PD models, many detection and pattern recognition technologies were researched [21–24]. Based on which, Wang identified the partial-discharge patterns of high-voltage devices with the recognition method extension neural network (ENN) [22]. And Si et al. applied Fromm's model to evaluate dielectric insulation and classify discharge signal from high-voltage equipments, using the artificial intelligence methods FCM and least square support vector machine (LS-SVM) [24]. But for corona discharge from HVDC transmission lines, it is stochastic without regular phase information, and the discharge interval is up to 1 ms, whose discharging number is too small compared with Fromm's and Gulski's model. So the feature evaluation parameters in Fromm and Gulski's models cannot produce expected results.

A remote detecting system of electrostatic discharge, based on the principle of differential noise reduction, is designed in this paper. With wide-band antennas and low-noise differential amplification module, electrostatic discharge can be detected in the band from 150 to 350 MHz. As a signal canceller in the system, the cancellation function of differential amplification module is effective for background and invalid for instantaneous signals. For those reasons, the system is sensitive for corona, and the detecting distance is up to 5 km. To eliminate interference from other sources, clustering and recognition methods were presented: Five evaluation parameters to describe electrostatic discharge were proposed; clustering method based on FCM and pattern recognition resorting to SVM method were conducted to distinguish spark discharge and corona discharge. In the following, the detecting system will be introduced at first.

## 2. DETECTING SYSTEM OF ELECTROSTATIC DISCHARGE

### 2.1. Analysis of Background Signals

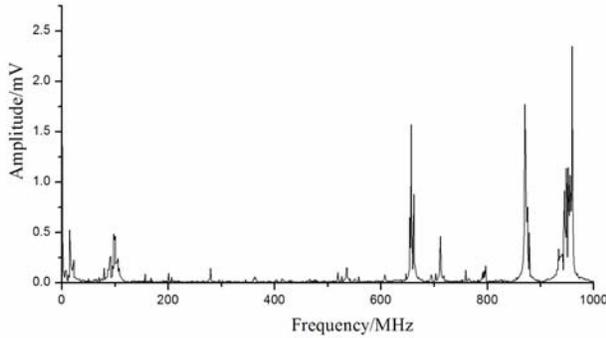
During the course of normal operation, corona discharge from high-voltage transmission lines produces electromagnetic emission over a wide range of frequencies, from 50 MHz to 1 GHz [2]. In general, the broadcast radio and broadcast television signals are narrow band, which are in certain range from 0 to 1 GHz.

Figure 1 provides the spectrum of background signal near a HVDC transmission line, which does not contain any electrostatic discharge. The signals have large amplitudes in the range from 0 to 100 MHz and 600 to 1000 MHz. They may be the signals of broadcast radio, broadcast television and direct electromagnetic radiation along the line. In addition, there is less noise between 100 and 600 MHz. Hence, the frequency range from 150 to 350 MHz is chosen to avoid the interference from broadcasting, and more pure information of corona discharge can be obtained.

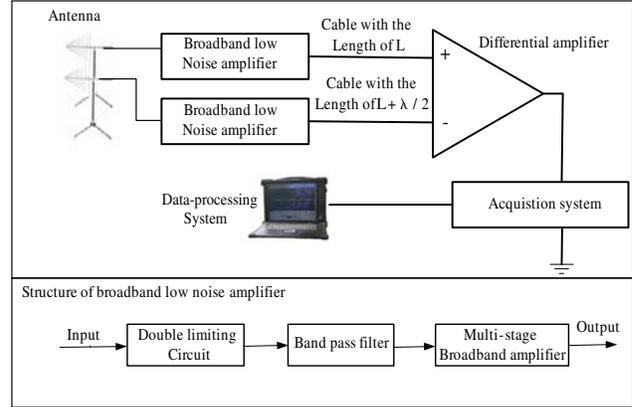
### 2.2. Design of Detecting System

An electrostatic discharge detecting system based on the principle of differential noise reduction is designed, and the structure of which is shown in Figure 2.

The system consists of two log-periodic antennas with a distance of 1 m between each other, a differential amplification module, an acquisition system and a data-processing system with clustering and recognition techniques. The antennas are log-periodic operated between 80 MHz and 5 GHz. The



**Figure 1.** Spectrum of background signal near a HVDC transmission line. Note that the amplitude of the noise is small between 150 and 350 MHz which is chosen as the frequency band of detecting system.



**Figure 2.** Electrostatic discharge detecting system.

broadband low noise amplifier (LNA) is composed of a double limiting circuit, band pass filters and a multi-stage broadband amplifier. They are employed to be pre-processing part of the system. The model of coaxial cable is RG58 with impedance  $50\Omega$ . And  $\lambda$  is set between 0.924 m and 1.014 m ( $\lambda = c/f$ ,  $c = 3 \times 10^8$  m/s and  $f_0 = 250$  MHz). The operation bandwidth of differential amplifier (Model THS4520) is 650 MHz. Sampling rate of acquisition system is 1 GSa/s. Data-processing system is designed to process signals collected by acquisition system. It contains three sections, which are features extraction, unsupervised clustering based on Fuzzy C-Means algorithm and support vector machine method, details of which are presented in Sections 3 and 4.

The detection principle is as follows:

In Figure 2, signals detected by the two antennas are filtered and amplified by two broadband LNA respectively.

$$\begin{cases} x_1(t) = n_1(t) + s_1(t) \\ x_2(t) = n_2(t) + s_2(t) \end{cases} \quad (1)$$

where  $x(t)$  is the signal output from LNA,  $n(t)$  the background signal in  $x(t)$ ,  $s(t)$  the electrostatic discharge signal, and subscripts 1 and 2 are numbers of the two signals.  $x_1(t)$  and  $x_2(t)$  are delivered to differential amplifier by two coaxial cables with the length difference of  $\lambda/2$ , where  $\lambda = c/f_0$ ,  $c = 3 \times 10^8$  m/s and  $f_0 = 250$  MHz. The two inputs of differential amplifier are given by:

$$\begin{cases} X_1(t) = x_1(t)e^{j\omega_1} \\ X_2(t) = x_2(t)e^{j\omega_2} \end{cases} \quad (2)$$

where  $\omega_1 = 2\pi fL/c$ ,  $\omega_2 = \pi(fL/c + f/f_0)$ ,  $L$  is the length of the coaxial cable. Compared with the detecting distance, the distance between the two antennas can be ignored. Therefore,  $x_1(t)$  and  $x_2(t)$  are approximately equal. Equation (2) can be given by:

$$X_1(t) \approx X_2(t)e^{j\pi f/f_0} \quad (3)$$

The output of differential amplifier is

$$Y(t) = X_1(t) - X_2(t) \quad (4)$$

Owing to the length difference, when  $f \in (0.5f_0, 1.5f_0)$ ,  $X_1(t)$  and  $X_2(t)$  are a couple of different-mode signals. The output  $Y(t)$  will be doubled by subtraction computing. In the contrary, signals not in the range  $(0.5f_0, 1.5f_0)$  as common-mode signals will be reduced.

Moreover, differential amplification module is a signal canceller. For slowly varying signals, such as background signals, the cancellation is effective. But for instantaneous signals (for instance electrostatic

discharge), the cancellation is invalid. Therefore, the system designed in this paper can fulfill narrow-band detection in the range from 150 to 350 MHz, getting rid of the interference from broadcasting and other noise. The structure of a broadband LNA is shown in Figure 2. It includes a double limiting circuit, a band pass filter and a multi-stage broadband amplifier. Double limiting circuit is designed for restraining transient interference pulse to protect the post stage circuit. A trap filter is set in bandpass filter to filter noise signal in the range from 219.25 to 222.75 MHz.

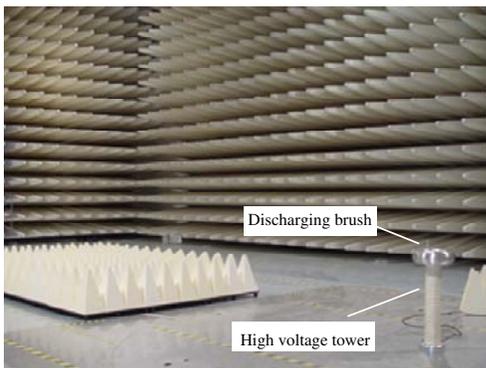
It should be noted that the background signal varies from country to country. The detecting frequency range should be selected based on the analysis of the background signal in advance. By varying the length of coaxial cables  $L$ , the operating frequency band of detecting system can be applied to other places in the world.

### 2.3. Detection Experiments

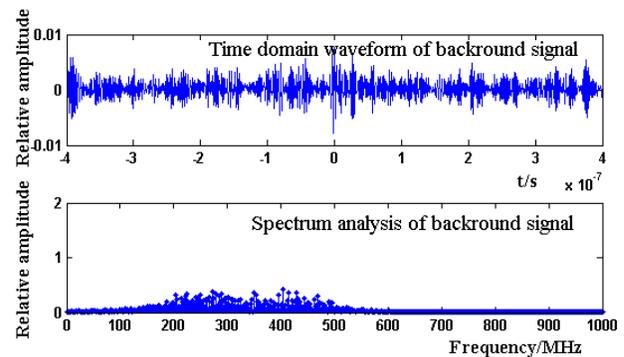
The corona is divided into positive and negative corona depending upon the polarity of the supply voltage. There is a huge difference between positive and negative corona. Above 50 MHz, EMI caused by negative corona predominates over that due to positive corona. For this reason, this paper is limited to a study of the detection of negative corona. To test the performance of detecting system, a simulated experiment is designed. HVDC transmission line is replaced by a corona discharge source which can veritably simulate the corona discharge of transmission lines. It consists of a high voltage tower and a discharging brush as shown in Figure 3. The high voltage tower is supplied by a DC high voltage source whose value can reach to  $\pm 300$  kV. The discharging brush, with numbers of wires as discharge points, is attached on the top of tower. When the voltage applied on brush increases to breakdown voltage, corona discharge occurs. As the voltage increases, the discharge energy and times will be enhanced apparently. Hence, along with its portable packaging, it can simulate any corona discharge with different distance and energy.

At first, background noise is measured by detecting system. Figure 4 provides time domain waveform and spectrum of the background signal. It can be seen from the frequency waveform that the background signal ranging from 0 to 100 MHz and 600 MHz to 1000 MHz were weakened to be negligible. So the interference from broadcast radio and broadcast television in this range is eliminated. The spectral range of detecting system is narrowed about 150 MHz to 550 MHz, which can fulfill narrow-band requirement mentioned above in the range from 150 MHz to 350 MHz.

Then, experiment of corona discharge is carried out. Detecting distance (distance between discharge source and detecting system) is set as 300 m. The working voltage of DC high voltage source is set as  $-40$  kV. Magnification of differential amplification module is 10 dB. Sampling rate of acquisition system



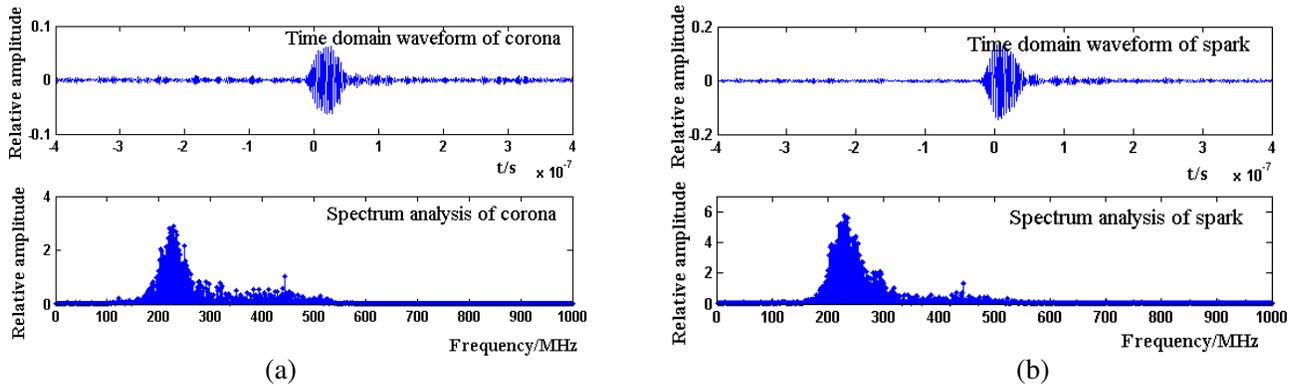
**Figure 3.** Electrostatic discharge source. High voltage tower is supplied by a high voltage source. The discharging brush, with numbers of wires as discharge points, is attached on the top of tower. When the voltage applied on brush increases to breakdown voltage, electrostatic discharge occurs.



**Figure 4.** Time domain and spectrum analysis of background signal.

is 1 GSa/s. On this occasion, the corona discharge has good repeatability and stability. Figure 5(a) provides the typical waveforms in time domain and frequency domain about one of the corona samples. Compared with Figure 4, the peak value of corona discharge is 0.062, which is 10 times larger than the background signal. And in frequency domain, signal intensity increases greatly, especially in the range from 150 MHz to 350 MHz. Therefore, corona discharge detected by the detecting system can be distinguished easily against the background. And the maximum detecting distance is up to 5 km.

However, we found that the discharges produced by surrounding automobile influence the corona detection, which is called the spark discharge [25]. Figure 5(b) shows a typical waveform of spark discharge generated by an engine of automobile. It is a type of electrostatic discharge different from corona discharge in generated mechanism, generally called “secondary ESD” [26]. Compared with Figure 5, the intensity of spark discharge is higher than corona discharge, but the wave shape and spectrum characteristic are similar from 150 to 350 MHz. When the detection distance of spark discharge becomes further, its intensity will decrease to a low level, which leads to the difficulty to distinguish the two discharges by eyes. Consequently, this secondary discharge will interfere to the detection of corona discharge.



**Figure 5.** (a) Time domain and frequency domain of the corona samples. (b) Time domain and frequency domain of the spark samples.

In addition, lighter sparking, switching the door of car and operating a mechanical switch can also generate spark discharge. Therefore, the pattern recognition methods are needed to distinguish the two kinds of discharges. In Section 3, methods of feature extractions are proposed as post-progress of pattern recognition.

### 3. FEATURE EXTRACTIONS

Feature extraction provides a compact and meaningful representation of the discharge characteristics embedded in waveform for post-processing. Corona discharge from HVDC transmission lines is stochastic without regular phase information, and the discharge interval is up to 1 ms. Hence, the feature evaluation parameters in Fromm and Gulski’s models are unsuitable. In this section, five features in single pulse are extracted. Since the spectrum characteristics of the two kinds of discharge mentioned above are too similar to distinguish, shape parameters and statistics parameters of time domain waveform are proposed to evaluate electrostatic discharge characteristics. They are peak factor, form factor, skewness, kurtosis and mean square error, which are expressed as  $P$ ,  $F$ ,  $Ske$ ,  $Kur$  and  $\sigma$ . The definitions of them are described as follows:

$$P = \frac{U_m}{U_{RMS}} \quad (5)$$

$$F = \frac{U_{RMS}}{U} \quad (6)$$

$$Ske = \frac{\frac{1}{N} \sum_i^N |U(\mathbf{t}(i)) - \bar{U}|^3}{S^3} \quad (7)$$

$$Kur = \frac{\frac{1}{N} \sum_i^N |U(\mathbf{t}(i)) - \bar{U}|^4}{S^4} \quad (8)$$

$$\sigma = \sqrt{\frac{1}{N} \sum_i^N |U(\mathbf{t}(i)) - \bar{U}|^2} \quad (9)$$

where  $U_m$ ,  $U_{RMS}$  and  $\bar{U}$  represent peak value, root mean square and mean value of signal, respectively.  $N$  is sample point number of single pulse and  $S$  the standard deviation given by

$$S = \sqrt{\frac{1}{N-1} \sum_i^N |U(\mathbf{t}(i)) - \bar{U}|^2} \quad (10)$$

100 discharge samples were collected, including 50 corona discharge samples and 50 spark discharge samples. The corona samples were generated by electrostatic discharge source shown in Figure 3 with  $-40$  kV operating voltage and 300 m detecting distance. The spark samples were generated by automobile engines nearby. The five features were extracted in batches as shown in Figure 6, where blue lines represent corona discharge, and green lines represent spark discharge.

In Figures 6(a), (b) and (e), more than 80% samples of corona discharge are greater than spark discharge in peak factor value and form factor value, and 86% samples are less than spark discharge in mean square error. In Figures 6(c) and (d), the values difference of the two discharges is small in skewness and kurtosis, but parameters of spark are steadily distributed in a small range. It can be seen that the five features have great distinguish capacity for the two discharges, especially peak factor, form factor and mean square error. So they can be made as feature evaluation parameters to be applied to clustering and pattern recognition.

#### 4. PATTERN RECOGNITION

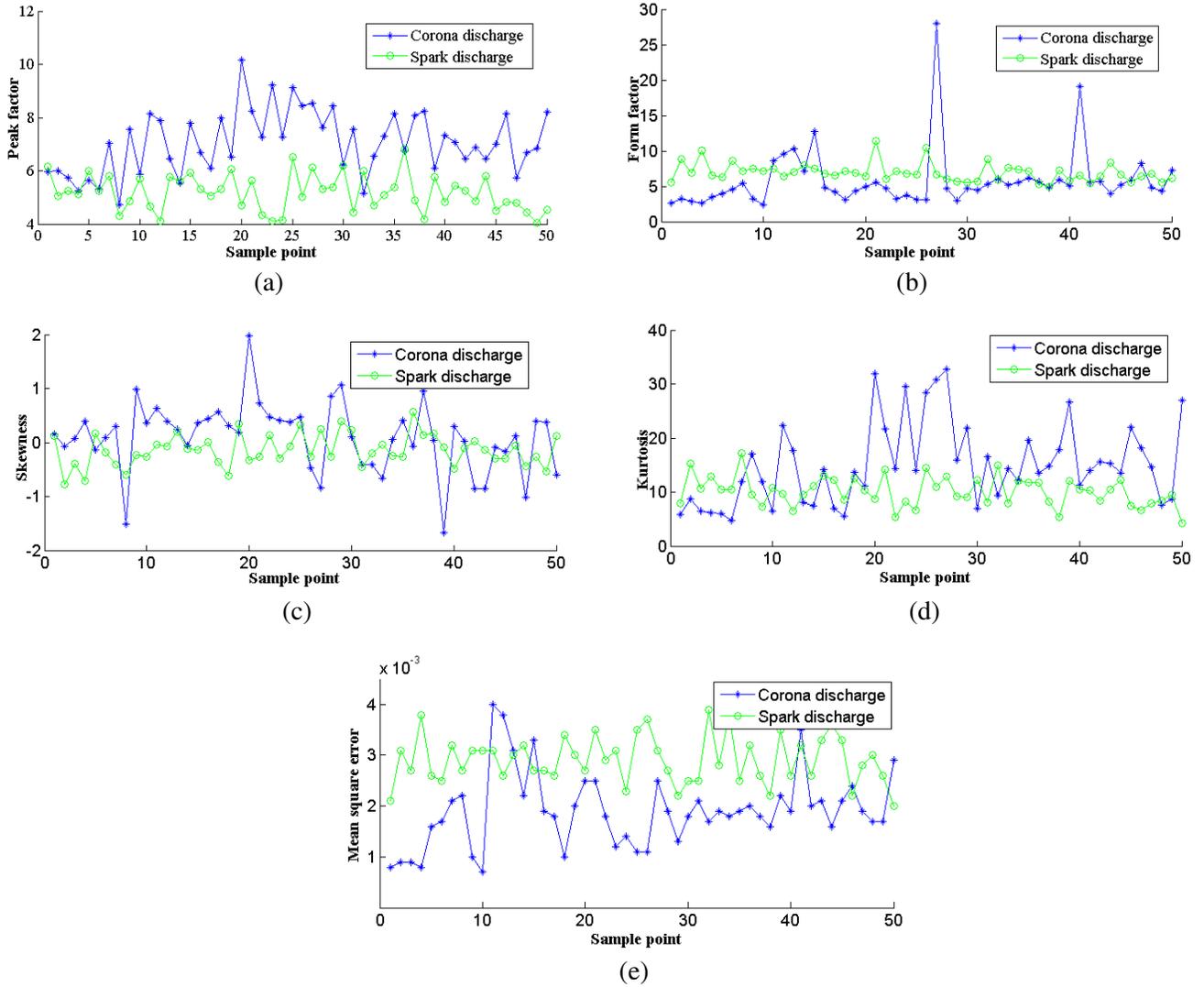
Pattern recognition is achieved by using unsupervised clustering and SVM (support vector machine). Unsupervised clustering of pulse sequence aims at grouping together discharge pulses characterized by similar waveforms. After clustering, the samples of discharge are split into two sub-groups as the training and testing sets of SVM. At first, the algorithms of clustering and SVM are introduced.

##### 4.1. Unsupervised Clustering

There are various approaches applied in clustering large databases, including Expectation Maximization (EM) [27, 28], K-Means [29] and Fuzzy C-Means (FCM) [30–32]. In this section, FCM algorithm is applied to process the feature evaluation parameters. This method was first developed by Dunn in 1973 [31] and then improved by Bezdek in 1981 [32], which is based on minimization of the objective function:

$$J_{FCM} = \sum_{i=1}^c \sum_{k=1}^n u_{ik}^m \|x_k - v_i\|^2 \quad (11)$$

where  $u_{ik}$  ( $i = 1, 2, \dots, c$ ;  $k = 1, 2, \dots, n$ ) is the membership degree of  $x_k$  in the cluster  $i$ ,  $m$  is the fuzzification parameter with the range  $m > 1$ ,  $x_k$  is the  $k$ th measured data point,  $v_i$  is the center of the



**Figure 6.** Results of the feature extractions. (a) Peak factor, (b) form factor, (c) skewness, (d) kurtosis, and (e) mean square error.

cluster  $i$ .  $u_{ik}$  and  $v_i$  are given by:

$$u_{ik} = \begin{cases} \left[ \sum_{j=1}^c \left( \frac{\|x_k - v_i\|}{\|x_k - v_j\|} \right)^{\frac{2}{m-1}} \right]^{-1}, & \|x_k - v_i\| \neq 0 \\ 1, & \|x_k - v_i\| = 0 \text{ and } k = j \\ 0, & \|x_k - v_i\| = 0 \text{ and } k \neq j \end{cases} \quad (12)$$

$$v_i = \frac{\sum_{k=1}^n (u_{ik})^m x_k}{\sum_{k=1}^n (u_{ik})^m} \quad (13)$$

Initialize the parameters of objective function, and perform the iteration by implementing the proposed Equations (11) ~ (13), with the update of the membership degree matrix  $u_{ik}$  and center of

the cluster  $v_i$ , until the procedure converges a local minimum of  $J_{FCM}$ . The sub-groups of samples are determined by the position of maximum value in matrix  $u_{ik}$ .

## 4.2. Pattern Recognition Based on SVM

Recently, many methods are available for pattern recognition, such as BP neural network, LVQ [33, 34] and SVM. BP and LVQ methods are supervised networks, which are frequently used in recognition. The former has multi-layer neurons and optimizes the weights and threshold of network by feeding forward the predicting errors. The latter is developed from competitive algorithm. It is a version of vector quantization and generates code vectors to produce near optimal decision. The two methods are performed in feedback manner that they rely mostly on the historical data, which is the major limitation of them. In this section, SVM method proposed by Vapnik and his co-workers [35] is selected for its powerful ability in general purpose pattern recognition without the limitation of BP and LVQ methods. The algorithm of SVM is described in [35, 36].

The SVM used in the paper is a two-class classifier constructed from sum of kernel functions, and the decision function of which is given by:

$$f(x) = \text{sign} \left[ \sum_j^N a_j y_j K(\mathbf{x}, \mathbf{x}_j) + b \right] \quad (14)$$

where  $(\mathbf{x}_j, y_j)$  is a training set,  $\mathbf{x}_j$  the five discharge characteristics as the  $j$ th input pattern,  $y_j \in \{-1, 1\}$  the output pattern,  $a_j$  the Lagrange multiplier,  $b$  a constant, and  $K(\mathbf{x}, \mathbf{x}_j)$  the kernel function which can be expressed as:

$$K(\mathbf{x}, \mathbf{x}_j) = \begin{cases} \mathbf{x}_j^T \mathbf{x} & \text{linear} \\ (\mathbf{x}_j^T \mathbf{x} + 1)^p & \text{polynomial} \\ \exp(-\gamma \|\mathbf{x} - \mathbf{x}_j\|^2) & \text{RBF} \\ \tanh(\delta \mathbf{x}_j^T \mathbf{x} + r) & \text{MLP} \end{cases} \quad (15)$$

where  $p$ ,  $\gamma$ ,  $\delta$  and  $r$  are constants. The two-class SVM classifiers are obtained as solution to the following optimization problem:

$$\begin{cases} \min \phi(\mathbf{w}, b) = \frac{1}{2} \|\mathbf{w}\|^2 \\ \text{s.t. } y_j [(\mathbf{w}^T \cdot \mathbf{x}_i) + b] \geq 1 \end{cases} \quad (16)$$

The solution to Eq. (16) is Lagrange function

$$L(\mathbf{w}, b, \mathbf{a}) = \frac{1}{2} \mathbf{w}^T \mathbf{w} - \sum_{j=1}^N a_j [y_j (\mathbf{w}^T \cdot \mathbf{x}_i + b) - 1] \quad (17)$$

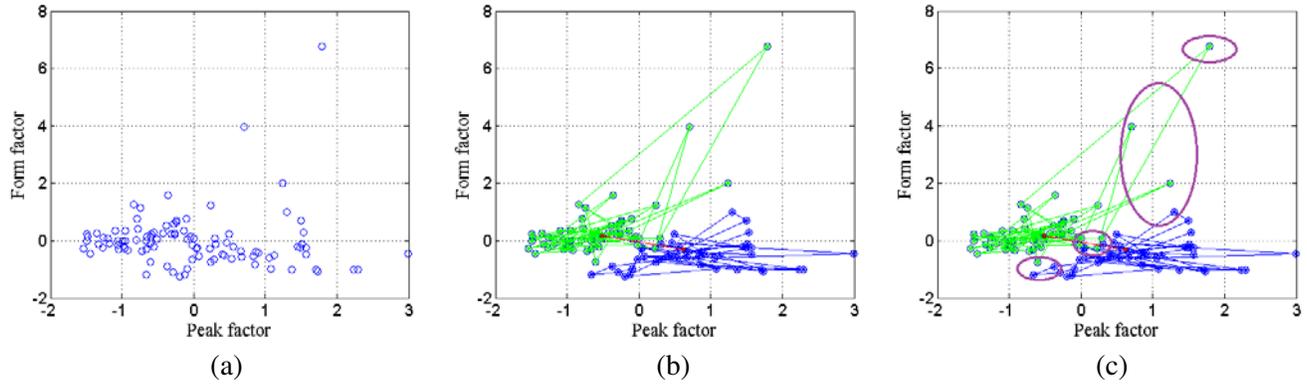
So the optimization problem of Eq. (16) can be solved by giving the results of the support values  $a_j$  and  $b$ .

## 5. RESULTS AND DISCUSSION

### 5.1. Unsupervised Clustering Results

100 discharge samples mentioned above are clustered by FCM algorithm. The data set are described by a 5-dimension feature space, comprising the five discharge characteristics (peak factor, form factor, skewness, kurtosis and mean square error). For visualization, the normalized data are in terms of two coordinates with peak factor as  $x$ -axis variables and form factor as  $y$ -axis variables. Figure 7 shows the clustering process of electrostatic discharge samples.

Figure 7(a) provides the 100 original sample points. Figure 7(b) is the clustering result. 100 samples are classified into two sub-groups, where blue points are corona discharges, and green points are spark



**Figure 7.** Clustering process of electrostatic discharge samples. (a) Describes the original sample points. It consists of 100 samples mixed together as clustering data, (b) is the clustering result, where blue points are corona discharge and green points are spark discharge. And the black points are clustering centers of the two sub-groups, (c) is selecting training set for pattern recognition. Points in purple circles will be taken out, for its coordinates far away from clustering center (the two circles above-mentioned) and the ones next to each other (the two circles below).

discharges. Meanwhile, clustering centers are calculated as  $(0.6197, -0.5055)$  and  $(-0.2936, 0.1670)$ , respectively. Note that some samples in one sub-group are also densely connected with another sub-group, and some samples coordinates are far from their clustering centers, as circled by purple lines in Figure 7(c). We deal with these samples as uncertain points. After clustering, the two sub-groups can be made as training set for pattern recognition.

### 5.2. Pattern Recognition Results

In the paper, the inputs of SVM  $\mathbf{x}_j(x_{1j}, x_{2j}, \dots, x_{5j})$  are the five feature evaluation parameters, where  $x_{1j} \sim x_{5j}$  are the peak factor, form factor, skewness, kurtosis and mean square error of the  $j$ th sample, respectively. The output  $y_j$  has two results: 0 or 1. When  $y_j = 1$ , the sample is recognized as corona discharge. When  $y_j = 0$ , the sample is recognized as spark discharge. 60 samples are selected as training set by taking out some sample points whose clustering results are uncertain, as circled in Figure 7(c). Besides, 40 samples which contain 20 corona discharge samples and 20 spark discharge samples are collect as test set. For comparison, BP and LVQ approaches are taken to test the recognition ability. BP neural network is set in there layers. The neuron numbers of input layer and output layer are 5 and 1, respectively. To get an optimizing recognition rate, the neuron number of hidden layer is tested by 3, 6 and 9. In LVQ, the neuron number of hidden layer is 20. And the neuron numbers of other two layers are set as BP. Table 1 shows the recognition results of the three approaches.

For BP neural network, the average recognition rate varies greatly with different neuron numbers in hidden layer in the range from 77.5% to 92.5%. The average recognition rate of LVQ is changed little

**Table 1.** BP, LVQ and SVM recognition rate on the test set.

Approach	Neurons in hidden layer	Recognition rate (%)		Performance (average recognition rate/%)
		corona discharge	spark discharge	
BP	3	65	90	77.5
	6	90	95	92.5
	9	85	95	90
LVQ	20	90	100	95
SVM	—	95	100	97.5

with the variation of neurons number in hidden layer, and the best result is 95%. Rate of SVM reaches to 97.5%. It should be noted that the training processing of SVM is simple, without adjusting structure parameters of SVM. And the recognition ability of SVM is the best one in comparison with BP and LVQ methods.

The above results show that the detecting system is capable of noise suppression and pattern recognition. The information of data progressing is extracted from signal pulse, which is applied in clustering and recognition algorithms as input samples. The clustering results and recognition rate of SVM prove that the five feature evaluation parameters are effective. Therefore, the effectiveness of clustering and pattern recognition depends on the pulse waveform. If the sampling rates of acquisition system and detection bandwidth are improved, which can be realized by high-speed acquisition and multi-channel with different operating band, more time domain information of pulse waveform will be obtained.

## 6. CONCLUSION

In this paper, a detecting system is designed to detect corona discharge from HVDC transmission lines in complex background. The system combines two log-periodic antennas, a differential amplification module, an acquisition system and a data-processing system. The differential amplification module is designed based on the principle of differential noise reduction, which can fulfill narrow-band detection breaking away interference from broadcasting and easily catch the electrostatic discharge signal. A detecting experiment is conducted using an Electrostatic discharge source to simulate the discharge from HVDC transmission lines. The results show that detecting platform can fulfill narrow-band detection ranging from 150 MHz to 350 MHz to get rid of interference from broadcasting and other noise. Moreover, with a typical characteristic in time and frequency domains, corona discharge can be caught easily against the background.

To get rid of interference from spark discharge, data progressing system containing feature extractions, clustering and recognition technologies are developed. Clustering is realized with extracting five discharge features, including peak factor, form factor, skewness, kurtosis and mean square error. The unsupervised clustering Fuzzy C-Means (FCM) method is used to achieve fast separation for electrostatic discharges and provide training set for pattern recognition. Pattern recognition resorts to Support Vector Machine (SVM) method. After feature extraction and clustering, experiment to test the recognition ability is taken, with BP and LVQ approaches as comparison. The results show that SVM recognizer with recognition rates of 97.5% achieves higher performance than BP and LVQ methods. We may conclude that the detecting system would be an interesting alternative for electrostatic discharge detection. And we expect that this new design will be employed with electrostatic discharge real time detection from HVDC transmission lines.

## ACKNOWLEDGMENT

This work was supported in part by the National Natural Science Foundation of China under Grant No. 61172035. The authors would like to thank Cui Yao-zhong, Liu Wei-dong and Prof. Wei Ming for the measurement help or discussion.

## REFERENCES

1. Nakano, Y. and Y. Sunaga, "Availability of corona cage for predicting radio interference generated from HVDC transmission line," *IEEE Trans. Power Del.*, Vol. 5, No. 3, 1436–1442, Jul. 1990.
2. Maruvada, P. S., *Corona Performance of High-voltage Transmission Lines*, Research Studies Press Ltd, England, 2000.
3. CISPR TR 18-1 ED.2, "Radio interference characteristics of overhead power lines and high-voltage equipment-part 1," 2010.
4. CISPR TR 18-1 ED.2, "Radio interference characteristics of overhead power lines and high-voltage equipment-part 2," 2010.

5. Zhou, C., Y. Liu, and X. Rui, "Mechanism and characteristic of rain-induced vibration on high-voltage transmission line," *Journal of Mechanical Science and Technology*, Vol. 26, No. 8, 2505–2510, 2012.
6. Bracken, T. D., R. S. Senior, and W. H. Bailey, "DC electric fields from corona-generated space charge near AC transmission lines," *IEEE Trans. Power Del.*, Vol. 20, No. 2, 1692–1702, 2005.
7. Hatanaka, G. K., "Field measurement of VHF noise from an operating 500-kV power line," *IEEE Trans. Power App. Syst.*, Vol. 100, No. 2, 863–872, 1981.
8. Huertas, J. I., R. Barraza, and J. M. Echeverry, "Wireless data transmission from inside electromagnetic fields," *Journal of Microwave Power and Electromagnetic Energy*, Vol. 44, No. 2, 88–97, 2010.
9. Chartier, V. L., R. Sheridan, J. N. Diplacido, et al., "Electromagnetic interference measurements at 900 MHz on 230 kV and 500 kV transmission lines," *IEEE Trans. Power Del.*, Vol. 2, No. 2, 140–149, 1986.
10. Chao, F., X. Cui, X. Zhou, et al., "Impact factors in measurements of ion-current density produced by high-voltage DC wire's corona," *IEEE Trans. Power Del.*, Vol. 28, No. 3, 1414–1422, 2013.
11. Radwan, R. M. and A. M. Mahdy, "Electric field mitigation under extra high voltage power lines," *IEEE Trans. Dielectr. Electr. Insul.*, Vol. 20, No. 1, 54–62, 2013.
12. Lv, F., S. You, Y. Liu, Q. Wan, and Z. Zhao, "AC conductors' corona-loss calculation and analysis in corona cage," *IEEE Trans. Power Del.*, Vol. 27, No. 2, 877–885, 2012.
13. Morris, R. M., A. R. Morse, J. P. Griffin, O. C. Norris-Elye, C. V. Thio, and J. S. Goodman, "The corona and radio interference performance of the Nelson river HVDC transmission lines," *IEEE Trans. Power App. Syst.*, Vol. 98, No. 6, 1924–1936, 1979.
14. Maryvada, P. S., "Electric field and ion current environment of HVdc transmission lines: Comparison of calculations and measurements," *IEEE Trans. Power Del.*, Vol. 27, No. 1, 401–410, 2012.
15. Vahidi, B., M. J. Alborzi, and H. Aghaeinia, "Corona detection on surfaces of insulators using ultrasound sensors and fibre-optic transmission system," *European Transactions on Electrical Power*, Vol. 15, No. 5, 413–424, 2005.
16. Fromm, U. and F. H. Kreuger, "Statistical behavior of partial discharge at DC voltage," *Japanese J. Appl. Phys.*, Vol. 33, 6708–6715, 1994.
17. Fromm, U., "Interpretation of partial discharges at DC voltage," *IEEE Trans. Dielectr. Electr. Insul.*, Vol. 2, No. 5, 761–769, Oct. 1995.
18. Fromm, U., "Partial discharge and breakdown testing at high dc voltage," Delft University of Technology (the Netherlands), Delft University Press, ISBN 90-407-1155-0, 1995.
19. Morshuis, P., M. Jeroense, and J. Beyer, "Partial discharge part XXIV: The analysis of PD in HVDC equipment," *IEEE Electr. Insul. Mag.*, Vol. 13, No. 2, 6–16, 1997.
20. Gulski, E., H. P. Burger, G. H. Vaillancourt, and R. Brooks, "PD pattern analysis during induced test of large power transformers," *IEEE Trans. Dielectr. Insul.*, Vol. 7, No. 1, 95–101, Feb. 2000.
21. Morshuis, P. H. F. and J. J. Smit, "Partial discharges at voltage: Their Mechanism, detection and analysis," *IEEE Trans. Dielectr. Insul.*, Vol. 12, No. 2, 763–808, 2005.
22. Wang, M. H., "Partial discharge pattern recognition of current transformers using an ENN," *IEEE Trans. Power Del.*, Vol. 20, No. 3, 1984–1990, 2005.
23. Wang, M. H., "Extension neural network-type 2 and its applications," *IEEE Trans. Neural Netw.*, Vol. 16, No. 6, 1352–1361, 2005.
24. Si, W., J. Li, P. Yuan, and Y. Li, "Digital detection, grouping and classification of partial discharge signals at DC voltage," *IEEE Trans. Dielectr. Insul.*, Vol. 15, No. 6, 1663–1674, 2008.
25. Yamamoto, S. and O. Ozeki, "Properties of high-frequency conducted noise from automotive electrical accessories," *IEEE Trans. Electromagn. Compat.*, Vol. 25, No. 1, 2–7, 1983.
26. Xiao, J., D. Pommerenke, J. L. Drewniak, H. Shumiya, T. Yamada, and K. Araki, "Model of secondary ESD for a portable electronic product," *IEEE Trans. Electromagn. Compat.*, Vol. 54, No. 3, 546–555, 2012.

27. Dempster, A. P., N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society, Series B*, Vol. 39, No. 1, 1–38, 1977.
28. Rohlfing, T., D. L. B. Russakoff, and C. R. Maurer, "Performance-based classifier combination in atlas-based image segmentation using expectation-maximization parameter estimation," *IEEE Trans. Med. Imaging*, Vol. 23, No. 8, 983–994, 2004.
29. Kanungo, T., D. M. Mount, N. S. Netanyahu, C. D. Piatko, and Y. Wu, "An efficient k-means clustering algorithm: Analysis and implementation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 24, No. 7, 881–892, Jul. 2002.
30. Zhang, S., R. Wang, and X. Zhang, "Identification of overlapping community structure in complex networks using fuzzy C-means clustering," *Physica A: Statistical Mechanics and its Applications*, Vol. 374, No. 1, 483–490, 2007.
31. Dunn, J. C., "A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters," *J. Cybernet.*, Vol. 3, No. 3, 32–57, 1973.
32. Bezdek, J. C., *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, New York, 1981.
33. Camastra, F. and A. Vinciarelli, "Cursive character recognition by learning vector quantization," *Pattern Recognition Letters*, Vol. 22, No. 6, 625–629, 2001.
34. Camastra, F. and A. Vinciarelli, "Combining neural gas and learning vector quantization for cursive character recognition," *Neurocomputing*, Vol. 51, 147–159, Apr. 2003.
35. Vapnik, V. N., *Statistical Learning Theory*, John Wiley & Sons, New York, 1998.
36. Sharkawy, R. M., R. S. Mangoubi, T. K. Abdel-Galil, M. M. Salama, and R. Bartnikas, "SVM classification of contaminating particles in liquid dielectrics using higher order statistics of electrical and acoustic PD measurements," *IEEE Trans. Dielectr. Electr. Insul.*, Vol. 14, No. 3, 669–678, 2007.