

Article

Entropy SVM-Based Recognition of Transient Surges in HVDC Transmissions

Guomin Luo * , Changyuan Yao, Yinglin Liu, Yingjie Tan and Jinghan He

School of Electrical Engineering, Beijing Jiaotong University, Beijing 100044, China; 16121566@bjtu.edu.cn (C.Y.); 17126021@bjtu.edu.cn (Y.L.); 17121497@bjtu.edu.cn (Y.T.); jhhe@bjtu.edu.cn (J.H.)

* Correspondence: gmluo@bjtu.edu.cn or guominluo@hotmail.com; Tel.: +86-10-5168-3740

Received: 19 April 2018; Accepted: 28 May 2018; Published: 31 May 2018



Abstract: Protection based on transient information is the primary protection of high voltage direct current (HVDC) transmission systems. As a major part of protection function, accurate identification of transient surges is quite crucial to ensure the performance and accuracy of protection algorithms. Recognition of transient surges in an HVDC system faces two challenges: signal distortion and small number of samples. Entropy, which is stable in representing frequency distribution features, and support vector machine (SVM), which is good at dealing with samples with limited numbers, are adopted and combined in this paper to solve the transient recognition problems. Three commonly detected transient surges—single-pole-to-ground fault (GF), lightning fault (LF), and lightning disturbance (LD)—are simulated in various scenarios and recognized with the proposed method. The proposed method is proved to be effective in both feature extraction and type classification and shows great potential in protection applications.

Keywords: HVDC transmission; frequency spectrum entropy; SVM; transient surge recognition

1. Introduction

High voltage direct current (HVDC) transmission plays an important role in power transmissions due to its advantages such as large transmission capacity and good performance in power flow control [1,2]. It has been widely applied in delivering large amount of power and connecting asynchrony power grids. Generally, traveling wave-based protection or voltage derivate based protection are used as the primary protection, and under-voltage protection or current-differential protection are adopted as the backup protections in the HVDC systems [3,4]. Traveling wave-based protection captures the transient traveling surges on transmission lines and make quick response in a very short time. As the shunt capacitors and the smoothing inductors in convertor station can effectively reflect traveling waves, traveling wave-based protection can easily distinguish faults beyond the protected zone. The time-domain features, such as magnitude and changing rate of electrical measurements, are commonly used in protection judgement [5,6].

However, the time domain-based method is sensitive to surge disturbances, for example, lightning strokes [7,8]. The transient waveforms of surge interferences look similar with the ones of ground faults in some cases. Such similarity makes them difficult to be discriminated. In order to improve the reliability of protection actions, the identification of transient surges is a critical function of protection algorithm, which includes two important aspects: feature extraction and classification algorithm.

To effectively identify transient surges, the unique features of various signals should be extracted, and the features should represent the signals with a stable and reliable performance. Frequency analysis that can provide more details on spectrum differences are often used to fully utilize the signal transient information in both time and frequency domains and reveal a better characterizing performance. Various frequency-based features are adopted and reported to generate good performance—for

example, wavelet energy spectrum [9], S-transform distribution [10,11], and frequency energy spectrum [12]. These amplitude-based or energy-based features heavily depend on the magnitude of spectrum, which can be distorted during propagation. Entropy, which only describes the distribution of frequency spectrum, was proved to be effective in characterizing transient surges in HVDC. Compared with traditional magnitude-based feature extractions that will be affected by unequally distortions, the entropy can characterize the distribution of energy in a certain range. It is a powerful tool to extract transient signal characteristics and is immune to variations of magnitude attenuations and distortions [13–16].

Besides good features, the classification algorithm is also an important factor in recognizing various transient signals. In most cases, a threshold is used to realize the signal classification, and the value of threshold is sometimes selected manually according to different signal characteristics. Such a threshold-based method is susceptible to the influence of line parameters, transition resistance, noise and so on [16,17]. Machine learning algorithms can effectively solve the problems of uncertain correspondence [18]. They show a number of advantages in pattern recognition, classification, and generalization and play an important role in the field of power system fault diagnosis [19–21]. A lot of classification algorithms are used in fault analysis, such as artificial neural networks (ANNs), support vector machines (SVMs), auto-encoders, expert systems, and so on [22–24]. The two most popular classifiers are ANN and SVM. When compared with SVMs, ANNs are reported to have slow algorithm convergence, poor adaptability, and high requirements for training samples. SVMs, which are proposed based on statistical learning theory, reveal better capability in solving classification problems with a small size of samples. For fault classification of power systems, where the number of samples are limited, SVM could be a better choice [25,26].

Combining the advantages of entropy and SVMs, this paper proposed a transient surge recognition method to improve the reliability of HVDC protections. Based on the analysis of transient waveforms of pole-to-ground faults (GFs), lightning faults (LFs), and lightning disturbances (LDs) in direct current (DC) transmission lines, the features of different transient signals are represented by frequency spectrum entropy (FSE) vectors. The FSE vector is then used as the input of a SVM structure to realize the recognition of transient interferences. A typical HVDC transmission line is modeled and simulated in different scenarios to demonstrate the effective performance of proposed method. The simulation results and the comparisons prove the potential of FES-SVM in protection applications.

2. HVDC and Transient Surges

2.1. Fundamentals of HVDC

The grounding method of a single polar DC system can lead to symmetric and asymmetric DC transmissions. Both kinds of grounding methods are widely used in practical projects. The symmetric DC transmission only employs one convertor to construct positive and negative poles and are more popular for voltage-source converters-high voltage direct current (VSC-HVDC) systems. A typical two-terminal VSC-HVDC system, as shown in Figure 1, is analyzed in this research and modeled on the platform of PSCAD/EMTDC. The midpoint of the DC supporting capacitor is grounded to form a symmetric DC transmission. Such a grounding method can reduce the insulation requirement of DC devices and avoid live currents in grounding loops under normal operation [27,28]. The bus voltage of VSC1 is controlled while the power of VSC2 is controlled.

Measuring units M are installed at the terminals of transmission lines to provide useful information for protection devices. The smoothing reactor L on both terminals of the transmission line can block high frequency components of the transient signal from convertors and thus reduce the influences of interferences beyond the protection zone of transmission lines. However, the transient interferences generated on transmission lines can still affect the protection judgement. To improve the reliability of DC line protections, the transients that can be detected by measuring units M

should be discussed. Therefore, the most commonly encountered interference-lightning-is analyzed. The waveforms of faults and lightning interferences are compared and discussed.

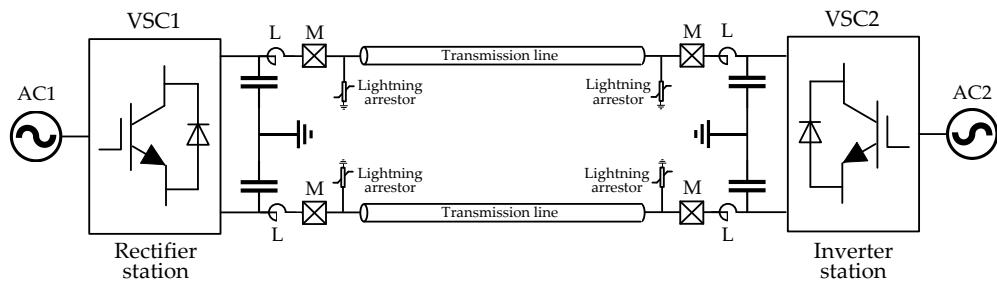


Figure 1. Diagram of a typical voltage-source converters (VSC)-high voltage direct current (HVDC) transmission line.

2.2. Pole-To-Ground Fault

GF is actually a kind of short circuit fault. It is the most commonly encountered fault in practical HVDC transmission lines. For the symmetric grounding VSC-HVDC transmission system mentioned above, the fault transient procedure includes two stages: discharge of supporting capacitor and current feeding of alternating current (AC) sources [17]. In the first stage, the traveling wave produced by faults moves quickly along the transmission line. When it reaches the converter station, the supporting capacitor discharges quickly. Such discharge causes a large amplitude decrease of DC bus voltage and quick rising fault currents. As the bus voltage decreases to AC phase voltages, the AC sources begin to feed the fault circuit and the second stage starts. The diodes on each leg of convertor are commutated in an uncontrolled commutation mode, and the overall current tends to be steady. Figure 2a illustrates five typical waveforms of GF, and Figure 2b shows their frequency spectrums. The details of these five GFs are as follows: (1) $Z_g = 20 \Omega$, $L_f = 10 \text{ km}$; (2) $Z_g = 20 \Omega$, $L_f = 50 \text{ km}$; (3) $Z_g = 10 \Omega$, $L_f = 100 \text{ km}$; (4) $Z_g = 1 \Omega$, $L_f = 150 \text{ km}$; (5) $Z_g = 0.01 \Omega$, $L_f = 200 \text{ km}$. Here, Z_g stands for the grounding impedance of fault, and L_f denotes the distance of fault point. Although these waveforms look quite different in the time domain, they have similar frequency spectrums. All of their spectrums decrease smoothly from lower frequencies to higher ones. Small ripples that distribute equally along the frequency range can be found.

2.3. Lightning Transients

For long distance transmission, overhead lines are preferred in the aspect of economy and maintenance. But overhead lines are generally erected high away from the earth and located in open area, where lightning disturbance occurs. The lightning strokes can produce quick rising transients that are interferences in some cases or faults when the insulations are breakdown [29]. Therefore, recognition of the lightning transient surges has great significance to improve the reliability of traveling wave based protection.

The lightning stroke is actually a kind of discharge of electric charges between clouds and the earth. It can be modeled by a current source and injects current into the transmission systems. In the aspect of transmission line protection, lightning surges can be divided into two kinds: lightning disturbance (LD) and lightning fault (LF). The formation of these two transient surges are similar, but the results are different. Both LD and LF are caused by overvoltage that is produced by direct strokes or indirect strokes. For overhead transmission lines, the direct strokes that hit the bare conductor can generate large overvoltage and result in short-circuit of lightning surge arrestors. Although a large overvoltage is produced, the current is not so large due to the characteristic impedance of the transmission line (around several hundred ohms) and the operation of lightning protection devices at the line terminals. This kind of lightning stroke only generates electrical interferences. The lightning-caused traveling wave continuously refracts due to discontinuities at both ends of the line and eventually decays to

zero [30]. The indirect strokes hit a point in the vicinity of the transmission lines—for example, the tower or shielding wire. If the lightning stroke current is high enough to cause tower-to-conductor flashover or shielding wire failure, LF occurs [31,32]. In this case, the induced currents in transmission line could be greater than those due to LDs.

Generally, the lightning strike is modeled mathematically with a double exponential equation, as shown below [33,34].

$$i(t) = AI_L \left(e^{-\alpha t} - e^{-\beta t} \right) \quad (1)$$

where A is the magnitude correction coefficient, I_L is the amplitude of lightning strike, and α and β are the waveform coefficients and stand for the rising and falling time of the lightning impulse, respectively. Different lightning strikes are added to produce both LD and LF transients. The current waveforms of grounding fault and disturbances due to lightning stroke are simulated and displayed in Figure 2c,e, and their frequency spectrums are shown in Figure 2d,f.

Five LDs with different parameters are analyzed. The current lightning waveform 8/20 μ s is used to simulate lightning strokes, and $\alpha = 7.714 \times 10^4$ s, $\beta = 2.49 \times 10^5$ s, $A = 2.331$ [34]. The other parameters of the five scenarios are (1) $I_L = 15$ kA, $L_f = 10$ km; (2) $I_L = 15$ kA, $L_f = 50$ km; (3) $I_L = 5$ kA, $L_f = 100$ km; (4) $I_L = 15$ kA, $L_f = 150$ km; (5) $I_L = 5$ kA, $L_f = 200$ km. Different from the GF waveforms, which keep increasing, the LD waveforms are composed of successive impulses, and their magnitudes tend to be zero after a long duration. The frequency spectrum of LDs also attenuates gradually from lower frequency bands to higher ones. However, unlike those of GFs, the ripples of LD are greater and more irregular.

The parameters of LFs shown in Figure 2e are (1) $I_L = 30$ kA, $L_f = 10$ km, $Z_g = 10 \Omega$; (2) $I_L = 100$ kA, $L_f = 50$ km, $Z_g = 10 \Omega$; (3) $I_L = 30$ kA, $L_f = 100$ km, $Z_g = 20 \Omega$; (4) $I_L = 100$ kA, $L_f = 150$ km, $Z_g = 20 \Omega$; (5) $I_L = 30$ kA, $L_f = 200$ km, $Z_g = 10 \Omega$. The average amplitudes of LFs increase with time due to the grounding component. The spectrum energy of LFs also decrease with frequency, but the ripples are much greater than those of the other two kinds of transient surges.

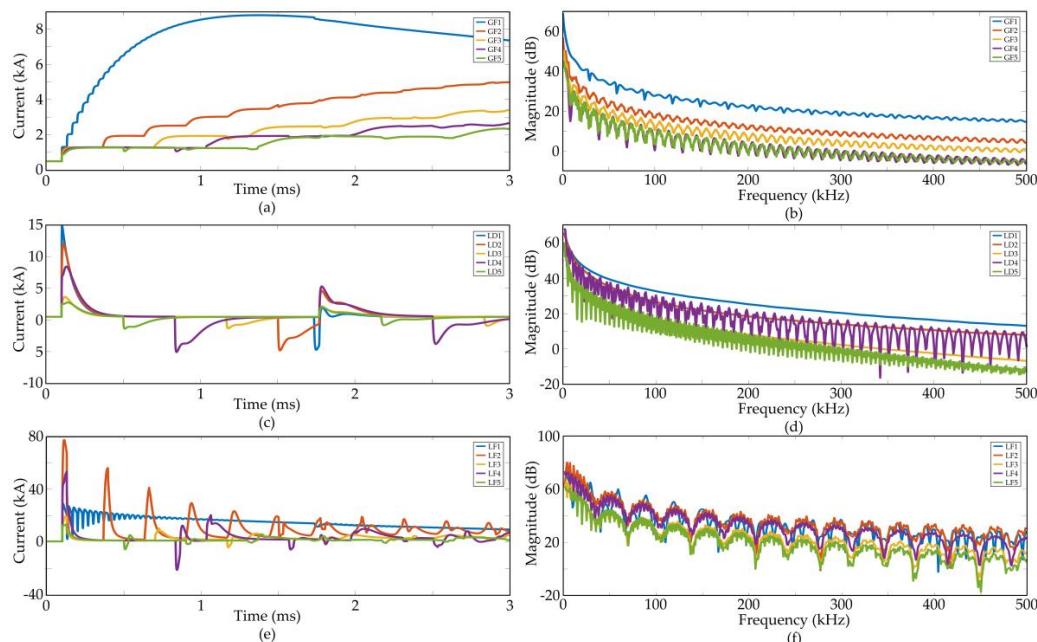


Figure 2. Comparison of different transient surges. (a) current waveforms of single-pole-to-ground faults (GFs); (b) frequency spectrums of GFs; (c) current waveforms of lightning disturbances (LDs); (d) frequency spectrums of LDs; (e) current waveforms of lightning faults (LFs); (f) frequency spectrums of LFs.

Generally, the LF and GF currents increase gradually when the LD current oscillates around normal operating value. An easy method to distinguish fault and disturbance is to integrate the current waveforms, but this method usually needs tens of milliseconds, which is too long for a DC protection to make a judgement. Differences between the three kinds of transient surges can be revealed obviously in the frequency domain even though the time window of transient signal is only a few milliseconds. Appropriate selection of distribution features in frequency spectrum can help to discriminate the transients in extremely short duration.

3. FSE-Based Feature Extraction

3.1. Definition of FSE

Entropy, a convenient tool for measuring the overall disorder of the system, has been effectively used in the field of signal processing [13,35]. If the frequency spectrum of any signal is considered as a system, its distribution can be characterized by entropy. In this paper, the frequency spectrum is generated by Fourier transform. The frequency spectrum is divided equally into m bands. The amplitude of the whole frequency spectrum is treated as a dataset, which is divided into n intervals to calculate the histogram of the frequency spectrum. The number of coefficients of i th ($1 \leq i \leq m$) band x_i in j th ($1 \leq j \leq n$) interval is denoted as x_{ij} , and the probability $p(x_{ij})$ of x_i is calculated according to Equation (2). The definition of FSE H_i is shown Equation (3). Each frequency band can produce one entropy value. A FSE vector H_{FSE} with a size of m will be formed, as shown in Equation (4).

$$p(x_{ij}) = \frac{x_{ij}}{\sum_{j=1}^n x_{ij}} \quad (2)$$

$$H_i = - \sum p(x_{ij}) \log_b(x_{ij}) \quad (j = 1, 2, \dots, n) \quad (3)$$

$$H_{FSE} = [H_1 \ H_2 \ \dots \ H_m] \quad (4)$$

3.2. FSE Representation of Transient Surges

To illustrate the performance of FSE in representing different transient surges, the entropy vectors are analyzed under various scenarios and tested by the simulation model shown in Figure 1. The transient signal is sampled at a rate of 100 kHz, and the time window size is 3 milliseconds. Therefore, the data segment contains only 300 values (100×10^3 sample per second $\times 3 \times 10^{-3}$ millisecond). After Fourier transform, the frequency spectrum from the lowest frequency (except 0 Hz) to 50 kHz is divided into six frequency bands. Each frequency band has a range around 8.5 kHz. Figure 3 illustrates the FSEs of three different transient surges. It is clear that the entropy distribution of these FSE vectors are different from each other. The distribution of the frequency spectrum of GFs are more chaos in lower frequency bands, while the LDs and LFs have more even FSE distributions in whole frequency bands.

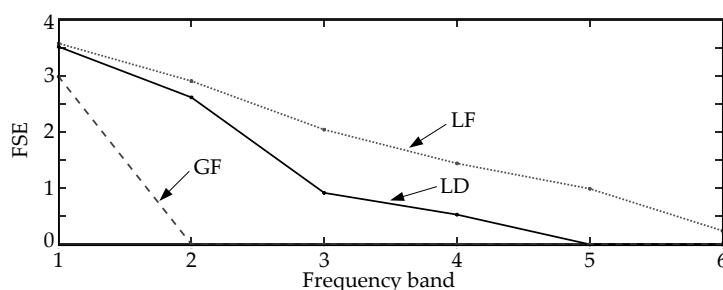


Figure 3. Frequency spectrum entropy (FSE) representation of different transient surges.

The reliability of FSE representation is also tested with different transition resistances and locations: transition resistance equals $0.01\ \Omega$, $1\ \Omega$ and $10\ \Omega$ in grounding faults, and transients occur at 10 km , 100 km , and 200 km , respectively. As illustrated in Figure 4, the trends of FSE distribution vary slightly when parameters change.

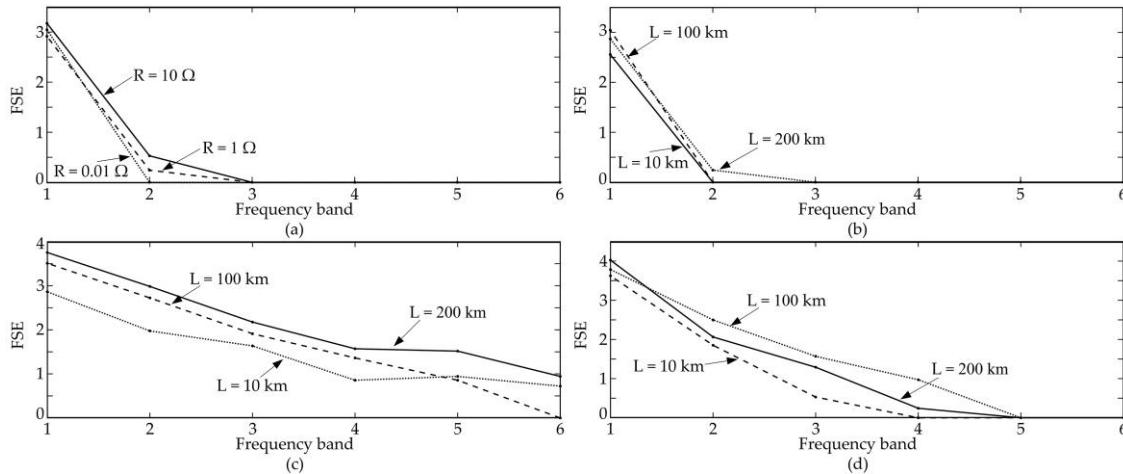


Figure 4. Effect of different factors on FSEs. (a) Effect of transition resistance of GFs; (b) effect of distance of GFs; (c) effect of distance of LFs; (d) effect of distance of LDs.

4. SVM-Based Recognition Method

4.1. Fundamentals of SVM

SVM is a kind of machine learning algorithm based on statistical learning theory, Vapnik–Chervonenkis theory, and structural risk minimization. It has unique advantages in solving small sample, non-linear, and high-dimensional pattern recognition problems and has been widely used in the fields of pattern recognition and regression analysis [25,36].

SVM is a non-probabilistic binary linear classifier. Its main idea is to establish a classification hyperplane as the decision-making plane that maximizes its distance to the data [26]. For linearly separable data with l training samples, the design algorithm for an SVM is reduced to a convex optimization problem, as described in Equation (5), and its binary classification can be represented by Equation (6).

$$\min_w \frac{1}{2} \|w\|^2 \text{ subject to } Y_i(w^T X_i + b) \geq 1 \quad (i = 1, 2, \dots, l) \quad (5)$$

$$\max_{\alpha_i} -\frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l Y_i Y_j \alpha_i \alpha_j \langle X_i, X_j \rangle \quad \text{subject to} \begin{cases} \sum_{i=1}^l Y_i \alpha_i = 0 \\ 0 \leq \alpha_i \quad (i = 1, 2, \dots, l) \end{cases} \quad (6)$$

where $X_i \in R^n$ is the i th feature, $Y_i \in \{-1, 1\}$ is the target label value (binary problem), $w \in R^m$ is the weight vector, α_i are the Lagrange coefficient, $\langle X_i, X_j \rangle$ is the inner product of the input features vector X_i and X_j , b is the bias term, and $d(w, b) = w^T X_i + b = 0$ defines the decision function (classification hyperplane). The weight vector w and the bias term b of decision function can be computed by Equations (7) and (8).

$$w = \sum_{i=1}^l Y_i \alpha_i X_i \quad (7)$$

$$b = \max w^T X_i + \min w^T X_j \quad \text{subject to} \begin{cases} 1 \leq i, j \leq l \\ Y_i = -1, Y_j = 1 \end{cases} \quad (8)$$

In practical applications, most kinds of data are not linearly separable in their original spaces. The original finite-dimensional space is then mapped to a much higher space to generate easier separation. The penalty parameter C and slack variables ε_i are added to the decision function, as shown in Equation (9).

$$\min_w \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l \varepsilon_i \quad \text{subject to} \begin{cases} Y_i(w^T X_i + b) \geq 1 - \varepsilon_i \\ 0 \leq \varepsilon_i \ (i = 1, 2, \dots, l) \end{cases} \quad (9)$$

Such optimization can be represented by a binary classification as shown in Equation (10).

$$\min_{\alpha_i} \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l Y_i Y_j \alpha_i \alpha_j K(X_i, X_j) - \sum_{j=1}^l \alpha_j \quad \text{subject to} \begin{cases} \sum_{i=1}^l Y_i \alpha_i = 0 \\ 0 \leq \alpha_i \leq C \ (i = 1, 2, \dots, l) \end{cases} \quad (10)$$

To amplify the differences or the margins between data, every inner product $\langle X_i, X_j \rangle$ that is related to the features vectors is replaced by a nonlinear kernel function, as shown in Equation (11).

$$K(X_i, X_j) \triangleq \langle \varphi(X_i), \varphi(X_j) \rangle \quad (11)$$

Here, $K(X_i, X_j)$ is the kernel function, φ is the nonlinear mapping. The use of kernel function allows the maximum-margin hyperplane to linearly separate data in transformed higher dimensional feature space. The kernel function is selected to suit the particular classification problem by testing the performance of kernel functions. The most commonly used kernel functions are listed in Table 1.

Table 1. Typical kernel functions.

| Type | Definition |
|------------|---|
| linear | $K(X_i, X_j) = X_i^T X_j$ |
| polynomial | $K(X_i, X_j) = (\gamma X_i^T X_j + r)^p, \gamma > 0$ |
| RBF | $K(X_i, X_j) = \exp(-\gamma \ X_i - X_j\ ^2), \gamma > 0$ |
| sigmoid | $K(X_i, X_j) = \tanh(\gamma X_i^T X_j + r)$ |

Where γ, r and p are kernel parameters.

Though originated from the processing of binary classification, SVM can solve the problem of multi-classification by construction [37–39]. The construction of SVM can be divided into two categories: direct method and indirect method. The direct method is generally completed by modifying the objective function. Such method has a high computational complexity and is a bit difficult in implementations. The indirect method is usually achieved by combining multiple binary classifiers. This solution is simple and easy to be used. “One-to-One” construction is one of the most commonly used indirect construction methods. It designs a SVM between any two types of samples, and determines the type of the unknown sample according to the category scores given by each SVM pair. This construction method greatly reduces the calculation complexity of each classification problem by increasing the number of binary classifiers, and the parallel computation of multiple classifiers improves the overall training speed and the classification accuracy. A “One-to-One” construction method is thus adopted to achieve multi-classification in this research.

4.2. Recognition Method

Combining the advantages of FSE and SVM, a transient surge recognition method is proposed. Its flowchart is shown in Figure 5. Four steps are included in this proposed method.

1. Signal acquisition

Since voltage is controlled in VSC-HVDC systems, the current measurements contain more transients than voltages ones and are thus employed in recognition. To avoid the influence from communications, only local measurements are used.

2. Data processing

A time window of 3 milliseconds is used to capture the starting part of the transient surges. Fourier transform is adopted to generate frequency spectrum. The FSE vector is calculated according to Equation (4).

3. SVM training

Training SVM is crucial for accurate discrimination of faults and disturbances. The structure of SVM is defined by using the “One-to-One” method.

4. Transient recognition

The trained SVM is tested with test samples and used for recognizing different kinds of transients.

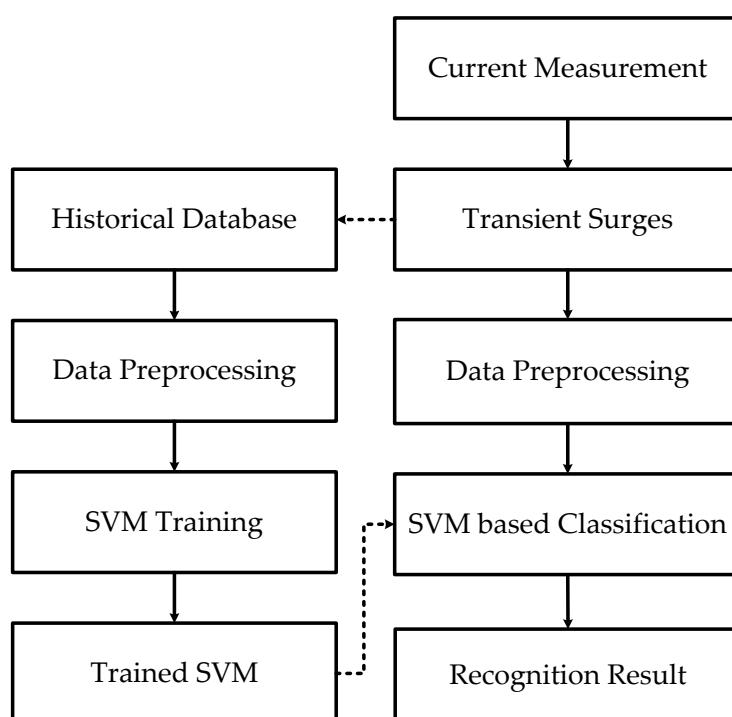


Figure 5. Flow chart of FSE- support vector machine (SVM)-based recognition method.

5. Simulations

5.1. Simulation Model

A two-terminal point-to-point VSC-HVDC system, as shown in Figure 1, is built on the platform of EMTP/PSCAD. Double closed-loop PI control is designed to stabilize the bus voltage. The transmission capacity of this system is 200 MW and the DC bus voltage is ± 200 kV. Coupled overhead transmission lines with frequency-dependent parameters are selected and the length of transmission line is 250 km. Transient current surges are simulated under different scenarios and with different parameters. The simulated transient sources are located equally along the transmission line with an interval of 10 km. The transition resistance of GFs and LFs varies randomly from 0.01Ω to 20Ω . The double exponential waveform, as demonstrated in Equation (1), is used to simulate lightning strokes. A typical current wave shape $8 \mu\text{s}/20 \mu\text{s}$ is adopted with $\pm 10\%$ variations in its parameters [34,40].

The amplitude of lighting strokes varies from 5 kA to 15 kA for LD, and from 30 kA to 100 kA for LF. The sampling rate is 100 kHz and the time window for surge capturing is 3 ms.

5.2. Data Processing

The frequency spectrum of measured current surges are generated by Fourier transform. The whole frequency spectrum (frequency range (0 Hz, 50 kHz]) is divided into 6 frequency bands, and the total amplitude is divided into 30 intervals for FSE calculation. So, the FSE vectors have a size of 6, $E_{FSE} = [E_1, E_2, E_3, E_4, E_5, E_6]$. For each kind of transients, 200 samples are collected, 100 samples of each kind are randomly selected to form the training sample set, and the rest 100 samples are used for testing. Since six-dimensional data cannot be demonstrated graphically, the E_{FSE} vector is decomposed into two vectors: $E_{FSE1} = [E_1, E_2, E_3]$ that represents the lower frequency distributions and $E_{FSE2} = [E_4, E_5, E_6]$ that suggests the higher frequency distributions. Figure 6 shows all samples used for training: 100 samples for each kind of transient surges. The feature map or the space distributions of two decomposed vectors E_{FSE1} and E_{FSE2} are shown in Figure 6a,b, respectively.

As illustrated in Figure 6, the decomposed E_{FSE1} and E_{FSE2} vectors are nonlinearly separable in their three-dimensional spaces. In Figure 6a, the lower frequency features E_{FSE1} of three kinds of transient surges are mixed together, especially, the features of LD and LF. The FSE features E_{FSE2} of LD and LF in higher frequency range are close to each other. A few of the E_{FSE2} of GF mixed with those of LD. Therefore, the original feature vectors E_{FSE} of the three kinds of samples are also linearly inseparable in six-dimensional spaces. A kernel function is thus needed to construct a decision surface in higher dimensional space.

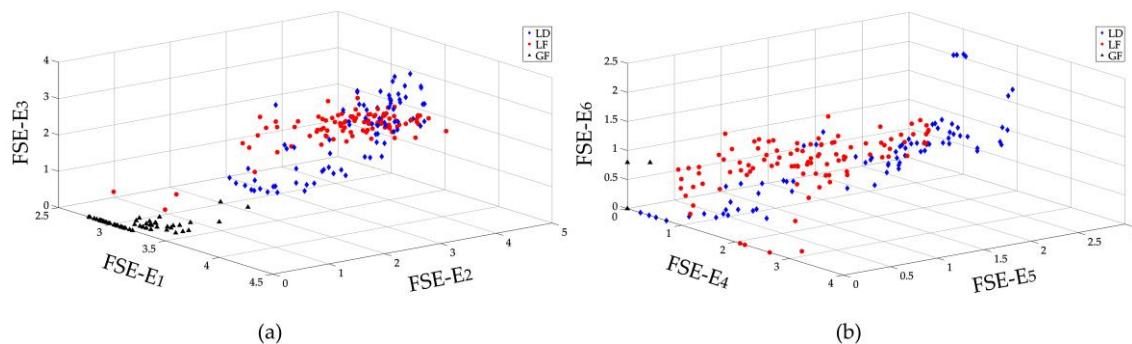


Figure 6. FSE feature map of different transient surges. (a) Space distribution of E_1 vs. E_2 vs. E_3 . (b) space distribution of E_4 vs. E_5 vs. E_6 .

5.3. SVM Training

As aforementioned, the “One-to-One” method is adopted in SVM training. Since there are 3 kinds of transient surges: LD, GF, and LF, three SVMs are employed to construct the following pairs: (i) 1st pair: LD-LF (SVM1), (ii) 2nd pair: LD-GF (SVM2), and (iii) 3rd pair: LF-GF (SVM3). The type of unknown transient can be determined by combining the results of each SVM. For example, an unknown transient surge can be determined to be LD only when both SVM1 and SVM2 produce LD classification results.

The selection of suitable kernel function is quite crucial for an excellent SVM classifier. The rate of correct recognition is used to evaluate the training performance. Four kinds of kernel functions—linear, polynomial, RBF, and sigmoid—are discussed. The K-fold Cross Validation (K-CV) is commonly used to choose parameter combinations to achieve highest classification accuracy, and avoid either over-learning or under-learning. The main idea of K-CV is to divide the original data into K groups, each of which includes both training and testing samples. The highest classification accuracy is taken

as the objective function to determine the parameters of the SVM classifier. Here, K equals to 5 in this research. The mean recognition rates of five-fold cross validations are listed in Table 2.

Table 2. Recognition rates of SVMs with different kernel functions.

| SVM | Kernel Function | | | |
|-----------------------|-----------------|------------|-------|---------|
| | Liner | Polynomial | RBF | Sigmoid |
| Mean recognition rate | SVM1 | 84% | 94% | 95.5% |
| | SVM2 | 100% | 99.5% | 100% |
| | SVM3 | 100% | 100% | 100% |

As illustrated by Table 2, the kernel function RBF can produce higher recognition rates than others. The RBF kernel function is thus selected in FSE-SVM based recognition.

Other parameters, such as penalty parameter C and kernel function parameter γ , are tested by K-CV ($K = 5$). The common practice for parameter selection is to take the relevant parameters within a certain range. Both parameters are performed within a range from 0 to 1000. The values that give highest mean recognition rate are kept and used as the parameters of trained SVM. Table 3 lists all the selected parameters of each SVM and the overall mean recognition rates.

Table 3. Optimal parameter selection results.

| SVM | C | γ | Mean Recognition Rate |
|------------------|-------|----------|-----------------------|
| SVM1 (LD vs. LF) | 5.66 | 1 | 95.5% |
| SVM2 (LD vs. GF) | 0.004 | 0.25 | 100% |
| SVM3 (LF vs. GF) | 0.004 | 0.5 | 100% |

To demonstrate the effectiveness of constructed SVMs, it is useful to illustrate the decision surface of each SVM. But for the proposed FSE vector which is 6 dimensional, only the decision surfaces of some selected features are displayed. The features that are more different from each other are selected for graphically illustrations. The decision surface of two distinguishing features of 3 SVMs that are used in this research are shown in Figures 7–9, respectively.

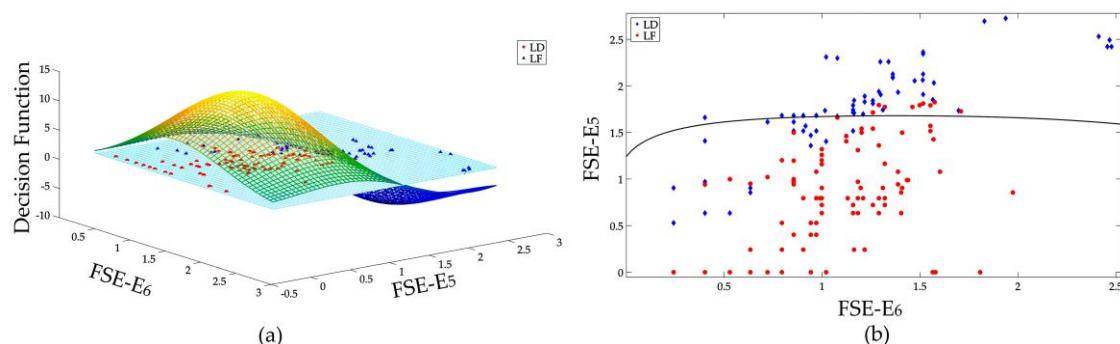


Figure 7. Binary classification of LD and LF, (a) decision surface in 3D space, (b) separation of LD and LF samples in 2D space.

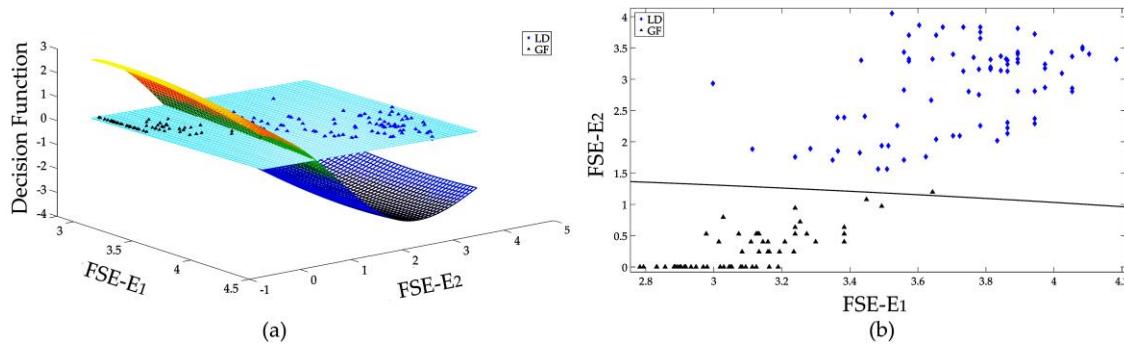


Figure 8. Binary classification of LD and GF, (a) decision surface in 3D space, (b) separation of LD and GF samples in 2D space.

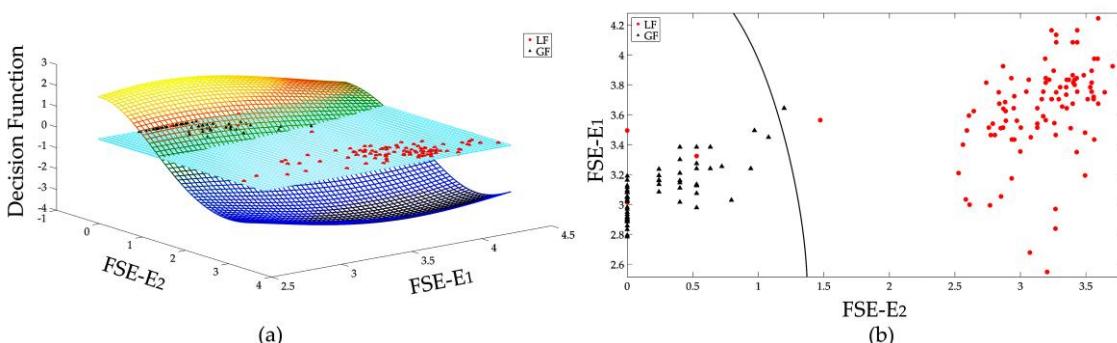


Figure 9. Binary classification of LF and GF, (a) decision surface in 3D space, (b) separation of LF and GF samples in 2D space.

Figure 7a displays the binary classification of LD and LF. As illustrated in Figure 6, the FSE features of LD and LF are quite similar. Two high frequency features: E_5 and E_6 are selected for illustration because their distributions are relatively far away from each other. The intersection of the decision function $d(w, b)$ with the plane of features defines the optimal separation hyperplane, as shown in Figure 7b. By calculating the sign of decision surface $d(w, b)$, the classification of LF and LD can be realized to some extent. As only two features are used for illustration, not all samples are effectively classified. With all 6 features, the classification results can be more effective.

Figures 8 and 9 show the binary classifications of LD vs. GF and LF vs. GF, respectively, through only the feature E_1 and E_2 . As shown in Figure 6, the large amplitude FSE features of GF gather more in the lower frequency range, which is different from those of transients caused by lightning strikes. As demonstrated by Figure 8b, the samples of LD and GF can be effectively classified with only E_1 and E_2 . Also, most samples of LF and GF can be correctly distinguished with only two features E_1 and E_2 . With the whole FSE vectors which includes six features, the GF can be effectively recognized from lightning-caused transients.

Hence, the SVM structure with appropriate kernel functions and parameters can effectively classify different kinds of transient surges.

5.4. Transient Recognition

The performance of the trained SVM is tested with test samples, and the recognition results are shown in Table 4. The GF can be discriminated with 100% recognition rates. Only four LF samples are classified to be LDs, and four LD samples are classified to be LFs. The overall recognition rate of proposed FSE-SVM based method is 97.33%, which shows great potential in protection application.

Table 4. Test results of FSE-SVM method.

| Type | Recognition Rate | Misjudgment | Overall Recognition Rate |
|------|------------------|-------------|--------------------------|
| LD | 96% (4/100) | 4 (4 LF) | |
| LF | 96% (4/100) | 4 (4 LD) | 97.33% (292/300) |
| GF | 100% (100/100) | 0 | |

6. Comparisons

To demonstrate the effectiveness of proposed FSE-SVM-based method, the feature FSE and the classifier SVM are compared with existing popular methods-energy distribution and artificial neural network (ANN), respectively.

6.1. Comparison of Features

Energy distribution is one of the commonly used methods for frequency domain analysis of signals, which has the advantages of simple calculation and intuitive expression. However, the energy distribution heavily depends on the amplitude of transient surges in certain frequency band. The energy distribution defined by Equation (13) is used to characterize the frequency spectrum of transient surges [41]:

$$E_i = \|c_i\| \quad (12)$$

$$E = [E_1 \ E_2 \ \dots \ E_M] \quad (13)$$

The frequency spectrum of transient surge is generated by Fourier transform. The whole frequency spectrum is divided into M bands. The energy E_i of i th frequency band is the norm or square root of the sum of all Fourier coefficients c_i . Here, c_i stands for all of the coefficients in i th frequency band. The M energy value E_i forms an energy distribution vector $\|E\|$.

The energy distribution vector $\|E\|$ is used as the feature, and the same SVM structure is adopted as classifier. The training procedure of SVM is as the same as the one discussed in Section 5. Table 5 shows the recognition results.

Table 5. Test results of energy distribution and SVMs.

| Type | Recognition Rate | Misjudgment | Overall Recognition Rate |
|------|------------------|-----------------|--------------------------|
| LD | 90% (90/100) | 10 (4 GF, 6 LF) | |
| LF | 100% (100/100) | 0 | 92.33% (277/300) |
| GF | 87% (87/1000) | 13 (13 LF) | |

As demonstrated in Table 5, the recognition results of energy representations are lower than the proposed FSE based ones. The overall recognition rate is only 92.33%. The energy-based feature can effectively discriminate LF from other surges. However, it has difficulty in distinguishing faults and disturbances caused by lightning strokes. Only 90% of LDs can be correctly recognized. Among the misjudgments, four LDs are classified as GFs, and six LDs are regarded as LFs. Up to 13% of GF samples are classified as LFs. This might due to the energy attenuation and distortion of transient surges during propagation. The magnitude of the energy value varies a lot. However, the distribution, or disorder, of the frequency spectrum changes a little. When compared with energy based feature, the entropy based spectrum distribution is more effective in representing transient surges.

6.2. Comparison of Classifiers

Back-propagation (BP) ANN, which is a multi-feedforward network trained by error inverse propagation algorithm, is one of the widely used neural network models [42]. Different from a single

SVM that can only distinguish two kinds of samples, a single BP ANN with proper design can realize recognition of multiple types. As the FSE dimension is six and the number of transient types is three, the number of neutrons of input and output layers of ANN are six and three, respectively. To achieve better performance, a lot of experiments are carried out to select the size of hidden layer, and six neutrons are finally chosen. Hyperbolic tangent function and linear function are selected to be the transfer function of hidden layer and output layer, respectively. The training function that based on gradient descent algorithm and dynamic adaptive learning rate is used.

The samples used in Section 5 are also characterized by FSE and recognized by ANN. Table 6 shows the recognition results of FSE-ANN based method.

Table 6. Test results of FSE and artificial neural network (ANN).

| Type | Recognition Rate | Misjudgment | Overall Recognition Rate |
|------|------------------|-------------|--------------------------|
| LD | 95% (95/100) | 5 (5 GF) | |
| LF | 80% (80/100) | 20 (20 LD) | 91.67% (275/300) |
| GF | 100% (100/100) | 0 | |

As shown in Table 6, the overall recognition rate of FSE-ANN based method is a bit lower than that of FSE-SVM based one, which is 97.33%. As the same as SVM based recognition results, misjudgments occur for both LD and LF when ANN is adopted. But more samples are misjudged.

7. Conclusions

This paper proposed a FSE-SVM-based method to distinguish three kinds of most commonly encountered transient surges in HVDC transmission lines. The proposed method can generate effective recognition results and help improving the reliability of protections with relative lower sampling frequency (100 kHz) and extremely short data segment (3 ms). Simulations and comparisons between the energy-based feature and the ANN classifier demonstrate the FSE is stable in characterizing the frequency spectrum of transient surges, and the “One-to-One” SVM structure is simple and effective for training and stable in performance. With training samples from precisely modeled simulation systems, the trained SVM can perform well and respond quickly in practical applications.

Author Contributions: G.L. developed the method and wrote the paper; C.Y. modeled the system and processed the simulated data; Y.L. help improving the performance of algorithm; Y.T. performed the comparisons; J.H. provided some suggestions in improving this manuscript.

Funding: This research is funded by the National Natural Science Foundation of China (Foundation No. 51507008), the Fundamental Research Funds for the Central Universities (No. 2018JBM056), and the National Key R&D Program of China (No. 2017YFB0902800).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Beerten, J.; Cole, S.; Belmans, R. Modeling of multi-terminal VSC HVDC systems with distributed DC voltage control. *IEEE Trans. Power Syst.* **2014**, *29*, 34–42. [[CrossRef](#)]
2. Zhang, L.; Harnefors, L.; Nee, H.P. Interconnection of two very weak AC systems by VSC-HVDC links using power-synchronization control. *IEEE Trans. Power Syst.* **2011**, *26*, 344–355. [[CrossRef](#)]
3. Li, Y.; Gong, Y.F.; Jiang, B. A novel traveling-wave-based directional protection scheme for mtdc grid with inductive DC terminal. *Electr. Power Syst. Res.* **2018**, *157*, 83–92. [[CrossRef](#)]
4. Hao, W.; Mirsaeidi, S.; Kang, X.; Dong, X.; Tzelepis, D. A novel traveling-wave-based protection scheme for LCC-HVDC systems using teager energy operator. *Int. J. Electr. Power Energy Syst.* **2018**, *99*, 474–480. [[CrossRef](#)]
5. Song, G.; Chu, X.; Gao, S.; Kang, X.; Jiao, Z. A new whole-line quick-action protection principle for HVDC transmission lines using one-end current. *IEEE Trans. Power Deliv.* **2015**, *30*, 599–607. [[CrossRef](#)]

6. Chamia, M.; Liberman, S. Ultra high speed relay for EHV/UHV transmission lines—development, design and application. *IEEE Trans. Power Appar. Syst.* **1978**, *97*, 2104–2116. [[CrossRef](#)]
7. Shu, H.C.; Zhang, B.; Zhang, G.B.; Duan, R.M. Identification of lightning disturbance in UHVDC transmission lines using correlation degree based on short time window data. In *Materials Science and Information Technology*; Zhang, C.S., Ed.; Trans Tech Publications Ltd.: Stafa-Zurich, Switzerland, January 2012; Volume 433–440, pp. 3787–3791.
8. Kong, F.; Hao, Z.; Zhang, S.; Zhang, B. Development of a novel protection device for bipolar HVDC transmission lines. *IEEE Trans. Power Deliv.* **2014**, *29*, 2270–2278. [[CrossRef](#)]
9. Jin, J.X. Protection of HVDC transmission lines based on wavelet transformation and analysis of energy spectrum. In Proceedings of the 2013 2nd International Symposium on Instrumentation and Measurement, Sensor Network and Automation, Toronto, ON, Canada, 23–24 December 2013; IEEE: New York, NY, USA, 2013; pp. 180–185.
10. Srikanth, P.; Chandel, A.K.; Naik, K.A. HVDC system fault identification using S-transform approach. In Proceedings of the 2010 International Conference on Power, Control and Embedded Systems, Allahabad, India, 29 November–1 December 2010; pp. 1–6.
11. Stockwell, R.G.; Mansinha, L.; Lowe, R.P. Localization of the complex spectrum: The S transform. *IEEE Trans. Signal. Process.* **1996**, *44*, 998–1001. [[CrossRef](#)]
12. Song, G.; Cai, X.; Gao, S.; Suonan, J.; Li, G. Natural frequency based protection and fault location for VSC-HVDC transmission lines. In Proceedings of the International Conference on Advanced Power System Automation and Protection, Beijing, China, 16–20 October 2011; pp. 177–182.
13. Luo, G.; Lin, Q.; Zhou, L.; He, J. Recognition of traveling surges in HVDC with wavelet entropy. *Entropy* **2017**, *19*, 184. [[CrossRef](#)]
14. Sharma, R.; Pachori, R.B.; Acharya, U.R. Application of entropy measures on intrinsic mode functions for the automated identification of focal electroencephalogram signals. *Entropy* **2015**, *17*, 669–691. [[CrossRef](#)]
15. Luo, G.; Zhang, D.; Koh, Y.; Ng, K.; Leong, W. Time-frequency entropy-based partial-discharge extraction for nonintrusive measurement. *IEEE Trans. Power Deliv.* **2012**, *27*, 1919–1927. [[CrossRef](#)]
16. Lin, S.; Gao, S.; He, Z.; Deng, Y. A pilot directional protection for HVDC transmission line based on relative entropy of wavelet energy. *Entropy* **2015**, *17*, 5257–5273. [[CrossRef](#)]
17. Li, H.-F.; Wang, G.; Zhao, J.-C. Study on characteristics and identification of transients on transmission lines caused by indirect lightning stroke. *Proc. CSEE* **2004**, *24*, 114–119.
18. Wu, T.; Bajwa, W.U. Learning the nonlinear geometry of high-dimensional data: Models and algorithms. *IEEE Trans. Signal. Process.* **2015**, *63*, 6229–6244. [[CrossRef](#)]
19. Samantaray, S.R.; Dash, P.K.; Upadhyay, S.K. Adaptive kalman filter and neural network based high impedance fault detection in power distribution networks. *Int. J. Electr. Power Energy Syst.* **2009**, *31*, 167–172. [[CrossRef](#)]
20. Khorashadi-Zadeh, H. Artificial neural network approach to fault classification for double circuit transmission lines. In Proceedings of the 2004 IEEE/PES Transmision and Distribution Conference and Exposition, Sao Paulo, Brazil, 8–11 November 2004; pp. 859–862.
21. Butler, K.L.; Momoh, J.A. A neural net based approach for fault diagnosis in distribution networks. In Proceedings of the 2000 Winter Meeting IEEE Power Engineering Society, Singapore, Singapore, 23 January–27 January 2000; Volume 1, pp. 1275–1278.
22. Praveenkumar, T.; Sabhrish, B.; Saimurugan, M.; Ramachandran, K.I. Pattern recognition based on-line vibration monitoring system for fault diagnosis of automobile gearbox. *Measurement* **2018**, *114*, 233–242. [[CrossRef](#)]
23. Jana, S.; De, A. A novel zone division approach for power system fault detection using ANN-based pattern recognition technique. *Can. J. Electr. Comput. Eng.* **2018**, *40*, 275–283.
24. Chen, K.; Hu, J.; He, J. Detection and classification of transmission line faults based on unsupervised feature learning and convolutional sparse autoencoder. *IEEE Trans. Smart Grid* **2018**, *9*, 1748–1758.
25. Geng, P.; Song, J.; Xu, C.; Zhao, Y. Fault pattern recognition method for the high voltage circuit breaker based on the incremental learning algorithms for SVM. In Proceedings of the 2016 International Conference on Condition Monitoring and Diagnosis (CMD), Xi'an, China, 25–28 September 2016; pp. 693–696.
26. Chen, M.-Y.; Hu, G.; Zhai, J.-Q. High impedance fault detection using hilbert transform and least square support vector machine for distribution feeders. *Int. Rev. Electr. Eng.* **2012**, *7*, 4013–4020.

27. Chang, B.; Cwikowski, O.; Barnes, M.; Shuttleworth, R. Point-to-point two-level converter system faults analysis. In Proceedings of the 7th IET International Conference on Power Electronics, Machines and Drives (PEMD), Manchester, UK, 8–10 April 2014; Institution of Engineering and Technology: Manchester, UK, 2014.
28. Yang, J.; Zheng, J.-C.; Tang, G.-F.; He, Z.-Y. Grounding design analysis of VSC-HVDC system. *Proc. CSEE* **2010**, *30*, 14–19.
29. Li, H.; Wang, G.; Liao, Z. Distinguish between lightning strikes and faults using wavelet-multi resolution signal decomposition. In Proceedings of the 2004 Eighth IEE International Conference on Developments in Power System Protection, Amsterdam, The Netherlands, 5–8 April 2004; Volume 8, pp. 80–83.
30. Wang, G.; Li, H.-F.; Zhao, J.-C.; Wu, M. Identification of transients on transmission lines caused by direct lightning strokes based on multiresolution signal decomposition. *Proc. CSEE* **2004**, *24*, 139–144.
31. Waring, J.B. The effects of lightning on overhead lines. In Proceedings of the IEE Seminar on Lightning Protection for Overhead Line Systems, London, UK, 11 December 2000.
32. Eriksson, A.J.; Meal, D.V.; Stringfellow, M.F. Lightning-induced overvoltages on overhead distribution lines. *IEEE Power Eng. Rev.* **1982**, *2*, 41. [CrossRef]
33. Thottappillil, R.; Uman, M.A. Comparison of lightning return-stroke models. *J. Geophys. Res. Atmos.* **1993**, *98*, 22903–22914. [CrossRef]
34. National Standards Authority of Ireland (NSAI). IEC 61000-4-5 Testingand measurement techniques-surgeimmunity test. In *Electromagnetic Compatibility (EMC)*; CENELEC: Dublin, Ireland, 2014.
35. Chen, B.D.; Wang, J.J.; Zhao, H.Q.; Principe, J.C. Insights into entropy as a measure of multivariate variability. *Entropy* **2016**, *18*, 196. [CrossRef]
36. Mousavi, M.; Butler-Purry, K. A novel condition assessment system for underground distribution applications. *IEEE Trans. Power Syst.* **2009**, *24*, 1115–1125. [CrossRef]
37. Guo, J.; Chen, Y.; Zhu, M.; Wang, S.; Liu, X. An efficient support vector machine algorithm for solving multi-class pattern recognition problems. In Proceedings of the 2010 Second International Conference on Computer Modeling and Simulation, Sanya, China, 22–24 January 2010; pp. 461–465.
38. Hsu, C.-W.; Lin, C.-J. A comparison of methods for multiclass support vector machines. *IEEE Trans. Neural Netw.* **2002**, *13*, 415–425. [PubMed]
39. Schwenker, F. Hierarchical support vector machines for multi-class pattern recognition. In Proceedings of the Fourth International Conference on Knowledge-Based Intelligent Engineering Systems and Allied Technologies, Brighton, UK, 30 August–1 September 2000; Volume 562, pp. 561–565.
40. International Electrotechnical Commission. *Protection against Lightning: Risk Management*; IEC: Geneva, Switzerland, 2006.
41. Luo, G.; Zhang, D.; Tseng, K.J.; He, J. Impulsive noise reduction for transient earth voltage-based partial discharge using wavelet-entropy. *IET Sci. Meas. Technol.* **2016**, *10*, 69–76. [CrossRef]
42. Li, X.; Xu, J. The improvement of bp artificial neural network algorithm and its application. In Proceedings of the 2010 International Conference on E-Business and E-Government, Guangzhou, China, 7–9 May 2010; pp. 2568–2571.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).