

Article

# Power Quality Disturbance Classification Using the S-Transform and Probabilistic Neural Network

Huihui Wang <sup>1,2,\*</sup>, Ping Wang <sup>1</sup> and Tao Liu <sup>3</sup>

<sup>1</sup> School of Electrical Engineering and Automation, Tianjin University, Tianjin 300072, China; pingw@tju.edu.cn

<sup>2</sup> School of Control and Mechanical Engineering, Tianjin Chengjian University, Tianjin 300384, China

<sup>3</sup> School of Electrical Engineering and Automation, Tianjin Polytechnic University, Tianjin 300387, China; taoliu@tju.edu.cn

\* Correspondence: homeworldt2014@gmail.com; Tel.: +86-135-122-71448

Academic Editor: Birgitte Bak-Jensen

Received: 19 September 2016; Accepted: 9 January 2017; Published: 17 January 2017

**Abstract:** This paper presents a transient power quality (PQ) disturbance classification approach based on a generalized S-transform and probabilistic neural network (PNN). Specifically, the width factor used in the generalized S-transform is feature oriented. Depending on the specific feature to be extracted from the S-transform amplitude matrix, a favorable value is determined for the width factor, with which the S-transform is performed and the corresponding feature is extracted. Four features obtained this way are used as the inputs of a PNN trained for performing the classification of 8 disturbance signals and one normal sinusoidal signal. The key work of this research includes studying the influence of the width factor on the S-transform results, investigating the impacts of the width factor on the distribution behavior of features selected for disturbance classification, determining the favorable value for the width factor by evaluating the classification accuracy of PNN. Simulation results tell that the proposed approach significantly enhances the separation of the disturbance signals, improves the accuracy and generalization ability of the PNN, and exhibits the robustness of the PNN against noises. The proposed algorithm also shows good performance in comparison with other reported studies.

**Keywords:** transient power quality; S-transform; width factor; feature extraction; probabilistic neural network (PNN)

## 1. Introduction

In recent years, with the development and increasing implementation of distributed generation and micro-grids, large numbers of high-speed switching devices and non-linear converters are integrated into the power system. They are sources of power quality (PQ) disturbances such as sag, swell, interruption, flicker, harmonic, and oscillatory transients, etc. These voltage disturbances degrade the end user's experience, and more importantly, it may damage modern precision devices and seriously affect the production efficiency of the manufacturing industry. Owing to the facts above, the problem of PQ disturbance detection and classification has received the attention of many scholars [1].

To understand the PQ disturbance, the time-frequency characteristics of disturbances signals need to be extracted and analyzed. Various signal processing algorithms, for example, Fourier transform (FT), short time Fourier transform (STFT) [2], wavelet transform (WT) [3,4], and Stockwell transform (S-transform) [5,6] have been utilized for this purpose. FT has the advantages of requiring a relatively less amount of calculation and broad applicability, however it has the disadvantages of spectrum leakage and fence effect, and it is limited to only detecting steady state signals. STFT transforms a

signal into a two-dimensional complex function which reveals the frequency characteristics of the signal at each time instant, however its time window is relatively fixed and it is difficult to balance the frequency domain resolution and time domain resolution. WT can accurately detect the singular point and the time instant at which the PQ disturbance occurs, however it is time consuming and susceptible to noises. To overcome the disadvantages of the WT, Stockwell et al. [7] proposed the S-transform, which is conceptually a hybrid of STFT and WT. Compared with WT, S-transform can be treated as an extension of WT, on the other hand, S-transform can be treated as STFT with a variable window. S-transform has superior characteristics to WT and STFT, and it has good time-frequency focusing characteristics, satisfactory time-frequency resolution, and noise immunity. The features extracted with the S-transform of the disturbance signals are directly useable in intelligent algorithms for classification [8,9].

To make the S-transform more effective for different particular application focuses, various forms of S-transform have been proposed and examined. For example, to enable a flexible regulation of the time resolution and frequency resolution, a width factor is introduced into S-transform, which makes the width of Gaussian window adjustable. To have a better control over the scale and the shape of the analyzing window, a parameter  $P$  is introduced into the Gaussian window function [10–12]. Moreover, in [13], the frequency domain is segmented into domains of different frequencies, low, medium, and high. Within in different frequency domain, the width factor of the S-transform is different, which brings different time-frequency resolution the corresponding frequency domain, so the accuracy of PQ classification can be improved. In mechanical engineering, to analyze the vibration data, asymmetrical window function instead of the regular symmetrical one is used in S-transform [14]. To improve the ability for resolving signals whose frequency changes with time, a complex Gaussian window function is adopted in S-transform [15].

For PQ disturbance classification, it has been proven that S-transform-based approaches are effective, but efforts are needed to make them more favorable [16]. Conventionally, the S-transform used for this purpose takes its basic form—a symmetrical real Gaussian window function without a width factor or with a width factor but it takes a fixed value throughout. This paper studies how to make the S-transform more applicable to PQ disturbance classification by investigating its role in the S-transform-based classification approach. Specifically, an S-transform with a width factor is used; the determination of the value of the width factor is handled in a systemic view of the PQ disturbance classification problem itself. In other words, the width factor is treated as a variable. Its value is made to be feature oriented considering the fact that different features react to the time resolution or the frequency resolution differently.

As for the classification analysis, PNN is adopted in our study. The considerations are briefly given as follows. So far, various intelligent algorithms have been used to classify the PQ disturbances, for example, artificial neural network (ANN) [17], fuzzy logic [18], support vector machine (SVM) [19], etc. Artificial neural networks are mainly used for pattern matching, classification, function approximation, optimization, and data clustering. Neural networks have obvious advantages and have been widely used in PQ classification and fault diagnosis [20,21]. Multi-layer perceptron (MLP) and PNN are the typical forward neural network structures. Compared with MLP, PNN has many advantages such as simple structure, fast training rate, high accuracy, good generalization ability, and robustness, moreover, the training process of PNN is a single-pass network training stage without any iteration for weight adjustment, and it can be easily retrained to adapt to any network topology changes. In the comparison between PNN and the other two well-known neural networks i.e., feed forward multilayer (FFML) back propagation and learning vector quantization (LVQ), PNN provides better performance in terms of the classification results [16,19,22,23]. Based on the advantages of PNN and the characteristic of PQ disturbances, PNN is considered to be more favorable for PQ disturbances classification.

In summary, a PQ disturbance classification approach based on S-transform with a feature oriented regulatory and PNN is proposed and examined. Contents of sections below include: firstly, four features are selected through investigating the impacts of the width factor on the

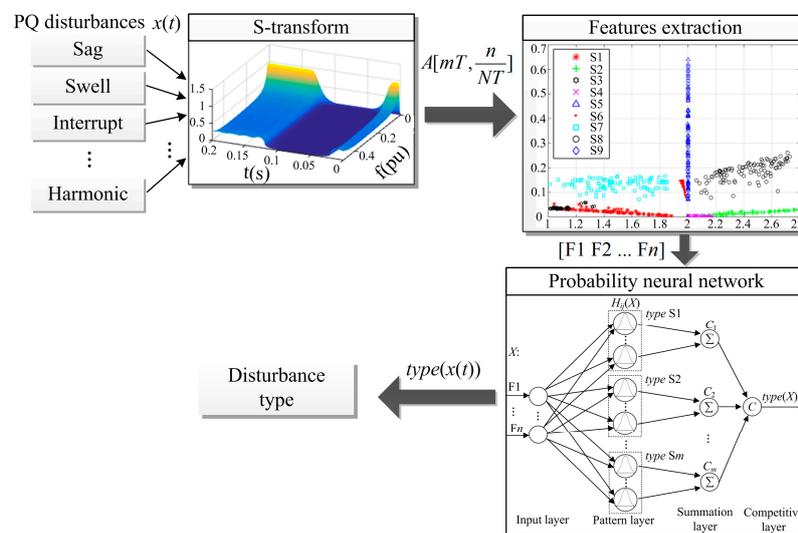
S-transform-Amplitude matrix of eight disturbance signals and the behavior of signal separation based on the features extracted; secondly, favorable value of the width factor used to extract each of the four features from the S-transform-amplitude matrix of disturbance signals is determined by examining the impacts of the width factor on the performance of PNN used for disturbance classification, and finally the proposed approach is tested, showing that it has higher classification accuracy and performs well with and without the noises.

## 2. Power Quality Disturbance Classification Based on Generalized S-Transform and Probabilistic Neural Network

In this section, first, the PQ disturbance classification algorithm based on S-transform and PNN is briefly described. Then, details of each functional unit of the algorithm are given, including the definition of eight disturbance signal types and the normal sinusoidal signal used in this study, mathematic description of S-transform without and with a width factor, definition of signal features and corresponding extraction formulation, and PNN. Finally, work of the authors is briefly introduced. Specifically, it is the implementation of the S-transform with a feature oriented optimal width factor to realize effective classification of the nine types of signals with a PNN.

### 2.1. Method Overview

The PQ disturbance classification based on S-transform and PNN is illustrated in Figure 1. It contains three steps: (1) S-transform, by which the time and frequency information of the given PQ disturbance signal are obtained; (2) feature extraction, by which the customized feature vectors are further extracted from the time and frequency information obtained in step (1); (3) disturbance classification with PNN, the input of which is the feature vector of the PQ disturbance signal and the output of which is the disturbances classification result.



**Figure 1.** The schematic diagram of power quality (PQ) disturbance classification using S-transform and probabilistic neural network (PNN).

### 2.2. Disturbance Signal Types

IEEE Standard 1159 and the European standard EN50160 give definitions for various PQ disturbances and the corresponding suggested threshold settings [24,25]. For instance, Figure 2 shows the disturbance defined as a sag in the mentioned standards, which occurred to a normal sinusoidal voltage during the time approximately from 0.02 s to 0.14 s. The disturbance signal was captured on 21 March 2014 in Jurong East which located in the West Region of Singapore. The official source said it was due to a customer installation fault at Jurong Gateway Road. The PQ will be

considered poor if the voltage magnitude is lower than the threshold set by the standard. In this specific case, the threshold should be 90% of the normal voltage magnitude.

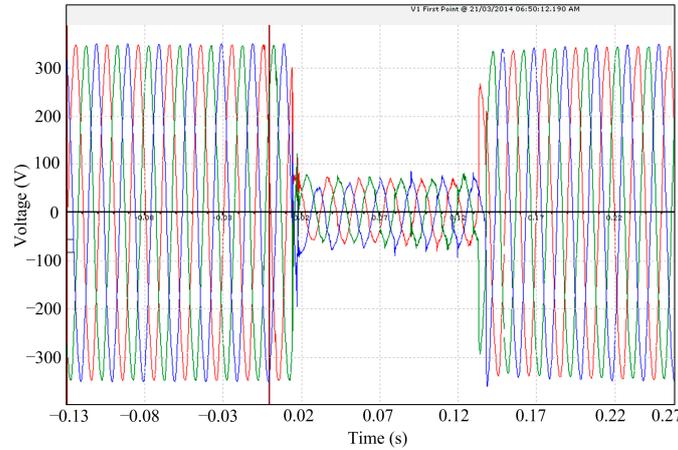


Figure 2. The waveform of sag captured from actual power system [26].

Table 1 lists the PQ disturbance types considered in this paper. Each disturbance type is described with an equation with parameters, with which the disturbance generated comply with the corresponding disturbance defined in IEEE standard 1159-1995 [24]. S1, S2, . . . , and S8 denote sag, swell, interruption, flicker, oscillatory transient, harmonic, sag & harmonic, and swell & harmonic, respectively, S9 denotes normal signal. Where  $f = 50$  Hz;  $\omega = 2\pi f$ ;  $T = 1/f$ . The parameters for the equations such as  $\alpha$ ,  $\alpha_3$ ,  $\alpha_5$ ,  $\alpha_7$ ,  $t_1$ ,  $t_2$ ,  $\alpha_f$ ,  $\beta$ ,  $\tau$ ,  $f_n$  are determined randomly within the thresholds.

Table 1. Mathematical model of PQ disturbances.

Symbol	Type of Disturbance Signal	Equations	Parameters
S1	Sag	$y(t) = (1 - \alpha(u(t - t_1) - u(t - t_2))) \sin(\omega t)$	$0.1 \leq \alpha \leq 0.9, T \leq t_2 - t_1 \leq 9T, t_1 \leq t_2$
S2	Swell	$y(t) = (1 + \alpha(u(t - t_1) - u(t - t_2))) \sin(\omega t)$	$0.1 \leq \alpha \leq 0.8, T \leq t_2 - t_1 \leq 9T$ $u(t) = \begin{cases} 1 & t \geq 0 \\ 0 & t < 0 \end{cases}$
S3	Interruption	$y(t) = (1 - \alpha(u(t - t_1) - u(t - t_2))) \sin(\omega t)$	$0.9 \leq \alpha \leq 1, T \leq t_2 - t_1 \leq 9T$
S4	Flicker	$y(t) = (1 + \alpha_f \sin(\beta \omega t)) \sin(\omega t)$	$0.1 \leq \alpha_f \leq 0.2, 5\text{Hz} \leq \beta \leq 20\text{Hz}$
S5	Oscillatory transient	$y(t) = \sin(\omega t) + \alpha e^{-(t-t_1)/\tau} \sin \omega_n (t - t_1)(u(t_2) - u(t_1))$	$0.1 \leq \alpha \leq 0.8, 0.5T \leq t_2 - t_1 \leq 3T,$ $8\text{ms} \leq \tau \leq 40\text{ms}, 300\text{Hz} \leq f_n \leq 900\text{Hz}$
S6	Harmonic	$y(t) = \alpha_1 \sin(\omega t) + \alpha_3 \sin(3\omega t) + \alpha_5 \sin(5\omega t) + \alpha_7 \sin(7\omega t)$	$0.05 \leq \alpha_3 \leq 0.15, 0.05 \leq \alpha_5 \leq 0.15,$ $0.05 \leq \alpha_7 \leq 0.15, \sum \alpha_i^2 = 1$
S7	Sag and harmonic	$y(t) = (1 - \alpha(u(t - t_1) - u(t - t_2))) (\alpha_1 \sin(\omega t) + \alpha_3 \sin(3\omega t) + \alpha_5 \sin(5\omega t))$	$0.1 \leq \alpha \leq 0.9, T \leq t_2 - t_1 \leq 9T,$ $0.05 \leq \alpha_3 \leq 0.15, 0.05 \leq \alpha_5 \leq 0.15,$ $\sum \alpha_i^2 = 1$
S8	Swell and harmonic	$y(t) = (1 + \alpha(u(t - t_1) - u(t - t_2))) (\alpha_1 \sin(\omega t) + \alpha_3 \sin(3\omega t) + \alpha_5 \sin(5\omega t))$	$0.1 \leq \alpha \leq 0.8, T \leq t_2 - t_1 \leq 9T,$ $0.05 \leq \alpha_3 \leq 0.15, 0.05 \leq \alpha_5 \leq 0.15,$ $\sum \alpha_i^2 = 1$
S9	Normal sine	$y(t) = \sin(\omega t)$	-

### 2.3. Generalized S-Transform and Feature Extraction

#### 2.3.1. Generalized S-Transform

The continuous S-transform of signal  $x(t)$  is described as:

$$S(\tau, f) = \int_{-\infty}^{\infty} x(t) w(\tau - t, f) e^{-i2\pi ft} dt \quad (1)$$

where  $f$  is the frequency of signal  $x(t)$ .  $w(\tau - t, f)$  is Gaussian window function, it is defined as:

$$w(\tau - t, f) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(\tau-t)^2}{2\sigma^2}} \quad (2)$$

The S-transform matrix is a complex matrix whose rows pertain to frequency and columns to time. According to the Heisenberg principle, the time and frequency resolution cannot be improved at the same time. The intent of using the Gaussian window function in Equation (1) is to allow a better balancing between the time resolution and the frequency resolution. In Equation (2),  $\tau$  is the point at the time axis at which the center of  $w(\tau - t, f)$  sits.  $\sigma$  is a scale factor, it is defined as:

$$\sigma = 1/|f| \quad (3)$$

To make the Gaussian window function more effective, a width factor  $\lambda$  is introduced to Equation (3):

$$\sigma = \lambda/|f| \quad (4)$$

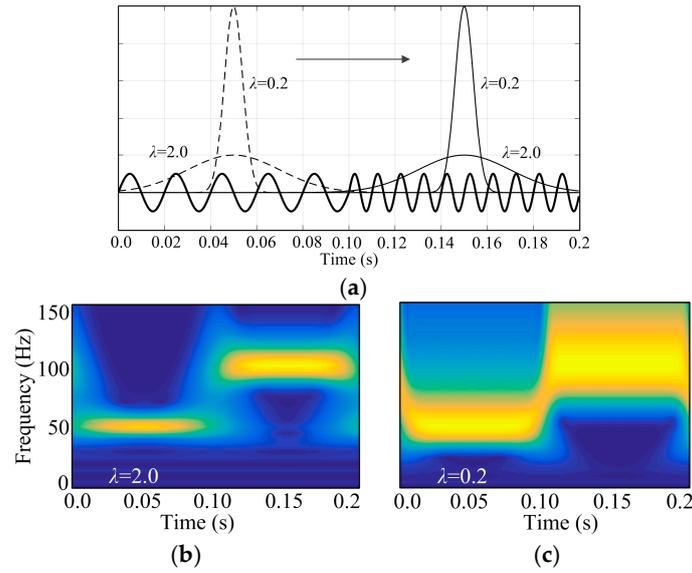
Changing the value of  $\lambda$  adjusts the value of scale factor  $\sigma$  and therefore makes the width of Gaussian window function adjustable. If a value greater than 1 is assigned to  $\lambda$ , for the same frequency point, Gaussian window will be wider than in the case where  $\lambda$  takes a value of 1. Thus, higher frequency resolution can be achieved in time-frequency domain. Inversely, if a value less than 1 is assigned to  $\lambda$ , for the same frequency point, Gaussian window will be narrower than in the case where  $\lambda$  takes a value of 1. Thus, higher time resolution can be achieved in time-frequency domain.

For instance, Figure 3 shows the effect of width factor on the time-frequency analysis of S-transform. The source signal, as shown in Figure 3a consists of two sinusoidal signals of different frequencies, 50 Hz for the signal from time 0 to 0.1 s, 100 Hz for the signal from 0.1 to 0.2 s. Width factors under examination are 0.2 and 2.0. In the case of  $\lambda > 1.0$ , the Gaussian window will be wider and cover more line frequency cycles, and this will help to improve the accuracy of frequency analysis of the signal. However, a wider Gaussian window will not make it easy to identify the time instant at which the frequency of the signal changes. In the case of  $\lambda < 1.0$ , the Gaussian window will be narrower and cover fewer line frequency cycles, this will decrease the accuracy of frequency analysis of the signal, but a narrower Gaussian window will make it easy to identify the time instant at which the frequency of the signal changes.

The result presented in Figure 3b shows that by using a width factor greater than 1.0, the frequency of two sine signals can be satisfactorily identified, as seen the two thin frequency bands in sharp yellow, however, this brings more uncertainty on the time of the signal changes because the changing part of the two frequency bands is not clear. This indicates a width factor greater than 1.0 helps to achieve the better frequency resolution, but the time resolution will get worse. On the contrary, the result presented in Figure 3c shows that using a width factor less than 1.0, the frequency of two sine signals cannot be clearly identified, as seen the two thick frequency bands in sharp yellow, however, the changing part of the two frequency bands is clearly identified. This indicates a width factor less than 1.0 results unsatisfactory frequency resolution, but the time resolution will get better.

Substituting Equation (4) into Equation (1), the continuous generalized S-transform takes the form:

$$S(\tau, f) = \int_{-\infty}^{\infty} x(t) \frac{|f|}{\lambda\sqrt{2\pi}} e^{-\frac{(t-\tau)^2 f^2}{2\lambda^2}} e^{-i2\pi ft} dt \quad (5)$$



**Figure 3.** Time-frequency analysis results with different setting value of the width factor. (a) The comparison of the Gaussian windows with different width factor; (b) Time-frequency spectrum with  $\lambda = 2.0$  and (c) Time-frequency spectrum with  $\lambda = 0.2$ .

Then, with the definition  $\tau = mT, f = n/NT$ , the discrete generalized S-transform expression can be obtained:

$$\begin{cases} S[mT, \frac{n}{NT}] = \sum_{k=0}^{N-1} X[\frac{k+n}{NT}] e^{-\frac{2\pi^2 \lambda^2 k^2}{n^2}} e^{i2\pi mk} & n \neq 0 \\ S[mT, 0] = \frac{1}{N} \sum_{k=0}^{N-1} X[\frac{k}{NT}] & n = 0 \end{cases} \quad (6)$$

where  $k, m, n = 0, 1, \dots, N - 1$ ,  $N$  is the total number of sampling points.  $T$  is the time interval between two consecutive sampling points.  $X$  is the discrete Fourier transform.

Further, for the purpose of extracting the features of disturbance signals, the S-transform-Amplitude (STA) matrix is calculated as follows:

$$A[mT, \frac{n}{NT}] = |S[mT, \frac{n}{NT}]| \quad m, n = 0, 1, \dots, N - 1 \quad (7)$$

### 2.3.2. Feature Extraction

Feature extraction is one of the essential steps of PQ disturbance classification. Below are definitions of 10 features extracted from STA matrix.

F1: maximum amplitude of  $TmA$ -plot (time-maximum amplitude plot), and it has:

$$F1 = \max\{TmA(m)\} \quad (8)$$

where  $TmA$ -plot is maximum amplitude versus time by searching columns of STA matrix given in Equation (7) at every frequency,  $TmA(m) = \max\{A[mT, \frac{n}{NT}]\}$ .

F2: minimum amplitude of the  $TmA$ -plot:

$$F2 = \min\{TmA(m)\} \quad (9)$$

F3: mean value of the  $TmA$ -plot:

$$F3 = \frac{1}{N} \sum_{m=1}^N TmA(m) \quad (10)$$

F4: standard deviation of  $TmA$ -plot:

$$F4 = \sqrt{\frac{1}{N} \sum_{m=1}^N (TmA(m) - F3)^2} \quad (11)$$

F5: the summation of the maximum and minimum of the  $TmA$ -plot:

$$F5 = F1 + F2 \quad (12)$$

F6: the standard deviation of  $FmA$ -plot (frequency maximum amplitude plot, which is maximum amplitude versus frequency by searching the rows of STA matrix at each time instant) for frequencies above four times the line frequency. It has:

$$F6 = \sqrt{\frac{1}{N} \sum_{n=1}^N (FmA(n) - \overline{FmA(n)})^2} \quad (13)$$

where  $FmA(n) = \max\{A[mT, \frac{n}{NT}]\}$  with  $\frac{n}{NT} > 200$ ,  $\overline{FmA(n)}$  is the mean value of  $FmA(n)$ .

F7: the subtraction of the maximum and minimum of the  $FmA(n)$ :

$$F7 = \max(FmA(n)) - \min(FmA(n)) \quad (14)$$

F8: the Skewness of  $FmA$ -plot:

$$F8 = \frac{1}{(N-1)F6^3} \sum_{n=1}^N (FmA(n) - \overline{FmA(n)})^3 \quad (15)$$

F9: the Kurtosis of  $FmA$ -plot:

$$F9 = \frac{1}{(N-1)F6^4} \sum_{n=1}^N (FmA(n) - \overline{FmA(n)})^4 \quad (16)$$

F10: standard deviation of  $mr$  plot, where  $mr$  is time-amplitude plot determined by the frequency which has maximum amplitude in STA matrix above the frequency of 200 Hz, and it has

$$F10 = \sqrt{\frac{1}{N} \sum_{m=1}^N (mr(m) - \overline{mr(m)})^2} \quad (17)$$

where  $mr(m) = A[mT, \frac{n_{mr}}{NT}]$ , and  $n_{mr} = \underset{n}{\operatorname{argmax}}\{A[mT, \frac{n}{NT}]\}$  with  $\frac{n}{NT} > 200$ .

Apparently, when more features are used, a better classification effect may be expected. However, increasing the number of features results in a substantial increase of the calculation time as the scale of ANN used for disturbances classification will be increased. Therefore, the number of features used for the PQ disturbance classification should be as few as possible without obviously decreasing the classification accuracy. In this paper, four features are chosen according to the characteristic of disturbance and width factor, which are presented in Section 3.1.

#### 2.4. Probabilistic Neural Network

As shown in Figure 1, a vector of features selected is used to be the input of PNN. The output of PNN will be the disturbance type. Figure 4 shows the schematic diagram of PNN. It contains four layers: input layer, pattern layer, summation layer, and competitive layer. The function of each layer is as follows.

Pattern layer calculates Euclidean distance between the feature vectors of PQ disturbance testing sample  $X$  and every PQ disturbance training sample  $X_{ij}$ , respectively. Below is the equation:

$$H_{ij}(X) = \frac{1}{(2\pi)^{n/2}\delta^n} e^{-\frac{\|X-X_{ij}\|^2}{2\delta^2}} \quad (18)$$

where  $X = [F1 \ F2 \ F3 \ \dots \ Fn]^T$  is the feature vector of an PQ disturbance testing sample, and  $F1, F2, F3, \dots, Fn$  are the features as defined in Section 2.3;  $n$  is the dimension of the feature vector;  $X_{ij}$  is the feature vector of  $i$ th training sample of PQ disturbance type  $S_j, S_j \in \{S1, S2, \dots, S9\}$ ;  $\delta$  is a smoothing parameter.

Summation layer makes summation of the results output from pattern layer, and the calculated conditional probability of  $X$  belonging to PQ disturbance type  $S_j$  is given below:

$$C_j(X) = \frac{1}{N_j} \sum_{i=1}^{N_j} H_{ij}(X) \quad (19)$$

where  $N_j$  is the number of training samples belonging to PQ disturbance type  $S_j$ .

In the competitive layer,  $X$  is assigned to the PQ disturbance type with maximum conditional probability. It has:

$$type(X) = \underset{j}{\operatorname{argmax}} \{C_j(X)\} \quad (20)$$

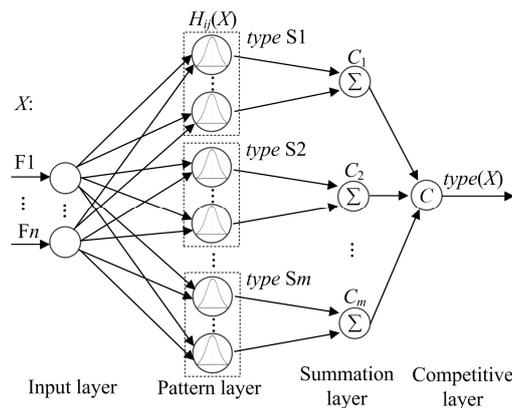


Figure 4. The schematic diagram of PNN.

### 3. Power Quality Disturbance Classification Based on Generalized S-transform with Feature Oriented Width Factor and Probabilistic Neural Network

The goal of this study is to realize an effective and efficient classification with high accuracy for eight types of disturbance signals plus the normal signal given in Section 2.2 using an approach based on the S-transform with a favorable width factor and PNN. Contents presented below include: selection of features used as inputs of PNN—four of ten are selected after examination; favorable value determination of the width factor for each of the four selected features by investigating the impacts of width factor on the PNN performance, and implementation of PNN for disturbance classification.

#### 3.1. Method Proposed—Considering the Width Factor to Be Feature Oriented

In conventional PQ disturbance classification based on S-transform, the width factor  $\lambda$  is treated as a constant. In other words, all features interested are extracted from the S-transform matrix obtained with a width factor of the same value. Our studies, presented here, reveal that considering the width factor  $\lambda$  to be feature oriented renders a more satisfactory result. Implementation data and results presented

below include: how does the STA-matrix of S-transform vary with  $\lambda$ , how does the effect of the feature separation behavior change with  $\lambda$ , and how does the classification accuracy of PNN vary with  $\lambda$ .

### 3.1.1. Effect of Width Factor on S-Transform-Amplitude-Matrix

Figure 5a–h shows the 3D plots of the STA-matrix of eight disturbance signals, listed in Table 1. For each disturbance signal, there are three 3D plots, which are graphical presentation of the STA-matrix obtained with three different values of the width factor  $\lambda$ , respectively. The STA-matrix is obtained per Equations (6) and (7); three values of  $\lambda$  are 0.1, 1.0, and 3.0. The values of  $\lambda$ , 0.1, 1.0, and 3.0 are selected per what was presented in Section 2.3.1: One less than 1.0, 0.1, which is used to examine its impacts on the low frequency domain resolution; one greater than 1.0, 3.0, which is used to examine its impacts on the high frequency domain resolution, and the value of 1.0, which is used as the base for comparison purpose.

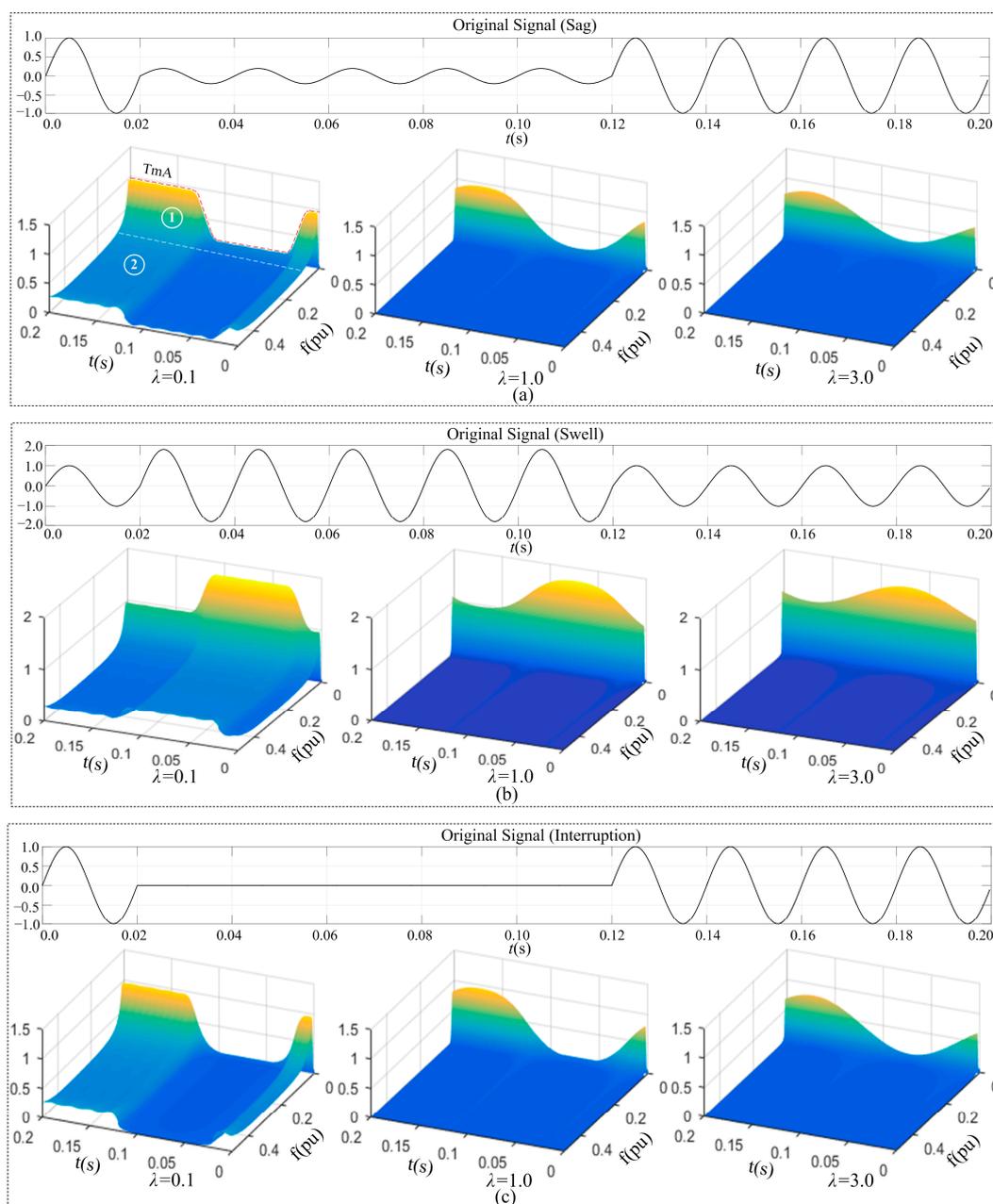


Figure 5. Cont.

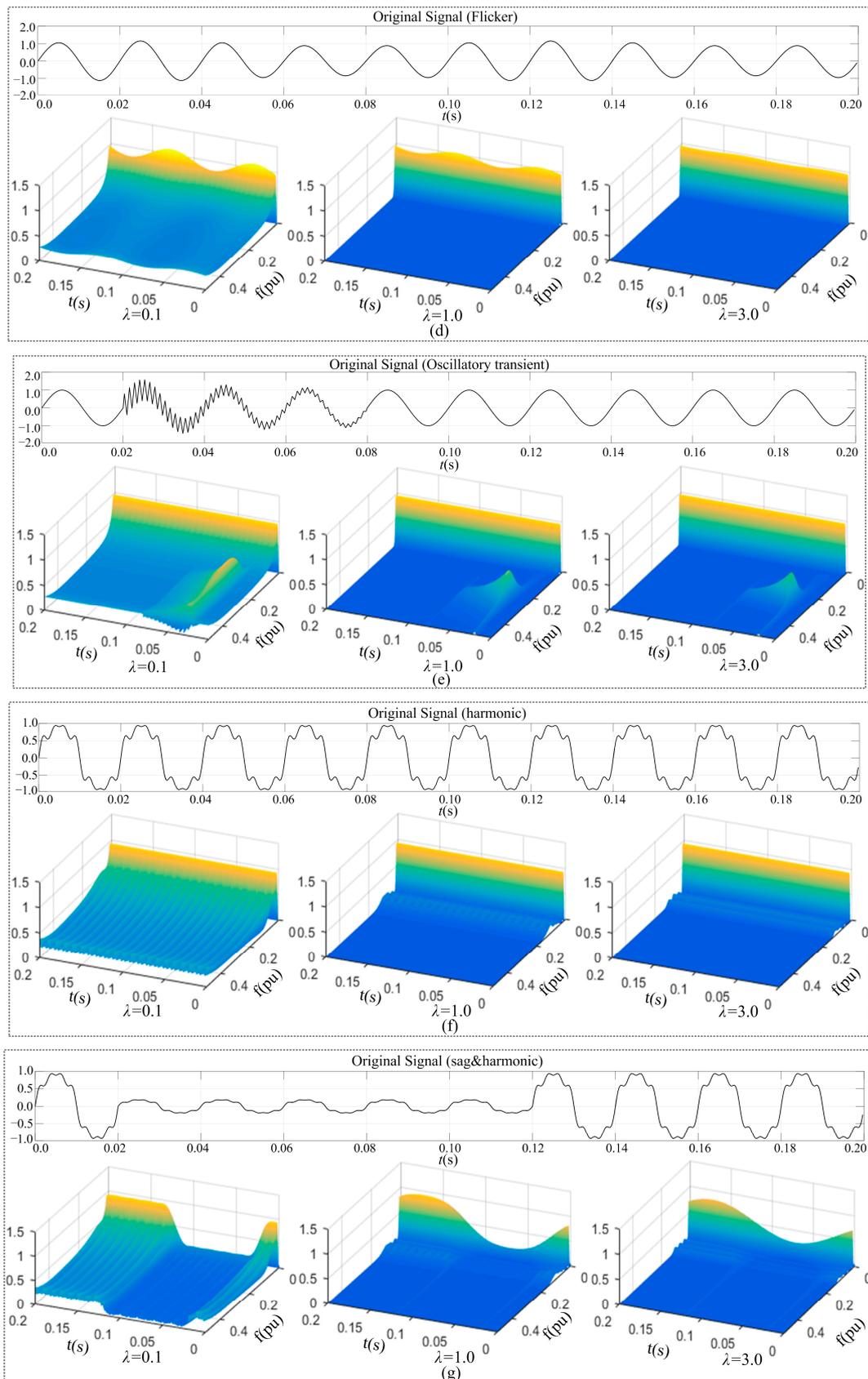
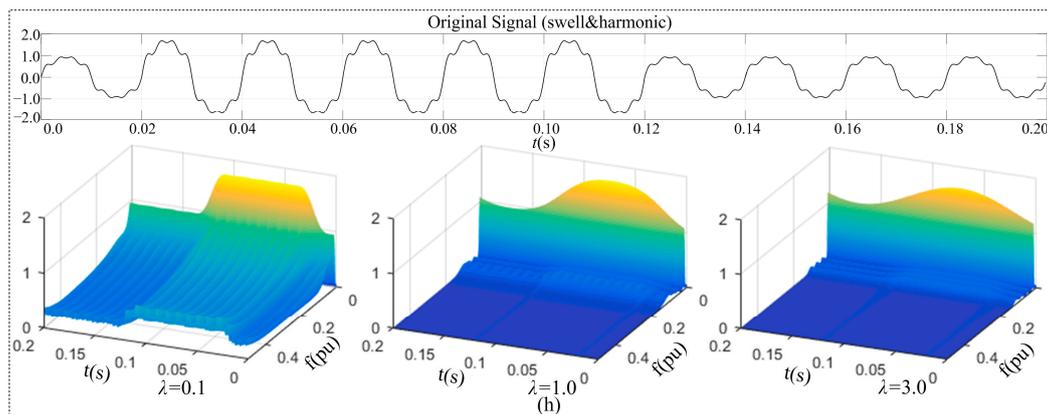


Figure 5. Cont.



**Figure 5.** Disturbance signal and 3-D time-frequency magnitude spectrum. (a) sag; (b) swell; (c) interruption; (d) flicker; (e) oscillatory transient; (f) harmonic; (g) sag & harmonic; and (h) swell & harmonic.

3D plots in Figure 5 illustrate the behavior of the time-frequency characteristic of STA-matrix of each specific disturbance signal under the condition that different values of  $\lambda$  are used. This information helps to understand the advantages of introducing the width factor  $\lambda$  and the impacts of different  $\lambda$  values on distinguishing eight disturbance signals, S1 to S8.

Disturbance signals S1–S8 can be divided into three groups in terms of their characteristics: group 1 which has S1–S4, group 2 which has S5 and S6, and group 3 which has S7 and S8. The characteristic of group 1 is that the disturbances occur to the amplitude of the signals at the line frequency for a short period of time. Plots in Figure 5a, as an example, show the S-transform results of sag, and it can be seen that:

- (1) When the width factor takes a value less than 1.0, the contour ( $TmA$  curve) represent the actual behavior of sag at the line frequency (shown by ①). This indicates that higher time resolution of the STA matrix is achieved. In other words, a less than 1.0 value of  $\lambda$  results in a satisfactory presentation of the behavior of sag at the line frequency; however, in the high-frequency domain (shown by ②), a less than 1.0 value of  $\lambda$  decreases the frequency resolution of the STA matrix, as one can find that obvious fluctuation, which is not the characteristic of the sag, appears in the high-frequency domain. That is, a less than 1.0 value of  $\lambda$  results in a wrong presentation of the behavior of sag in the high-frequency domain.
- (2) When the width factor takes a value greater than 1.0, one can see that the high-frequency domain is almost flat without any high-frequency component, which is consistent with the actual behavior of sag in the high-frequency domain. This indicates that higher frequency resolution of the STA matrix is achieved. That is, a greater than 1.0 value of  $\lambda$  results a better presentation of the behavior of sag in the high-frequency domain. However, a greater than 1.0 value of  $\lambda$  decreases the time resolution of the STA matrix at the line frequency. It can be seen that the change of the  $TmA$  curve around line frequency becomes smoother, which may lead to a wrong identification of sag due to the inaccurate presentation of its amplitude variation with time. That is, a greater than 1.0 value of  $\lambda$  results an unsatisfactory presentation of the behavior of sag at line frequency. Similar conclusion can be obtained if examining plots in Figure 5b–d.

The characteristic of group 2 is that the disturbance appears to be the occurrence of harmonic components for a certain period of time or the harmonic components existing over the entire time range. The common feature of these two disturbance signal types is that the amplitude of the line frequency component of them stays unchanged. The difference between these two signals and the others are in the high-frequency domain. Plots in Figure 5e–f show the S-transform results of oscillatory transient and harmonic, and it can be seen that: (1) when the width factor takes a value less than 1.0,

the frequency resolution of the STA matrix gets significantly reduced; High-frequency components which apparently do not belong to the original signals appear in the high-frequency domain; (2) these high frequency-components gradually fade away as  $\lambda$  increases. In the case when  $\lambda$  takes a value greater than 1.0, the high-frequency characteristics of the oscillatory transient and harmonic are well represented in the high-frequency domain.

Based on what are presented above, one can draw a conclusion that: A less than 1.0 value of  $\lambda$  results a better presentation of the behavior of the original signal at the line frequency, but distort the presentation of the behavior of the original signal in the high-frequency domain; A greater than 1.0 value of  $\lambda$  results a distorted presentation of the behavior of the original signal at the line frequency, but improves the presentation of the behavior of the original signal in the high-frequency domain. This indicates that combination of features is needed for the classification of different signals. As the signals of sag, swell, interruption, and flicker can be distinguished from others by the characteristic at the line frequency, a less than 1.0 value of  $\lambda$  is needed in the setting of the corresponding features; As the signals of oscillatory transient and harmonic can be distinguished from others by the characteristic in the high-frequency domain, a greater than 1.0 value of  $\lambda$  is needed in the setting of the corresponding features.

Characteristic of group 3 combines the characteristics of groups 1 and 2. Special attention needs to be paid to disturbance types S7 and S8, which are the combination of sag & harmonics and swell & harmonics, respectively. Investigation given above tells neither a value less than 1.0 nor a value greater than 1.0 ensures a successful separation of S7 or S8 from S1–S6, so in this condition, it needs combination of features to distinguish these disturbances from others.

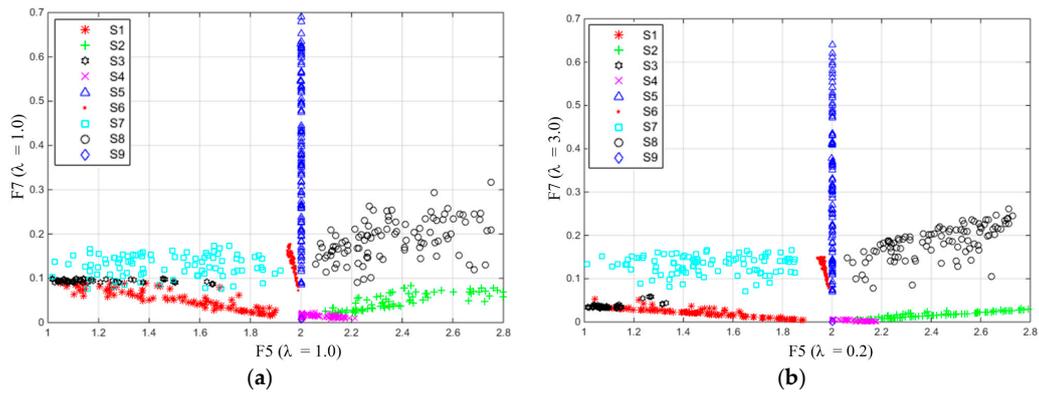
In summary, to realize classification of the nine signal types defined in Table 1, a combination of features is needed; to make the classification more effective and efficient, making the width factor feature oriented may be considered.

### 3.1.2. Effect of Width Factor on Feature Distribution Behavior

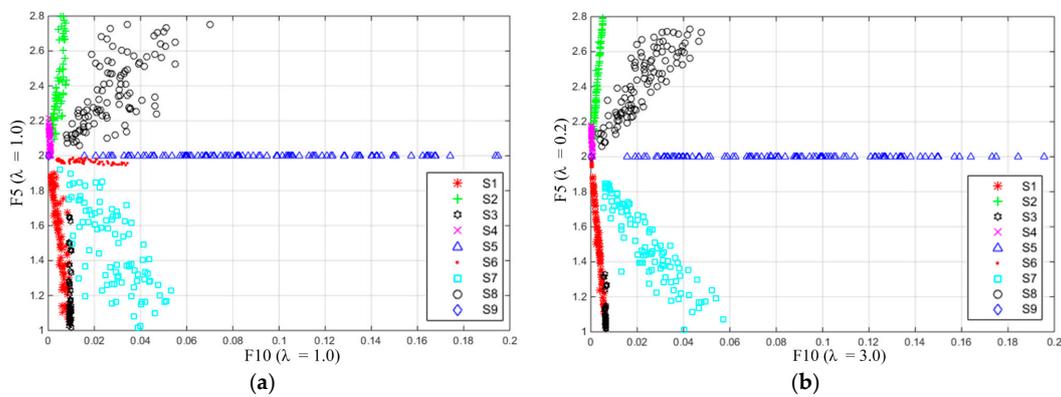
Different features represent different characteristics of signals. Among the 10 features defined in Section 2.3.2, features F1–F5, give an intuitive description of disturbance signals S1–S9 at the line frequency. Differently, features F6–F10, give an intuitive description of signals in the high-frequency domain. Specifically, F6–F9 expresses the frequency characteristic of signals in the high-frequency domain and F10 expresses the time characteristic of signals in the high-frequency domain. For a successful distinction of PQ disturbances S1–S9, commonly experienced in power system, a combination of features is needed per what are presented above. The goal is to use the least number of features to realize the classification with high accuracy. Various combination possibilities were examined. Combination of F2, F5, F7, and F10 was found to be the most applicable to our purpose:

- (1) F2 and F5 embody the line frequency characteristics of disturbance signals. F2 contributes to the separation of normal signal, sag, swell, and interruption by assigning a value less than 1.0 to width factor  $\lambda$ . Additionally, F5 is used to distinguish flicker from others by assigning a value less than 1.0 to width factor  $\lambda$ .
- (2) F7 embodies the frequency characteristic characteristics of signals in the high-frequency domain, and it contributes to distinguishing S1–S4 (sag, swell, flicker and interruption) from S5–S6 (harmonic, transient oscillation) by assigning a greater than 1.0 value to width factor  $\lambda$ .
- (3) F10 embodies the time characteristic of signals in the high-frequency domain, and it contributes to the separation of transient oscillation and harmonic by assigning a greater than 1.0 value to width factor  $\lambda$ .

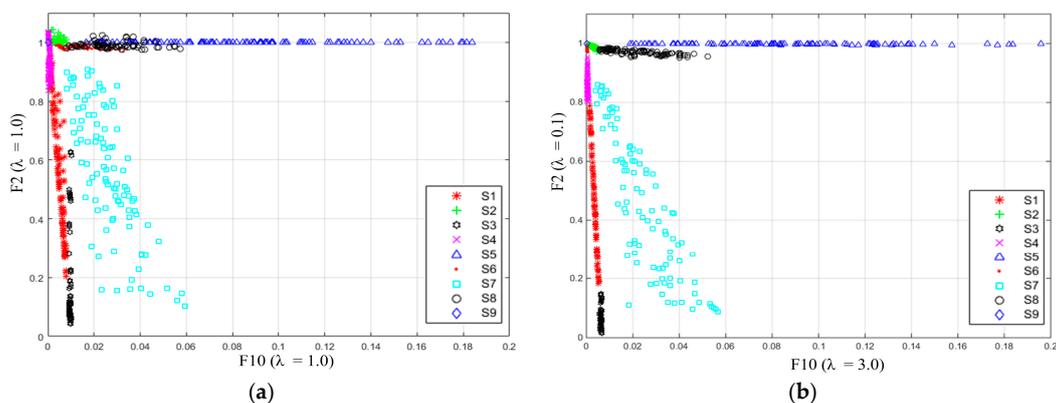
Based on the features F2, F5, F7, F10 above, firstly, to get intuitionistic vision, the distributions of PQ disturbances are shown in plots with the features combination F5 & F7, F5 & F10, F2 & F10. By analyzing the separation degree of eight types of PQ disturbances and the normal signal with different setting of  $\lambda$  ( $\lambda_{F2}$ ,  $\lambda_{F5}$ ,  $\lambda_{F7}$ ,  $\lambda_{F10}$ ), the distributions of PQ disturbances are shown in Figures 6–8.



**Figure 6.** Comparison of F5 and F7 distribution behavior: (a)  $\lambda_{F5} = 1.0$ ,  $\lambda_{F7} = 1.0$ ; and (b)  $\lambda_{F5} = 0.2$ ,  $\lambda_{F7} = 3.0$ .



**Figure 7.** Comparison of F5 and F10 distribution behavior: (a)  $\lambda_{F10} = 1.0$ ,  $\lambda_{F5} = 1.0$ ; and (b)  $\lambda_{F10} = 3.0$ ,  $\lambda_{F5} = 0.2$ .



**Figure 8.** Comparison of F10 and F2 distribution behavior: (a)  $\lambda_{F10} = 1.0$ ,  $\lambda_{F2} = 1.0$ ; and (b)  $\lambda_{F10} = 3.0$ ,  $\lambda_{F2} = 0.1$ .

Figure 6 shows distributions of F5 & F7 of signals S1–S9. Figure 6a is the distribution obtained with both  $\lambda_{F5}$  and  $\lambda_{F7}$  taking a value of 1.0, which is the case of the features extracted from traditional S-transform. It shows that there is obvious overlapping between areas taken by S1, S3, and S7, respectively. Also, the areas taken by S2 and S8 respectively are too close to be easily separated. This will reduce the accuracy of classification algorithm. For comparison, Figure 6b shows the results obtained with  $\lambda_{F5} = 0.2$  and  $\lambda_{F7} = 3.0$ , determined per discussion above. It can be seen that areas taken by S1, S3, and S7 are nicely separated, and there appears a clear space between areas taken by S2 and

S8. Feature F7 is used to distinguish S1–S4 from S5–S6. Comparison between results obtained with  $\lambda_{F7} = 1.0$  and  $\lambda_{F7} = 3.0$ , it can be seen the separation between S1–S4 and S5–S6 along the longitudinal axis (the direction of F7 axis) are much better in the case which  $\lambda_{F7}$  is equal to 3.0.

Figure 7 shows distributions of F10 & F5 of signals S1–S9. Figure 7a is the distribution obtained with both  $\lambda_{F10}$  and  $\lambda_{F5}$  taking a value of 1.0. It shows that there are obvious overlapping between areas taken by S1, S3 and S7; Figure 7b shows the results obtained with  $\lambda_{F10} = 3.0$  and  $\lambda_{F5} = 0.2$ . It shows that areas taken by S1, S2, S3, S7, S8 are nicely separated. F10 is used to separate transient oscillation S5 and harmonic S6, F2 and F10 distribution map given in Figure 7a shows that there is obvious overlapping between S5 and S6 along the horizontal axis (the direction of F10 axis) with  $\lambda_{F10} = 1.0$  while Figure 7b, there appears a clear space between areas taken by S5 and S6 with  $\lambda_{F10} = 3.0$ .

Figure 8 shows distributions of F10 & F2 of signals S1–S9. Feature F2 contributes to the separation of normal signal, sag, swell and interruption. The result in Figure 8a shows that, along the longitudinal axis (the direction of F2 axis), there is obvious overlapping between areas taken by S1 and S3, and also some overlapping between areas taken by S2 and S4 with  $\lambda_{F10} = 1.0$ ; the result in Figure 8b shows that areas taken by S1 and S3, S2 and S4 are nicely separated with  $\lambda_{F10} = 3.0$ . Moreover, comparison between Figure 8a,b says the S7 can be nicely separated from other signals with  $\lambda_{F10} = 3.0$ ,  $\lambda_{F2} = 0.1$ .

In summary, results presented in Figures 6–8 confirm that a combination of features is needed to separate S1 to S9. The results also verify that making  $\lambda$  feature oriented and using a favorable value instead of 1.0 for  $\lambda$  helps to achieve a satisfactory separation of S1–S9. In addition, for features which work better with  $\lambda$  greater than 1.0,  $\lambda$  values greater than 3.0 were tested and the result tells the improvement of the separation of signals S1 to S9 is insignificant. Similar results were obtained if values less than 0.1 are applied to  $\lambda$  for those features which work better with  $\lambda$  less than 1.0. Therefore, the variation range of  $\lambda$  used in the PNN section below is set to be from 0.1 to 3.0.

#### 4. Determination of the Favorable Value of Feature Oriented Width Factor with the Use of Probabilistic Neural Network

This section presents the determination of favorable width factor set ( $\lambda_{F2}$ ,  $\lambda_{F5}$ ,  $\lambda_{F7}$ , and  $\lambda_{F10}$ ) with PNN. Inputs of PNN are features F2, F5, F7, and F10 of the signal being classified. F2, F5, F7, and F10 are calculated from the S-transform matrix generated with width factor ( $\lambda_{F2}$ ,  $\lambda_{F5}$ ,  $\lambda_{F7}$ , and  $\lambda_{F10}$ ); output of PNN is the classification result. The objective is to obtain the favorable value of ( $\lambda_{F2}$ ,  $\lambda_{F5}$ ,  $\lambda_{F7}$ , and  $\lambda_{F10}$ ), with which and a trained PNN the classification of disturbances which falls into S1–S9 can be achieved with high accuracy. Steps for finding the favorable width factor ( $\lambda_{F2}$ ,  $\lambda_{F5}$ ,  $\lambda_{F7}$ , and  $\lambda_{F10}$ ) are given in Figure 9.

$6^4 = 1296$  combinations of width factor  $\lambda_{F2}$   $\lambda_{F5}$   $\lambda_{F7}$   $\lambda_{F10}$ , see the external loop of the flowchart in Figure 9, are examined, which are generated by assigning each element of [ $\lambda_{F2}$   $\lambda_{F5}$   $\lambda_{F7}$   $\lambda_{F10}$ ] with 0.1, 0.3, 0.6 1.0, 2.0, and 3.0, respectively. For each width factor combination set [ $\lambda_{F2}$   $\lambda_{F5}$   $\lambda_{F7}$   $\lambda_{F10}$ ]: (1) in PNN training, 900 samples of signals are used, which are obtained by randomly generating 100 signals for each of nine signal types; in PNN testing, similarly, 900 randomly generated samples of signals are used; (2) the feature vectors [F2 F5 F7 F10] are extracted from the 1800 samples of signals by using the S-transform with the corresponding width factor combination [ $\lambda_{F2}$   $\lambda_{F5}$   $\lambda_{F7}$   $\lambda_{F10}$ ]; (3) the PNN are trained and tested with the feature vectors [F2 F5 F7 F10] (900 for training; 900 for testing), and the classification error for the corresponding width factor combination [ $\lambda_{F2}$   $\lambda_{F5}$   $\lambda_{F7}$   $\lambda_{F10}$ ] is evaluated.

In the process explained above, the 200 signals (100 for PNN training; 100 for PNN testing) of each disturbance type are generated by randomly selecting the value of parameters used by each disturbance signal type, seen the parameter column of Table 1. To mitigate the possible randomness-related classification error, the internal loop is repeated six times,  $q = 6$ , as shown in the flowchart in Figure 9. The classification errors of PNN are calculated six times, and the average of the obtained six classification errors is taken as testing error of the PNN. The obtained testing errors of PNN are presented in Figure 10 with spheres, the size of which is proportional to the value of the corresponding error.

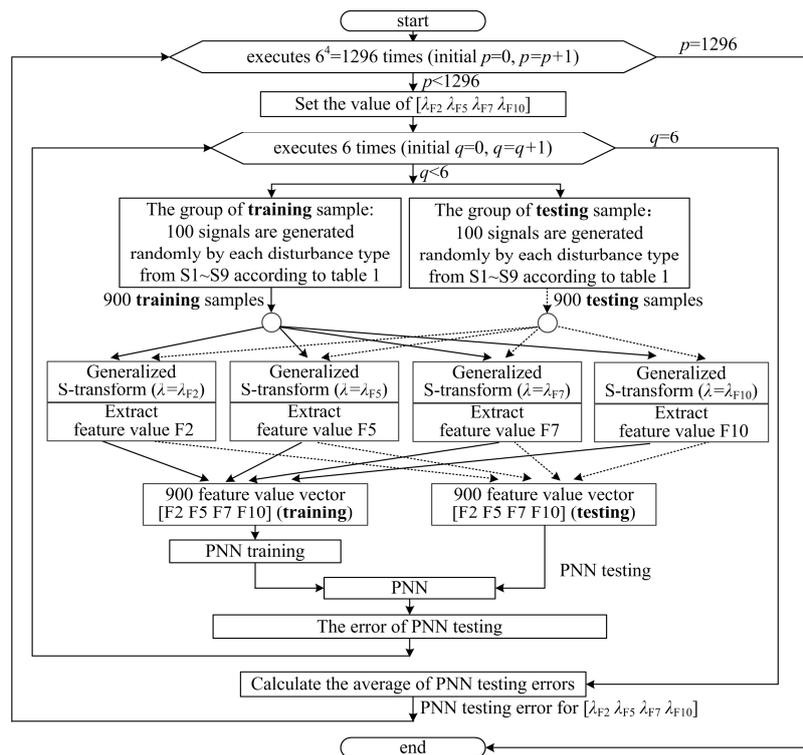


Figure 9. The flow chart for the PNN classification error.

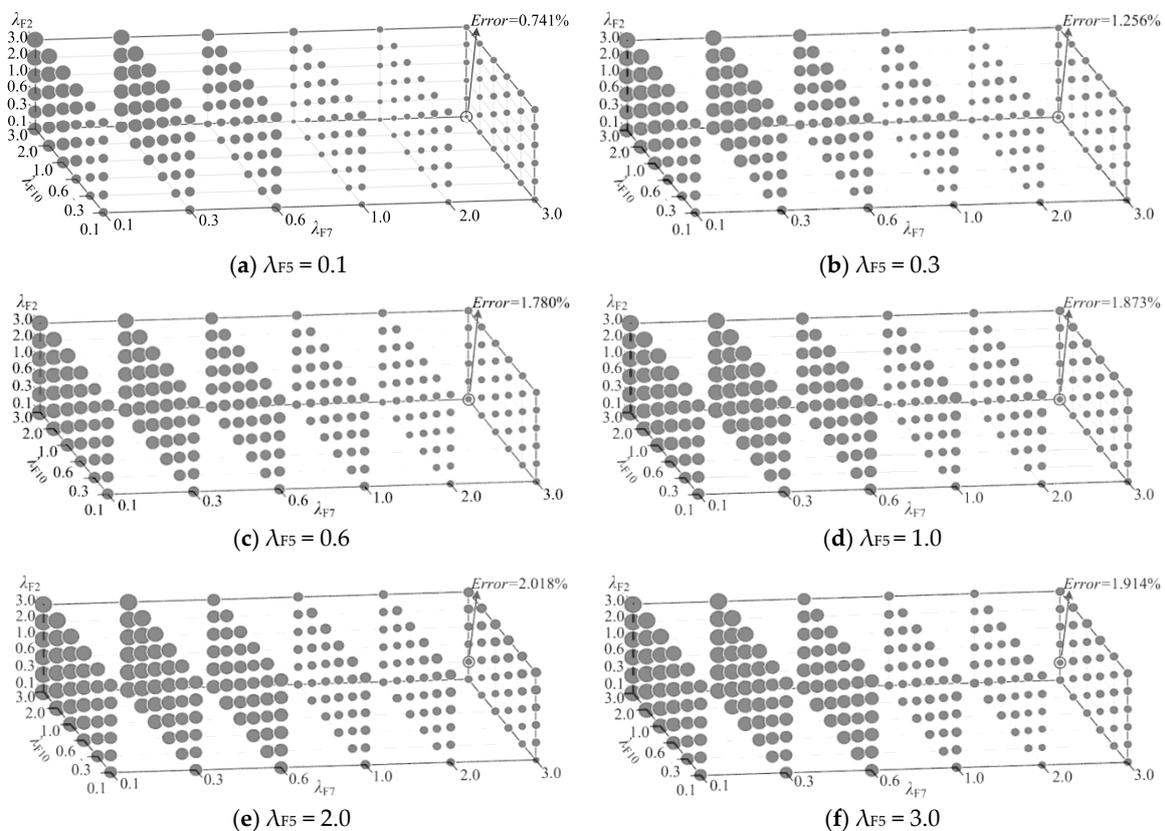


Figure 10. The error of PNN classification with the width factors combination  $[\lambda_{F2} \lambda_{F5} \lambda_{F7} \lambda_{F10}]$  varying in the range of 0.1–3.0 (the smaller the radius of sphere, the smaller its corresponding classification error).

Firstly, take a look at Figure 10a. The size of spheres becomes smallest when  $\lambda_{F2}$  takes the smallest value 0.1,  $\lambda_{F7}$  and  $\lambda_{F10}$  takes their greatest value 3.0. Similarity exists in Figure 10b–f. In other words, the smallest spheres on Figure 10a–f all locate at the lower right corner in the back. This tells us that favorable values for  $\lambda_{F2}$ ,  $\lambda_{F7}$  and  $\lambda_{F10}$  are 0.1, 3.0, and 3.0, respectively. Then, comparing the size of the smallest sphere of each individual subfigure, one can see that the sphere at the lower right corner in the back of Figure 10a has the smallest size. This tells us that the PNN classification error becomes the least if  $\lambda_{F5}$  takes a value of 0.1. Numbers in percentage format shown on the upper right corner in the back of each subfigure are the classification error of PNN corresponding to the sphere of the smallest size of each individual subfigure. It shows that the trained PNN will have the classification error be less than 1% (0.741% on Figure 10a) if  $\lambda_{F2}$ ,  $\lambda_{F5}$ ,  $\lambda_{F7}$ , and  $\lambda_{F10}$  take their values to be 0.1, 0.1, 3.0, and 3.0, respectively. The trained PNN, with features F2, F5, F7, F10 as inputs, may provide a satisfactory classification of PQ disturbance S1 to S8 and the normal sinusoidal signal if F2, F5, F7, F10 are extracted from S-transform matrix obtained with width factors  $\lambda_{F2}$ ,  $\lambda_{F5}$ ,  $\lambda_{F7}$ ,  $\lambda_{F10}$  equal to 0.1, 0.1, 3.0, 3.0, respectively.

### 5. Accuracy of the Proposed Power Quality Disturbance Classification Approach

Without noise, classification results of PNN corresponding to the favorable width factor combination are shown in Table 2, and the classification accuracy is 99.259%. To analyze the effect of noise on the classification errors, different levels of noise are added to the nine types of signals. The results are listed in Table 3. The level of noise is expressed by the signal-to-noise ratio (SNR), and  $SNR = 20\log_{10}(A_S/A_N)$ , where  $A_S$  and  $A_N$  are the maximum amplitude of the signal and noise, respectively.

**Table 2.** PNN classification results with the favorable width factor combination.

Category	S1	S2	S3	S4	S5	S6	S7	S8	S9	Accuracy
S1	592	0	8	0	0	0	0	0	0	98.67%
S2	0	583	0	17	0	0	0	0	0	97.17%
S3	0	0	600	0	0	0	0	0	0	100%
S4	0	9	0	591	0	0	0	0	0	98.50%
S5	0	0	0	0	600	0	0	0	0	100%
S6	0	0	0	0	0	600	0	0	0	100%
S7	6	0	0	0	0	0	594	0	0	99.00%
S8	0	0	0	0	0	0	0	600	0	100%
S9	0	0	0	0	0	0	0	0	600	100%

PNN classification accuracy = 99.259%

**Table 3.** PNN classification accuracy with noise conditions.

$\lambda_{F2}, \lambda_{F5}, \lambda_{F7}, \lambda_{F10}$	PNN Classification Accuracy (%)			
	Pure	40 dB	30 dB	20 dB
0.1, 0.1, 3.0, 3.0	99.26	99.13	98.63	98.38
1.0, 1.0, 1.0, 1.0	96.92	96.78	96.65	95.94
0.1, 2.0, 3.0, 0.6	97.42	96.80	96.62	96.17
2.0, 1.0, 0.1, 1.0	91.40	91.27	91.00	90.07
3.0, 1.0, 2.0, 1.0	97.04	95.98	95.87	95.68

As can be seen from Table 3, for the PQ disturbances without noise, the classification accuracy corresponding to the optimal width factor is 1.84%–7.86% higher than that of other width factor settings. The favorable width factor still maintains a good classification accuracy under noise (20 dB) and the classification accuracy is 2.21%–8.31% higher than the other width factor settings, including the traditional width factor settings ( $[\lambda_{F2} \lambda_{F5} \lambda_{F7} \lambda_{F10}] = [1.0, 1.0, 1.0, 1.0]$ ).

## 6. Performance Comparison

In order to evaluate the effectiveness and feasibility of the proposed algorithm, Table 4 shows the comparison between the obtained results in this paper and the reported results by other studies [4,8,16,17].

**Table 4.** Performance comparison in terms of percentage of correct classification results.

Category	Classification Accuracy (%)				
	Proposed Method	[4]	[8]	[16]	[17]
S1	98.67	88	98	95	100
S2	97.17	96.5	92	91	100
S3	100	85.55	100	99	100
S4	98.50	-	98	98	94
S5	100	-	100	100	98
S6	100	100	92	96	98
S7	99.00	100	93	98	98
S8	100	100	90	98	97
S9	100	100	100	100	100
Average classification accuracy	99.26	95.71	95.5	97.22	98.33

In [8,16], the classification accuracy of each PQ disturbance is lower than that of the proposed algorithm. For [4], the classification accuracy of each PQ disturbance is lower than that of the proposed algorithm except S7. For [17], the classification accuracy of each PQ disturbance is lower than that of the proposed algorithm, except for S1 and S2. The average classification accuracy shows the ratio of correctly classified PQ disturbances to the total number of PQ disturbances, and the proposed method gives the best classification results for this case.

The classification accuracy comparison between the proposed algorithm and other reported studies is shown in Table 5. For the PQ disturbances without noise, the classification accuracy corresponding to the proposed algorithm is 99.26% higher than that of other algorithms. For the PQ disturbances with low level noise condition (40 dB), the classification accuracy corresponding to the proposed algorithm is 99.13% slightly lower than that of the algorithm in [27], but higher than that of other algorithms. For the PQ disturbances with high level noise condition (30 dB), the classification accuracy corresponding to the proposed algorithm is 98.63% higher than that of other algorithms.

**Table 5.** Comparison between the proposed algorithm and other algorithms with noise conditions.

PQ Classification Studies	Classification Accuracy (%)		
	Pure	40 dB	30 dB
[4]	95.71	93.64	91.85
[27]	-	99.7	98.5
[28]	98.5	96.875	93.625
[29]	-	98.8	97.49
Proposed method	99.26	99.13	98.63

## 7. Conclusions

This paper proposed a PQ disturbance classification approach based on S-transform with a feature-oriented width factor and PNN. By introducing a width factor into the conventional S-transform, the time resolution and the frequency resolution presented by the STA matrix of the signal being analyzed is made adjustable. In this way, the impact of the width factor on the 3D-STA time-frequency magnitude spectrum of eight disturbance signals are studied and the overall picture of how the regulator factor affects the description accuracy of the signal in the low frequency domain and

the high frequency domain is obtained. On the basis of this and according to the joint consideration of the characteristics of eight disturbance types in frequency domain and the definition of 10 features, four out of 10 features are selected to be used for the disturbance classification. Three combinations of four selected features are investigated in terms of the 2D distribution behavior of their values for the eight disturbance signal types; the influence of the width factor on the separation of data points denoting the values of each feature combination is presented. From there, association between each feature and the width factor value favorable for the corresponding feature is established. Further, it is verified with PNN by examining the classification accuracy with a wide variation range of each width factor, from 0.1 to 3.0. Furthermore, a PNN satisfactorily trained is obtained. Simulation tells it renders high classification accuracy (less than 1% error) for 8 type disturbance signals by using only four features as inputs, which are extracted from the S-transform amplitude matrix with corresponding favorable width factor. In addition, the obtained PNN shows satisfactory robustness under various noise conditions. Finally, the proposed algorithm shows better performance in comparison with those presented in other research studies.

**Acknowledgments:** The authors are grateful to the financial support from the Specialized Research Fund for National Key Research and Development Plan (No.2016YFB0900204), and doctoral Program of Higher Education of China (No. 20120032110070).

**Author Contributions:** The paper was collaborative effort among the authors. The authors contributed collectively to the theoretical analysis and manuscript preparations.

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

## References

1. Sainib, M.K.; Kapoor, R. Classification of quality disturbances event—A review. *Int. J. Electr. Power Energy Syst.* **2012**, *43*, 11–19. [[CrossRef](#)]
2. Gu, Y.H.; Bollen, M.H.J. Time-frequency and time scale domain analysis of voltage disturbance. *IEEE Trans. Power Deliv.* **2000**, *15*, 1279–1284. [[CrossRef](#)]
3. Biswal, B.; Biswal, M.; Hasanc, S. Nonstationary power signal time series data classification using LVQ classifier. *Appl. Soft Comput.* **2014**, *18*, 158–166. [[CrossRef](#)]
4. Uyara, M.; Yildirim, S.; Gencoglu, M.T. An effective wavelet-based feature extraction method for classification of power quality disturbance signals. *Electr. Power Syst. Res.* **2008**, *78*, 1747–1755. [[CrossRef](#)]
5. Uyara, M.; Yildirim, S.; Gencoglu, M.T. An expert system based on S-transform and neural network for automatic classification of power quality disturbances. *Expert Syst. Appl.* **2009**, *36*, 5962–5975. [[CrossRef](#)]
6. Ray, P.K.; Kishor, N.; Mohanty, S.R. Islanding and power quality disturbance detection in grid-connected hybrid power system using wavelet and S-transform. *IEEE Trans. Smart Grid* **2012**, *3*, 1082–1094. [[CrossRef](#)]
7. 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](#)]
8. Bhende, C.N.; Mishra, S.; Panigrahi, B.K. Detection and classification of power quality disturbances using S-transform and modular neural network. *Electr. Power Syst. Res.* **2008**, *78*, 122–128. [[CrossRef](#)]
9. Dash, P.K.; Panigrahi, B.K.; Panda, G. Power quality analysis using S-transform. *IEEE Trans. Power Deliv.* **2003**, *18*, 406–411. [[CrossRef](#)]
10. Pinnegar, C.R.; Mansinha, L. The S-transform with windows of arbitrary and varying shape. *Geophysics* **2003**, *68*, 381–385. [[CrossRef](#)]
11. Djurović, I.; Sejdić, E.; Jiang, J. Frequency-based window width optimization for S-transform. *Int. J. Electr. Commun.* **2008**, *62*, 245–250. [[CrossRef](#)]
12. Sejdić, E.; Djurović, I.; Jiang, J. A window width optimized S-transform. *EURASIP J. Adv. Signal Process.* **2007**, *2008*, 672941. [[CrossRef](#)]
13. Huang, N.; Zhang, S.X.; Cai, G.W.; Xu, D.G. Power quality disturbances recognition based on a multiresolution generalized S-transform and a PSO-improved decision tree. *Energies* **2015**, *8*, 549–572. [[CrossRef](#)]
14. McFadden, P.D.; Cook, J.G.; Forster, L.M. Decomposition of gear vibration signals by the generalized S-transform. *Mech. Syst. Signal Process.* **1999**, *13*, 691–707. [[CrossRef](#)]

15. Pinnegar, C.R.; Mansinha, L. Time-local Fourier analysis with a scalable, phase-modulated analyzing function: The S-transform with a complex window. *Signal Process.* **2004**, *84*, 1167–1176. [[CrossRef](#)]
16. Mishra, S.; Bhende, C.N.; Panigrahi, B.K. Detection and classification of power quality disturbances using S-transform and probabilistic neural network. *IEEE Trans. Power Deliv.* **2008**, *23*, 280–287. [[CrossRef](#)]
17. Valtierra-Rodriguez, M.; Romero-Troncoso, R.D.J.; Osornio-Rios, R.A.; Garcia-Perez, A. Detection and classification of single and combined power quality disturbances using neural networks. *IEEE Trans. Ind. Electr.* **2014**, *61*, 2473–2481. [[CrossRef](#)]
18. Biswal, B.; Dash, P.K.; Panigrahi, B.K. Power quality disturbance classification using fuzzy c-means algorithm and adaptive particle swarm optimization. *IEEE Trans. Ind. Electr.* **2009**, *56*, 212–220. [[CrossRef](#)]
19. Mohanty, S.R.; Ray, P.K.; Kishor, N.; Panigrahi, B.K. Classification of disturbances in hybrid DG system using modular PNN and SVM. *Electr. Power Energy Syst.* **2013**, *44*, 764–777. [[CrossRef](#)]
20. Kumar, R.; Singh, B.; Shahani, D.T.; Chandra, A.; Al-Haddad, K. Recognition of power-quality disturbances using S-transform -based ANN classifier and rule-based decision tree. *IEEE Trans. Ind. Appl.* **2015**, *51*, 1249–1258. [[CrossRef](#)]
21. Meng, K.; Dong, Z.Y.; Wang, D.H.; Wong, K.P. A self-adaptive RBF neural network classifier for transformer fault analysis. *IEEE Trans. Power Syst.* **2010**, *25*, 1350–1360. [[CrossRef](#)]
22. Specht, D.F. Probabilistic neural networks and the polynomial Adeline as complementary techniques for classification. *IEEE Trans. Neural Netw.* **1990**, *1*, 111–121. [[CrossRef](#)] [[PubMed](#)]
23. Lin, W.M.; Lin, C.H.; Sun, Z.C. Adaptive multiple fault detection and alarm processing for loop system with probabilistic network. *IEEE Trans. Power Deliv.* **2004**, *19*, 64–69. [[CrossRef](#)]
24. *IEEE Recommended Practices for Monitoring Electric Power Quality*; IEEE Std 1159-1995; IEEE-SASB Coordinating Committees: New York, NY, USA, 1995.
25. *EN 50160: Voltage Characteristics of Electricity Supplied by Public Distribution System*; European Standards: Pilsen, Czech Republic, 2002.
26. Bohari, M.N. Voltage Dip 21/3/2014 0650hrs. Available online: <http://powerquality.sg/wordpress/?p=194> (accessed on 21 March 2014).
27. Zhao, F.Z.; Yang, R.G. Power-quality disturbance recognition using S-transform. *IEEE Trans. Power Deliv.* **2007**, *22*, 944–950. [[CrossRef](#)]
28. Hooshmand, R.; Enshae, A. Detection and classification of single and combined power quality disturbances using fuzzy systems oriented by particle swarm optimization algorithm. *Electr. Power Syst. Res.* **2010**, *80*, 1552–1561. [[CrossRef](#)]
29. Biswal, M.; Dash, P.K. Detection and characterization of multiple power quality disturbances with a fast S-transform and decision tree based classifier. *Digit. Signal Process.* **2013**, *23*, 1071–1083. [[CrossRef](#)]



© 2017 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/>).