

Article

Combined Power Quality Disturbances Recognition Using Wavelet Packet Entropies and S-Transform

Zhigang Liu*, Yan Cui and Wenhui Li

School of Electrical Engineering, Southwest Jiaotong University, Chengdu, Sichuan 610031, China

* Author to whom correspondence should be addressed; E-Mail: liuzg@swjtu.cn;
Tel.: +86-28-8760-3229.

Academic Editor: Raúl Alcaraz Martínez

Received: 29 May 2015 / Accepted: 10 August 2015 / Published: 12 August 2015

Abstract: Aiming at the combined power quality +disturbance recognition, an automated recognition method based on wavelet packet entropy (WPE) and modified incomplete S-transform (MIST) is proposed in this paper. By combining wavelet packet Tsallis singular entropy, energy entropy and MIST, a 13-dimension vector of different power quality (PQ) disturbances including single disturbances and combined disturbances is extracted. Then, a ruled decision tree is designed to recognize the combined disturbances. The proposed method is tested and evaluated using a large number of simulated PQ disturbances and some real-life signals, which include voltage sag, swell, interruption, oscillation transient, impulsive transient, harmonics, voltage fluctuation and their combinations. In addition, the comparison of the proposed recognition approach with some existing techniques is made. The experimental results show that the proposed method can effectively recognize the single and combined PQ disturbances.

Keywords: wavelet packet; Shannon entropy; Tsallis entropy; modified incomplete S-transform; combined disturbances; recognition

1. Introduction

With the dramatic increase of non-linear and unbalanced loads being connected to the power grid, such as new energy outlets, smart grids and high-speed trains, the issue of power quality (PQ) becomes more and more serious. In order to ensure high quality power supply of the power grid, it is very

important to analyze and recognize these PQ disturbances. PQ disturbance recognitions can be divided into single and combined disturbance recognitions. The single disturbance refers to that only containing one disturbance in PQ signals. According to the IEEE power quality standard, the single disturbance includes voltage sag, voltage swell, voltage interruption, transient impulse, oscillation transients, harmonics and flicker. At present, researches on single disturbance recognition have formed some mature theoretical and experimental methods. Actually, PQ disturbances in power systems are often mixed disturbances, which are the combinations of several single disturbances, and the various components and types of these disturbances are more complex. In addition, the interaction between certain single disturbances may cause feature aliasing and even failure characteristics, and will result in wrong evaluation and low recognition accuracy. Therefore, considering that feature extraction is very difficult, more universal recognition methods need to be further analyzed.

Recently, some methods have been proposed, especially for the combined disturbance recognitions. In [1], a new dual neural-network-based methodology to detect and classify single and combined PQ disturbances is proposed. The adaptive linear network for harmonic and interharmonic estimation that allows computing the root-mean-square voltage and total harmonic distortion indices is adopted. In [2], an improved classification for PQ disturbances considering load changes and environmental factors is proposed. The hyperbolic S-transform is adopted, out of which the optimal features are selected using a genetic algorithm. These optimal features are used for combined various forms of PQ disturbances classification by employing support vector machine (SVM) and decision tree classifiers. In [3], a new method is presented for detection and classification of single and combined PQ disturbances using the sparse signal decomposition on the overcomplete hybrid dictionary matrix. In [4], an integrated approach using discrete wavelet transform and hyperbolic S transform is presented. The orthogonal forward selection by incorporating the Gram Schmidt procedure and forward selection are adopted for the selection of the best subset features, and the variable parameters of classifiers are optimized using particle swarm optimization (PSO). In [5], a method based on single channel independent component analysis for single and multiple power quality disturbance classification is proposed. The proposed method decouples the power system signal into its independent components, which are classified by specialized classifiers. The classifier outputs are combined by using a logic that gives the final classification. In [6], a method based on independent component analysis (ICA) is proposed to adaptively decompose signals containing multiple power quality disturbances. In [7], a multiresolution generalized S-transform (MGST) approach is presented to improve the ability of analyzing and monitoring the power quality in a microgrid. In [8], a combination method is proposed for the classification of combined power quality disturbances based on ensemble empirical mode decomposition (EEMD) and multilabel learning. In [9], a modified technique for the recognition of single stage and multiple power quality disturbances is introduced. An algorithm, based on Stockwell's transform, artificial neural network-based classifier and rule-based decision tree, is proposed. Some of the latest research on combined disturbance recognition can be found in [10–14].

S-transform is a kind of reversible local time-frequency domain transform developed from Fourier transform. It possesses no cross terms interference, frequency adaptive Gauss window and maximum time-frequency resolution [15]. More detailed time-frequency features of high-frequency components can be extracted. Modified incomplete S-transform (MIST) is a simplified and improved S-transform to reduce the computational complexity and efficiency of time frequency traversal [16]. Wavelet

transform is a multi-scale time-frequency transform with multi-resolution analysis, which is widely used to analyze PQ disturbance signals [17]. Compared with traditional wavelets, the wavelet packet has better properties. In addition, wavelet packet transform can yield more frequency sub-bands. However, due to the complexity of combined PQ disturbance signals, the actual research shows that there are problems with feature extraction, such as energy leakage and aliasing between adjacent scales with the application of various methods based on wavelet transform. In [18,19], combining wavelet decomposition with Shannon entropy, the concept of wavelet energy entropy (WEE), wavelet time entropy (WTE), wavelet singular entropy (WSE), multi-wavelet packet coefficient entropy (MWPCE) were presented and discussed. In [20,21], because of the advantages of Tsallis entropy, as opposed to Shannon entropy, other wavelet entropies were defined, namely Tsallis wavelet energy entropy (TWEE), Tsallis wavelet time entropy (TWTE), Tsallis wavelet singular entropy (TWSE), and wavelet packet Tsallis singular entropy (WPTSE). The application of these entropies opened up a new way of thinking to apply wavelet transform in PQ analysis. The application results show that wavelet entropies can more accurately characterize the features of complex signals, and their ability for feature extraction is better. Compared with other classifiers, the decision tree is a direct-vision method. It takes a shorter amount of time than most of the other classifiers. If the suitable extracted eigenvectors and reasonable decision rules are chosen, the decision tree can obtain ideal recognition accuracy [9,22,23].

In this paper, based on wavelet packet Tsallis singular entropy, energy entropy, modified incomplete S-transform and ruled decision tree, a new recognition method of combined PQ disturbances is proposed. This paper is structured as follows. Section 2 provides the introduction of the wavelet packet entropies and modified incomplete S-transform. Section 3 proposes the recognition plan including the extracting algorithm and ruled decision tree. Section 4 establishes simulation models and analyzes corresponding experimental results. Section 5 draws some conclusions.

2. Wavelet Packet Entropies and Modified Incomplete S-Transform

2.1. Wavelet Packet Decomposition

When wavelet transform is used to decompose the signals, the high-frequency resolution of signals is very low. Wavelet packet is a generalization of wavelet bases by taking linear combinations of the traditional wavelet functions to form a function cluster. Wavelet packet transform can provide more detailed decomposition of high frequency components. These decompositions are redundant and have no omissions. The results of wavelet packet transform can represent the complexity of signals more accurately and can obtain better time-frequency analysis. The discrete wavelet packet recursive decomposition is listed as follows.

$$\begin{cases} d_{i,2j}(t) = \sqrt{2} \sum_k g(k) d_{i-1,j}(2t-k) \\ d_{i,2j-1}(t) = \sqrt{2} \sum_k h(k) d_{i-1,j}(2t-k) \\ d_{0,0}(t) = S_0(t) \end{cases} \quad (1)$$

where $S_0(t)$ is the original signal, $h(k)$ is the high-pass filter, $g(k)$ is the low-pass filter, and $d_{i,j}(k)$ is the reconstruction signal of wavelet packet decomposition at the i^{th} level for the j^{th} node.

For the wavelet packet decomposition, at level 1, let L1 be the low frequency part and H1 be the high frequency part of wavelet transformation. At level 2, let LL2 and LH2 respectively be low and high frequency parts of L1, and HL2 and HH2 respectively be low and high frequency parts of H1. At level 3, let LLL3 and LLH3 respectively be low and high frequency parts of LL2, let LHL3 and LHH3 respectively be low and high frequency parts of LH2, let HLL3 and HLH3 respectively be low and high frequency parts of HL2, and let HHL3 and HHH3 respectively be low and high frequency parts of HH2. L1 and H1 can be divided into four bands at level 2. Similarly, LL2, LH2, HL2 and HH2 can be divided into eight bands at level 3. The framework of the 3-level wavelet packet decomposition is shown in Figure 1.

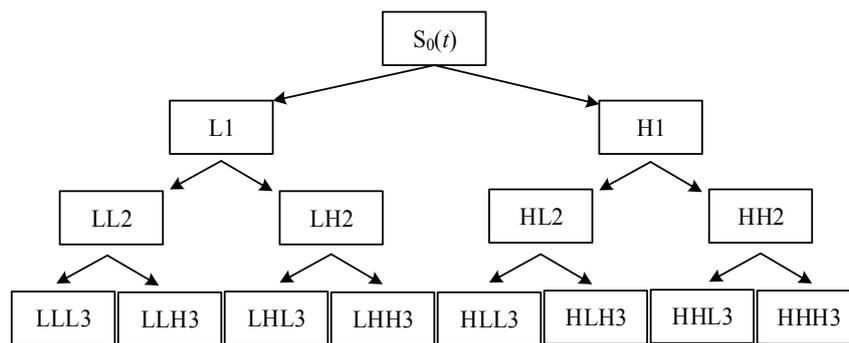


Figure 1. Framework of wavelet packet decomposition.

Based on the characteristics of the wavelet packet decomposition, the choice of decomposition level and wavelet function must be considered. In addition, in order to avoid the redundancy representation for the signal with the wavelet packet decomposition, the best wavelet packet basis must be considered. For the wavelet packet decomposition, there are many methods for choosing the best basis, such as the number of coefficients exceeding some threshold, $l_p(p \leq 2)$ norm, logarithm entropy and energy entropy. Considering the existing research results, Daubechies series wavelets are more sensitive to irregular signal, thus DB4 wavelet function based three level wavelet packet decomposition is chosen. Shannon entropy is adopted as the cost function to find the best wavelet packet base.

For the combined PQ disturbance signals, the high-frequency features are very important to characterize the transient components of disturbance signals. Wavelet packet transform is adopted as one of the time-frequency analysis tools. Compared with the traditional wavelet decomposition, wavelet packet decomposition can extract richer and more sophisticated features of combined PQ disturbances. Moreover, Shannon entropy and Tsallis entropy can be considered as the system's information measurement, and can be adopted to estimate the complexity of random signals, since the combined PQ disturbances contain a great deal of transient, random and uncertain information.

2.2. Shannon Entropy and Wavelet Packet Energy Entropy

Entropy is the measurement of disorder such as the imbalance, uncertainty and randomness. The uncertainty of any event is associated with its states and probabilities. For some uncertain system, if

the random variable $X \{x_1, x_2, \dots, x_L\}$ represents its state characteristics, where the probability of x_j is $p_j = P\{X = x_j\}, j = 1, \dots, L \in N, \sum_{j=1}^L p_j = 1$.

The information for some result of X can be represented with $I_j = \log(1/p_j)$. The information entropy of X can be defined by Shannon.

$$H(X) = -\sum_{j=1}^L p_j \log(p_j) \tag{2}$$

If $p_j = 0$, then $p_j \log(p_j) = 0$. Through a i -level wavelet packet decomposition, set $E_{(i,k)} = |D_j(k)|^2$ as wavelet packet energy at the j^{th} scale, where $k=1,2,\dots,N$ is the sampling point of original signal. Then, $E_j = \sum_{k=1}^N E_{(i,k)}$ represent the sum energy at the j^{th} scale. Set relative wavelet packet energy $p_{(j,k)} = E_{(i,k)} / E_j$, according to energy conservation principle, $\sum_k p_{(j,k)} = 1$. Based on fundamental principle of information entropy, wavelet packet energy entropy (WPEE) distributing along the scale is defined as follows.

$$W_{EE} = -\sum_k P_{(j,k)} \log p_{(j,k)} \tag{3}$$

2.3. Tsallis Entropy and Wavelet Packet Tsallis Entropy

Tsallis entropy is one of the nonextensive entropies, which can provide the correct physical expression for the non-additive system with a mixed or irregular fragment [24]. It can characterize the original signal more accurately. Its continuous expression is defined as follows.

$$S_q = \frac{c}{q-1} \left(1 - \int f(x)^q dx \right), \quad q \in R \tag{4}$$

where $f(x)$ is the probability density distribution function, $\int f(x)dx = 1$, c is the conventional positive constant, and q is the nonextensive parameter. The discrete expression is shown below.

$$S_q = \frac{c}{q-1} \left(1 - \sum_{i=1}^n p(i)^q \right), \quad q \in R, \quad n \in N \tag{5}$$

where $p(i) (\sum_{i=1}^n p(i) = 1)$ is the probability density distribution function of random variable i .

For the calculation of Tsallis entropy, the nonextensive parameter q needs to be considered. When $q \rightarrow 1$, Tsallis entropy is equal to Shannon entropy, which can describe the system with an extensivity property. Tsallis entropy is the expansion of Shannon entropy. After the wavelet packet decomposition for combined PQ disturbance signals, the decomposition coefficients at each scale generally show nonextensivity. For the wavelet packet transform, the frequency aliasing and energy leakage may be produced, which will result in the loss of the extensivity property for Shannon entropy. If the proper nonextensivity parameter q is chosen and Tsallis entropies are calculated with wavelet coefficients over time, the time-frequency characteristics of combined PQ disturbances based on the variety of entropy values can be obtained. In this paper, WPTSE is adopted as one of the feature extraction methods. Based on the existing research results in [20], when the nonextensive parameter q is chosen

in the ranges of (0.2, 1) and (1, 3), the influence of q values is very small. In this paper, set $q = 0.8$ considering the nonextensive degree of disturbance signals.

The reconstruction signals of each node for wavelet packet decomposition can form a matrix $D = \{d_{i,j}(k), k = 1, \dots, L, 1 \leq i \leq N, j = 1, \dots, 2^i\}$. If a moving data window is defined on the reconstruction signals of wavelet packet node, suppose the window width is $w \in N$ and moving factor be $\delta \in N$, then the moving window can be represented as follows.

$$W(m, w, \delta) = \begin{Bmatrix} d_{i,1}(1+m\delta) & d_{i,1}(2+m\delta) & \dots & d_{i,1}(w+m\delta) \\ d_{i,2}(1+m\delta) & d_{i,2}(2+m\delta) & \dots & d_{i,2}(w+m\delta) \\ \vdots & \vdots & \ddots & \vdots \\ d_{i,2^i}(1+m\delta) & d_{i,2^i}(2+m\delta) & \dots & d_{i,2^i}(w+m\delta) \end{Bmatrix} \quad (6)$$

where $m = 1, 2, \dots, M$, $M = (L - w) / \delta$, m is the moving number of data window, M is the length of wavelet entropy, and L is the signal length.

The reconstructed signals of 2^i wavelet packet nodes in moving data window $W(m, w, \delta)$ form a matrix $D_{2^i \times w}$. Based on the singular value decomposition theory [18], matrix $D_{2^i \times w}$ can be decomposed as follows.

$$D_{2^i \times w} = U_{2^i \times g} A_{g \times g} V_{g \times w} \quad (7)$$

The diagonal elements $\lambda_s (s = 1, \dots, g, g \leq \min(2^i, w))$ of diagonal matrix A are the singular values of $D_{2^i \times w}$, and $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_g \geq 0$. According to the reconstructed signal characteristics of wavelet packet node, if the correlation is stronger, the frequency components are more similar. When the reconstructed signals of neighbor nodes are approximately consistent, the corresponding singular value approaches zero. Otherwise, if the frequency components are quite different, there are a few nonzero singular values on the main diagonal. The number of nonzero singular values can reflect the complexity of frequency components in tested signals. In order to represent the rule, at the moment $(m\delta + w/2)$, the WPTSE at i th scale is formulated as follows.

$$W_{TsallisPSE}(m) = \frac{1}{q-1} \left(1 - \sum_{j=1}^g \Delta P_m(i, j)^q \right), \quad q \in R \quad (8)$$

where $\Delta P_m = \lambda_j / \sum_{s=1}^g \lambda_s$, and q is the nonextensive parameter.

Based on the Equation (8), if there are more nonzero singular values, the WPTSE value will be large. With the combination of redundant information, the singular entropy in wavelet packet spaces can directly reflect the uncertainty of characteristic energy distributions in the time-frequency domain of signals.

2.4. Modified Incomplete S-Transform

S-transform is a kind of reversible local time-frequency domain transform developed from Fourier transform. It possesses interference with no cross terms, frequency adaptive Gauss window and maximum time-frequency resolution. However, the frequency domain ergodicity of S-transform leads

to low computation speed and efficiency. MIST is a simplified and improved S-transform. The discrete form expression is below.

$$S_{MI}[j, n_d] = \sum_{m=0}^{N-1} H[m + n_d] e^{-\frac{2\pi^2 m^2 \lambda_{n_d}}{n_d}} e^{i \frac{2\pi m j}{N}} \quad n_d = l_1 \sim l_L \quad (9)$$

where λ_{n_d} is the window width coefficient chosen for different frequency points, and $l_1 \sim l_L$ are main frequency points detected by the dynamic measurement using power spectrum envelope. L is generally 1~4. The conversion relationship of actual frequency is $f = n_d/NT$, and T is the sampling period. The rapid calculation procedure of MIST is shown in Figure 2.

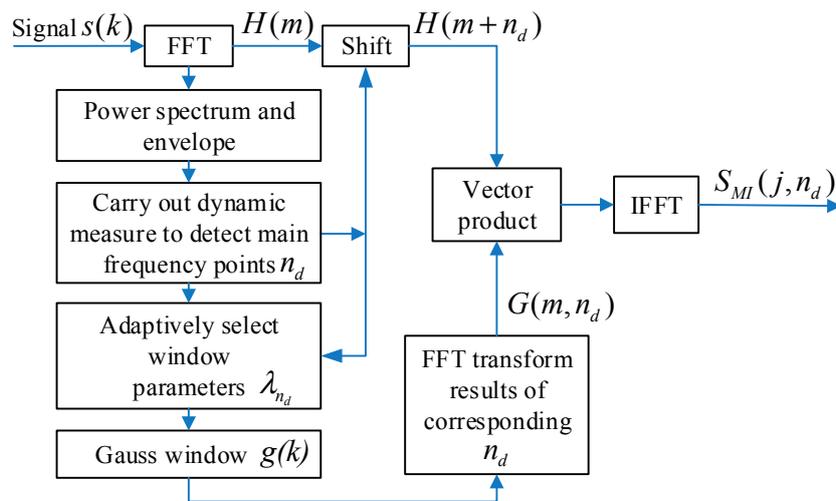


Figure 2. The rapid calculation procedure of MIST.

In Figure 2, FFT and IFFT respectively represent fast Fourier transform and inverse fast Fourier transform. $H(m)$ is FFT spectrum, and m is the point of FFT spectrum.

3. Recognition Plan

The recognition plan of combined PQ disturbances is shown in Figure 3.

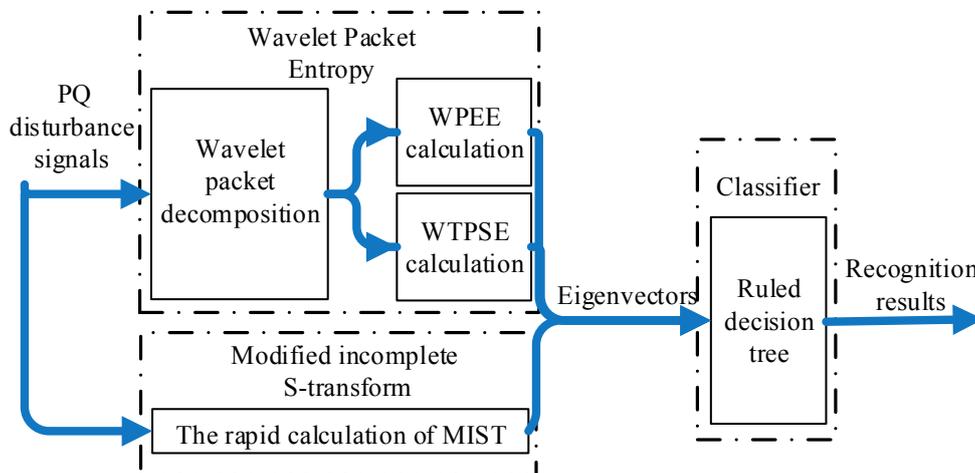


Figure 3. Combined PQ disturbance recognition plan.

3.1. Feature Extraction

In this paper, 13 significant features of PQ disturbances are extracted. According to different time-frequency analysis methods, these features can be divided into three classes.

(1) WPEE features

This class contains three features, named E_{av} , E_{std} and E_{bias} . They respectively represent the mean value, standard deviation and the bias of WPEE. According to Equation (3), E_{av} , E_{std} and E_{bias} can be calculated using the following formulas.

$$E_{av} = \frac{1}{L} \sum_{l=1}^L W_{EE} \tag{10}$$

$$E_{std} = \sqrt{\frac{1}{L} \sum_{l=1}^{l=L} [W_{EE} - E_{av}]^2} \tag{11}$$

$$E_{bias} = 2 \max \{ \max(W_{EE}) - 0.5, 0.5 - \min(W_{EE}) \} \tag{12}$$

where L is the length of WPEE.

(b) WPTSE features

This class contains two features, named T_{av} and T_{bias} . They respectively represent the mean value and the bias of WPTSE. According to Equation (8), T_{av} and T_{bias} can be calculated using the following formulas.

$$T_{av} = \frac{1}{L} \sum_{l=1}^L W_{TsallisPSE} \tag{13}$$

$$T_{bias} = 2 \max \{ \max(W_{TsallisPSE}) - 0.5, 0.5 - \min(W_{TsallisPSE}) \} \tag{14}$$

(c) MIST features

There are eight features in total, S_{av} , S_{bias} , S_{max} , S_{-1} , S_{-2} , N_f , N_h and N_1 . In this class, S_{av} , S_{bias} , S_{max} and S_{-1} respectively represent the average amplitude, the bias, the maximum of fundamental component and the amplitude fluctuation of fundamental component obtained after MIST. The fundamental time-amplitude curve obtained after MIST is

$$V_{fb}(t) = A(t, f_b) \tag{15}$$

where t is the sampling moment, and f_b is the fundamental frequency.

S_{av} , S_{bias} , S_{max} , and S_{-1} can be calculated using the following formulas.

$$S_{av} = \frac{1}{N} \sum_{n=1}^N A(t, f_b) \tag{16}$$

where N means the total sampling points.

$$S_{bias} = 2 \max \{ \max[A(t, f_b)] - 0.5, 0.5 - \min[A(t, f_b)] \} \tag{17}$$

$$S_{max} = \max[A(t, f_b)] \tag{18}$$

$$S_{-1} = \left\{ \left(\sum_{i=1}^{n_{\max}} d_{\max}(i) - \max(d_{\max}) \right) / (n_{\max} - 1) \right\} - \left\{ \left(\sum_{i=1}^{n_{\min}} d_{\min}(i) - \min(d_{\min}) \right) / (n_{\min} - 1) \right\} \quad (19)$$

In addition, the other four features S_{-2} , N_f , N_h and N_1 are extracted via dynamic measurement of power spectrum obtained after FFT. S_{-2} represents the symmetry of main frequency points obtained through dynamic measurement of FFT power spectrum. The values of S_{-2} are 1 or 0, which can be judged by the following formula.

$$\left| |f_{down} - f_b| - |f_{up} - f_b| \right| \leq 5 \quad (20)$$

where f_b is the fundamental frequency, f_{down} corresponds to the single frequency that is lower than fundamental frequency, and f_{up} is the frequency that is higher than f_b . The main function of S_{-2} is to reflect whether there is voltage fluctuation in the disturbance signal that contains voltage sag, voltage swell and voltage interruption.

N_f , the number of main frequency points, is first used to identify oscillation transient and harmonic by judging whether $N_f = 1$ or not. N_h can characterize the existence of the main frequency point in the high frequency band of the signal, which is the criterion of oscillation transient. $N_1 = 0$ or $N_1 = 1$, can represent whether the harmonic exists.

Table 1. Description of significant features.

No	Method	Name	Description	Threshold	Function
1	WPEE	E_{av}	Mean value	1.2	Oscillation assistant judgment
2		E_{std}	Standard deviation	0.17, 0.8	Impulsive/oscillation assistant judgment
3		E_{bias}	Bias	1.1, 4.8	Interruption/impulsive assistant judgment
4	WPTSE	T_{av}	Mean value	0.091, 0.35	Oscillation/impulsive assistant judgment
5		T_{bias}	Bias	0.8	Harmonic assistant judgment
6	MIST	N_f	Number of main frequency points	-	Oscillation/harmonic initial judgment
7		N_h	Whether it contains high frequency	0, 1	Oscillation initial judgment
8		N_1	Whether it contains Harmonic	0, 1	Harmonic initial judgment
9		S_{av}	Average amplitude of the fundamental components	0.475, 0.495	Swell/sag/interruption initial judgment
10		S_{bias}	Bias of the fundamental component	0.19, 0.85	Swell/sag/interruption initial judgment
11		S_{\max}	Maximum of the fundamental component	0.4807, 0.57	Swell/sag/interruption assistant judgment
12		S_{-1}	Amplitude fluctuation of fundamental components	0, 1	Fluctuation assistant judgment
13	S_{-2}	The symmetry of main frequency points	0, 1	Fluctuation assistant judgment	

Note: sign “-” means no Threshold.

The further descriptions of all the features are listed in Table 1. In order to test all the features, the normal signal and seven single PQ disturbances are chosen as the test objects, including normal signal, voltage swell, voltage sag, voltage interruption, impulsive transient, oscillation transient, harmonics, and voltage fluctuation (labeled as **R0~R7** respectively for the convenience of expression, namely, **R0**-normal signal, **R1**-voltage swell, **R2**-voltage sag, **R3**-voltage interruption, **R4**-impulsive transient, **R5**-oscillation transient, **R6**-harmonics, **R7**-voltage fluctuation). Referring to the related IEEE standards in [25], 200 samples of each single disturbance are randomly produced as feature extraction test signals. The fundamental frequency of the signals is 50 Hz, the sampling frequency and point number are 6.4 kHz and 2048, respectively. Examples are shown in Figure 4, and the 13 dimensional features for the examples in Figure 4 are listed in Table 2.

Using the extraction algorithm given in this paper, the distributions of main features of single PQ disturbances are shown in Figure 5a to Figure 5h. After the feature extraction, a 13-dimension vector is obtained as the input of the classifier.

Table 2. The 13-dimensional features for the examples in Figure 4.

Disturbances	E_{av}	E_{std}	E_{bias}	T_{av}	T_{bias}	N_f	N_h	N_l	S_{av}	S_{bias}	S_{max}	S_1	S_2
Normal (R0)	0.0006	0.0009	0.9991	0.0882	0.8392	1	0	0	0.4806	0.0387	0.4806	0	0
Swell (R1)	0.0016	0.0041	0.9991	0.1377	0.8398	1	0	0	0.5576	0.4072	0.7036	0	0
Sag (R2)	0.0022	0.0065	0.9991	0.1374	0.8824	1	0	0	0.4392	0.2841	0.4806	0	0
Interruption (R3)	0.0986	0.4007	4.8617	0.2519	3.5527	1	0	0	0.3537	0.9615	0.4806	0	1
Impulsive (R4)	0.0686	0.3500	3.4339	0.1809	3.4694	1	0	0	0.4767	0.1130	0.4806	0	0
Oscillation (R5)	1.5776	1.0600	3.7398	0.6735	2.1716	2	1	0	0.4811	0.0387	0.4830	0	0
Harmonics (R6)	0.0409	0.0027	0.9322	0.5730	0.1735	2	0	1	0.4806	0.0387	0.4806	0	0
Fluctuation (R7)	0.0007	0.0010	0.9991	0.0891	0.8423	1	0	0	0.4808	0.1402	0.5314	1	1

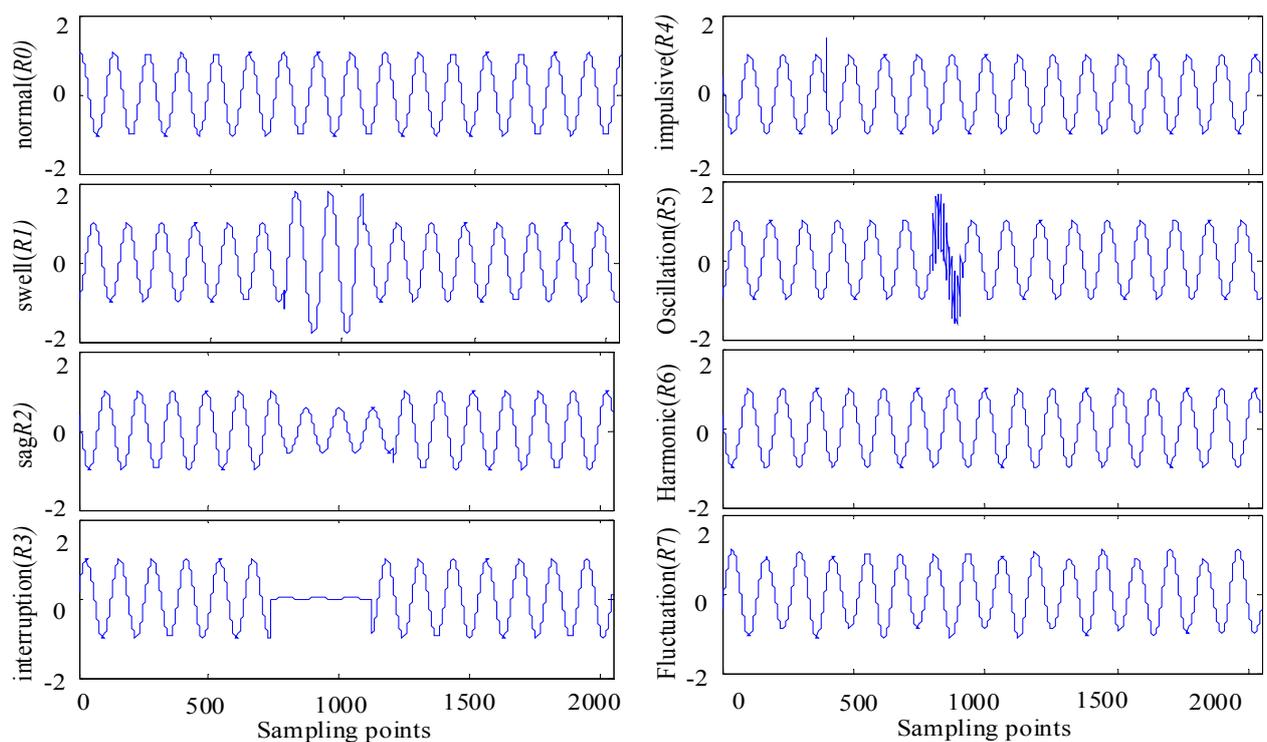


Figure 4. Examples of each single PQ disturbance.

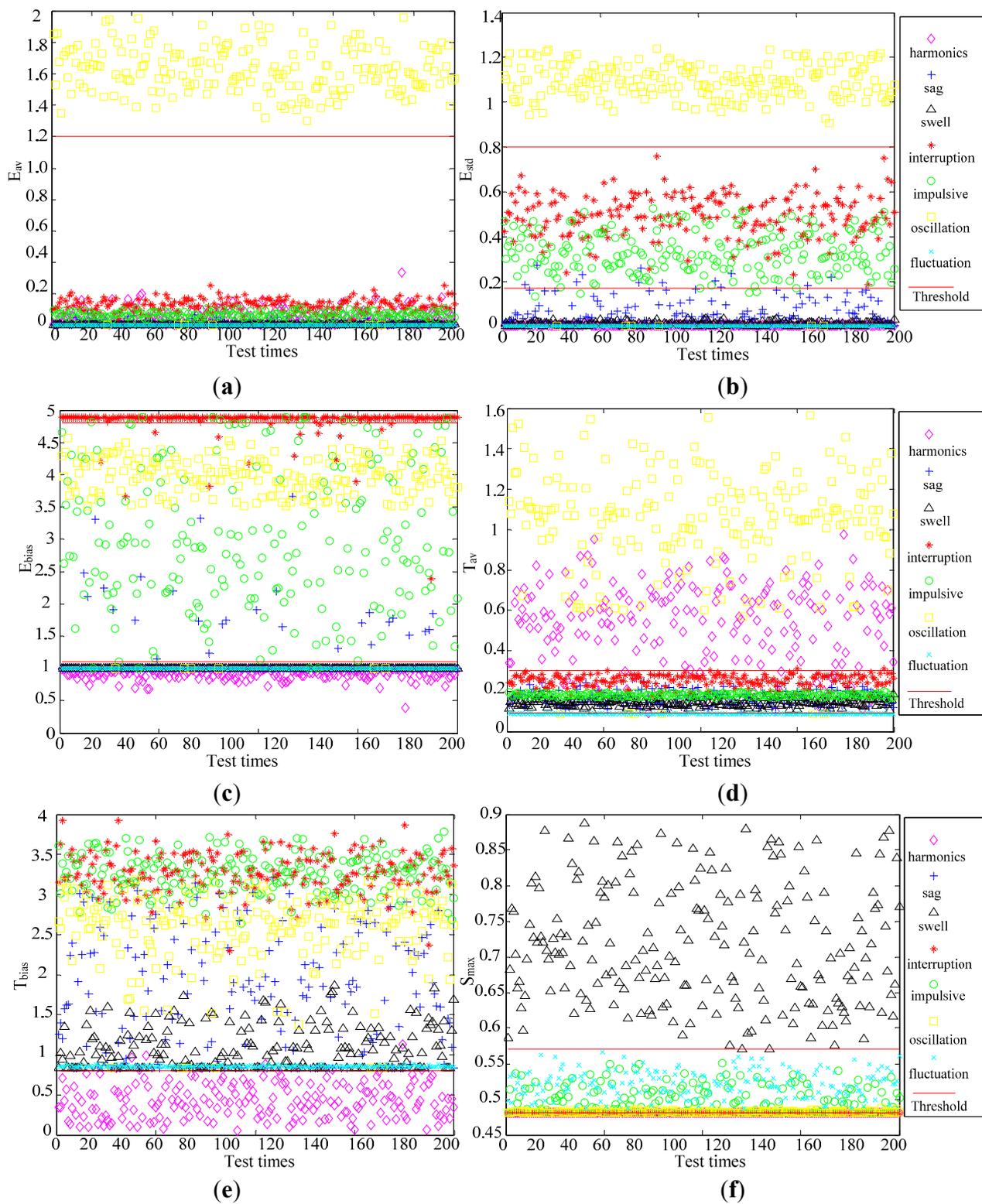


Figure 5. Cont.

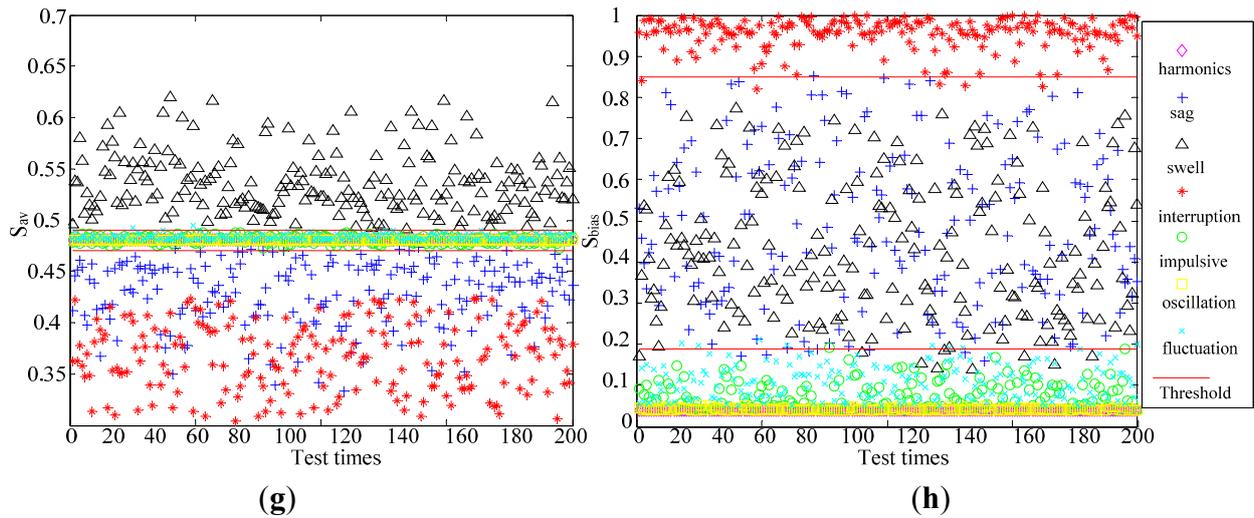


Figure 5. The distribution of each feature quantity (a) Add WPEE feature E_{av} distribution; (b) WPEE feature E_{std} distribution; (c) WPEE feature E_{bias} distribution; (d) WPTSE feature T_{av} distribution; (e) WPTSE feature T_{bias} distribution; (f) MIST feature S_{max} distribution; (g) MIST feature S_{av} distribution; (h) MIST feature S_{bias} distribution.

3.2. Ruled Decision Tree

Compared with other classifiers, the decision tree is a direct-vision method. Moreover, it takes a shorter amount of time than most of the other classifiers with an ideal classification result. In this paper, the interactions between different disturbances and the failure situation of features are fully considered. Through sufficient analysis of the feature extracting in the last section, a ruled decision tree is designed. The rules are listed in Table 3. All the threshold values selected in Table 3 are obtained through specific analysis and multiple repeat tests, considering both the threshold coverage and the recognition accuracy.

Table 3. Rules of the decision tree.

Rule	Description
Rule1	if $S_{av} > 0.495$ & $0.19 < S_{bias} < 0.85$ & $S_{max} > 0.57$ then $R1 = 1$
Rule2	if $S_{av} < 0.475$ & $0.19 < S_{bias} < 0.85$ & $S_{max} < 0.4807$ then $R2 = 1$
Rule 3	if $S_{av} < 0.475$ & $S_{bias} > 0.85$ & $S_{max} < 0.4807$ then $R3 = 1$
Rule 4	if $0.17 < E_{std} < 0.8$ & $0.091 < T_{av} < 0.35$ & $1.1 < E_{bias} < 4.8$ & $N_h = 0$ & $N_1 = 0$ then $R4 = 1$
Rule 5	if $N_f > 1$ & $N_h = 1$ & $E_{av} > 1.2$ & $E_{std} > 0.8$ then $R5 = 1$
Rule 6	if $N_f > 1$ & $N_1 = 1$ & $T_{bias} < 0.8$ then $R6 = 1$
Rule 7	if $R1 R2 R3 = 1$ & $R4 = 1$ then $R7 = S_1$ & S_2 else if $R1 R2 R3 = 1$ & $R4 = 0$ then $R7 = S_2$ else if $R1 R2 R3 = 0$ & $R4 = 1$ then $R7 = S_1$ else $R1 R2 R3 = 0$ & $R4 = 0$ then $R7 = S_1 S_2$

3.3. Recognition Flow

The recognition flow of the proposed approach is shown in Figure 6.

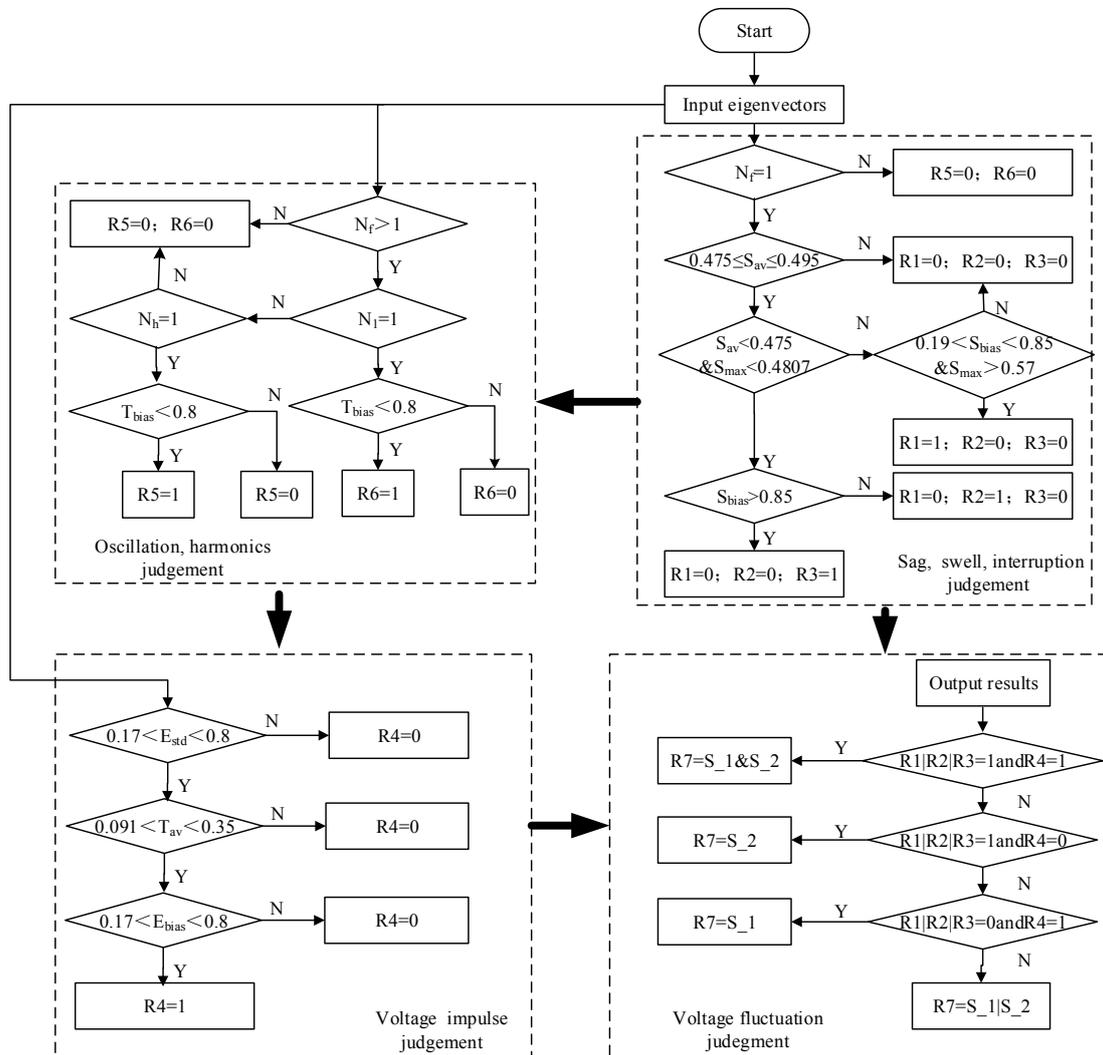


Figure 6. PQ disturbances recognition flow

4. Experimental Results

The combined PQ disturbances in the power system are variable and complex. It is difficult to obtain all kinds of real life signals. In order to verify the effectiveness of the recognition method proposed in this paper, MATLAB 2013a is adopted to simulate the disturbance signals according to standard of the IEEE. The seven signal PQ disturbances introduced above are chosen, and 14 mixed signals combined with the seven single disturbances are produced. The fundamental frequency is 50 Hz, the sampling frequency and point number are 6.4 kHz and 2048, respectively. Two hundred samples of each disturbance under SNR 40 dB are randomly produced. The sample sum is 200×22 . It means there are in total 22 (1 + 7 + 14) types of disturbances as listed in Table 4 and Table 5. Each kind of disturbance will produce 200 samples used to test the recognition accuracy. The vector length for each signal is 13. The experimental results are shown in Tables 4 and 5.

Table 4. The recognition results of single disturbances.

Disturbance Type	Recognition Results							Number of Right Samples	Accuracy/%	Time/s
	R1	R2	R3	R4	R5	R6	R7			
swell	191	0	0	0	0	0	0	191	95.5	0.014
sag	0	189	10	0	0	0	0	189	94.5	0.018
interruption	0	5	195	0	0	0	0	195	97.5	0.014
impulsive	0	0	0	185	0	0	0	185	92.5	0.014
oscillation	0	0	0	0	194	0	0	194	97	0.011
harmonics	0	0	0	0	0	199	0	199	99.5	0.011
fluctuation	0	0	0	0	0	0	200	200	100	0.013

Total sample identification accuracy: 96.64%; the mean consuming time to recognize a power quality disturbance: 0.0000678571 s.

Table 5. The recognition results of combined disturbances.

Disturbance Type	Recognition Results							Number of Right Samples	Accuracy/%	Time/s
	R1	R2	R3	R4	R5	R6	R7			
R1 + R6	193	0	0	0	0	199	0	193	96.5	0.017
R2 + R5	0	194	0	0	196	0	0	194	97	0.017
R2 + R6	0	195	0	0	0	199	0	195	97.5	0.015
R2 + R7	0	188	0	0	0	0	197	188	94	0.013
R3 + R6	0	0	195	0	0	195	0	195	97.5	0.016
R5 + R6	0	0	0	0	195	196	0	195	97.5	0.011
R5 + R7	0	0	0	0	194	0	200	194	97	0.013
R6 + R7	0	0	0	0	0	200	200	200	100	0.013
R2 + R5 + R6	0	192	0	0	198	199	0	192	96	0.013
R2 + R5 + R7	0	186	0	0	193	0	191	186	93	0.014
R2 + R6 + R7	0	182	0	0	0	199	194	182	91	0.016
R3 + R5 + R6	0	0	188	0	195	180	0	180	90	0.018
R5 + R6 + R7	0	0	0	0	195	197	200	195	97.5	0.020
R1+R4+R6+R7	181	0	0	194	0	198	196	181	90.5	0.020

Total sample identification accuracy: 95.36%; the mean consuming time to recognize a power quality disturbance: 0.0000817857s.

In order to evaluate the proposed disturbance recognition method more reasonably, two recognition indexes are defined as follows.

- (1) Sample recognition accuracy. This index considers the overall recognition accuracy of the sample. It is a traditional pattern recognition evaluation method. The calculation formula is as follows.

$$Sample\ recognition\ accuracy = \frac{the\ number\ of\ correct\ recognition}{the\ total\ number\ of\ samples} \times 100\% \tag{21}$$

- (2) Label error (leak) rate. This index considers the number of recognition error and leakage in recognition results of all samples. It reflects the stability of the proposed recognition method for single disturbance in the case of different combined disturbances. The calculation formula is as follows.

$$\text{Label error (leak) rate} = \frac{\text{the number of error and leakage recognition}}{\text{the total number of samples}} \times 100\% \tag{22}$$

As seen in Tables 4–6, the accuracy of the method proposed in this paper is high (the average of single disturbances and combined disturbances are respectively 96.64% and 95.36%). The statistic of recognition error and leakage of single disturbance in combined disturbances shows that the total recognition error and leakage rate is 0.770%. In addition, due to the small amplitude and noise effect, several cases cannot be completely accurately classified.

Table 6. The recognition error and leakage of single disturbance in combined disturbances.

Disturbance type	R1	R2	R3	R4	R5	R6	R7
Total sample number	14 × 200 = 2800						
Recognition error and leakage number	26	31	17	6	27	37	7
Recognition error and leakage rate/%	0.928	1.107	0.607	0.214	0.964	1.321	0.250

Total recognition error and leakage rate: 0.77%.

In order to further test the performance of the proposed method, some real-life signals are adopted. Figure 7a–d are the normalized real-life signals from a 110 kV bus of a 220 kV substation in North China power grid. The sampling frequency is 25.6 kHz and the interception is 15 cycles in 0.3 s. The recognition results are shown in Table 7. It can be found that the proposed method can efficiently recognize the real-life combined PQ disturbance signals, such as swell, impulsive, sag + oscillation and interruption + oscillation.

In order to show the advantage of the proposed method in this paper, some comparisons are made. The recognition methods based on improved incomplete S-transform with decision tree [26], wavelet transform with neural network [27] and EEMD (Ensemble Empirical Mode Decomposition) and MIST with automatic classification [28] are respectively applied. The test samples produced previously in Section 4 are also used. Their recognition performance is listed in Table 8.

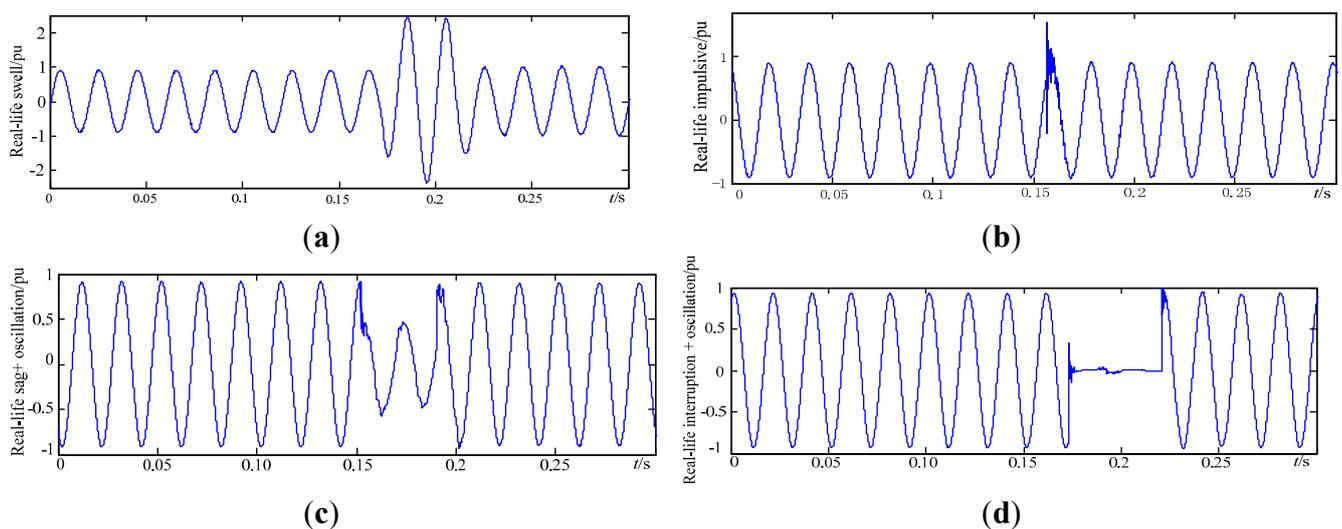


Figure 7. The normalized real-life signals (a) swell signal; (b) impulsive signal; (c) sag + oscillation signal; (d) interruption + oscillation signal.

Table 7. The recognition results of real life disturbance signals.

Disturbances	<i>R1</i>	<i>R2</i>	<i>R3</i>	<i>R4</i>	<i>R5</i>	<i>R6</i>	<i>R7</i>
Swell	1	0	0	0	0	0	0
Impulsive	0	0	0	1	0	0	0
Sag + Oscillation	0	1	0	0	1	0	0
Interruption + Oscillation	0	0	1	0	1	0	0

Table 8. Recognition performance comparison with some existing methods.

Method	Accuracy/%	
	Single	Combined
Improved incomplete S-transform with decision tree	81.86	88.93
Wavelet transform with neural network	94.42	83.33
EEMD and MIST with automatic classification	97.70	88.70
Wavelet Packet Entropies and MIST with decision tree (proposed)	96.64	95.36

According to the results in Table 8, the proposed method in this paper shows better classification ability both for single and combined PQ disturbances. Some existing methods are not able to recognize specific disturbances. For example, the method based on improved incomplete S-transform with decision tree is invalid for the recognition of impulsive transient. Some existing methods show poor recognition accuracy for the combined PQ disturbances, such as the classifier using wavelet transform with neural network and EEMD and MIST with automatic classification.

5. Conclusions

The main contribution of this paper is to present a new recognition approach of combined power quality disturbances. In order to extract the features of combined disturbances, the wavelet packet decomposition is combined with information entropies (Shannon entropy and Tsallis entropy). The multi-resolution analysis (MRA) ability of wavelet packet decomposition and the complexity estimation ability of information entropies ensure reliable feature extraction. In addition, MIST is used to obtain more detailed time-frequency features. After feature extraction, a 13-dimension vector is obtained as the input of a ruled decision tree. Simulation experiments and a real life signal test show its effectiveness and practicability.

Acknowledgments

This work is supported in part by the National Nature Science Foundation (U1434203, U1134205, 51377136).

Author Contributions

The individual contribution of each co-author to the reported research and writing of the paper are as follows. Zhigang Liu conceived the idea, Yan Cui and Wenhui Li performed experiments and data analysis, and Zhigang Liu and Yan Cui wrote the paper. All authors have read and approved the final manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Valtierra-Rodriguez, M.; de Jesus Romero-Troncoso, R.; Osornio-Rios, R.A.; Garcia-Perez, A. Detection and classification of single and combined power quality disturbances using neural networks. *IEEE Trans. Ind. Electron.* **2014**, *61*, 2473–2482.
2. Ray, P.K.; Mohanty, S.R.; Kishor, N.; Catalao, J.P.S. Optimal feature and decision tree-based classification of power quality disturbances in distributed generation systems. *IEEE Trans. Sustain. Energy* **2014**, *5*, 200–208.
3. Manikandan, M.S.; Samantaray, S.R.; Kamwa, I. Detection and classification of power quality disturbances using sparse signal decomposition on hybrid dictionaries. *IEEE Trans. Instrum. Meas.* **2015**, *64*, 27–38.
4. Hajian, M.; Foroud, A.A. A new hybrid pattern recognition scheme for automatic discrimination of power quality disturbances. *Measurement* **2014**, *51*, 265–280.
5. Ferreira, D.D.; Seixas, J.M.D.; Cerqueira, A.S. A method based on independent component analysis for single and multiple power quality disturbance classification. *Electr. Power Syst. Res.* **2015**, *119*, 425–431.
6. Lima, M.A.A.; Coury, D.V.; Cerqueira, A.S.; Nascimento, V.H. A method based on Independent Component Analysis for adaptive decomposition of multiple power quality disturbances. *J. Control Autom. Electr. Syst.* **2014**, *25*, 80–92.
7. Huang, N.; Zhang, S.; Cai, G.; Xu, D. Power quality disturbances recognition based on a multiresolution generalized S-transform and a PSO-improved decision tree. *Energies* **2015**, *8*, 549–572.
8. Liu, Z.G.; Cui, Y.; Li, W.H. A classification method for complex power quality disturbances using EEMD and rank wavelet SVM. *IEEE Trans. Smart Grid* **2015**, *6*, 178–1685.
9. 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*, 1249–1257.
10. Yong, D.D.; Bhowmik, S.; Magnago, F. An effective power quality classifier using wavelet transform and support vector machines. *Expert Syst. Appl.* **2015**, *42*, 6075–6081.
11. Dalai, S.; Dey D.; Chatterjee B.; Chakravorti, S. Cross-spectrum analysis based scheme for multiple power quality disturbance sensing device. *Sens. J. IEEE* **2015**, *15*, 3989–3997.
12. Costa, F.B. Boundary wavelet coefficients for real-time detection of transients induced by faults and power-quality disturbances. *IEEE Trans. Power Deliv.* **2014**, *29*, 2674–2687.
13. Kanirajan, P.; Suresh Kumar, V. Power quality disturbance detection and classification using wavelet and RBFNN. *Appl. Soft Comput.* **2015**, *35*, 470–481.
14. Rodriguez, A.; Aguado, J.A.; Martin, F.; Ruiz, J.E. Rule-based classification of power quality disturbances using S-transform. *Electr. Power Syst. Res.* **2012**, *86*, 113–121.

15. Zhao, F.Z.; Yang, R.A. Power-quality disturbance recognition using S-transform. *IEEE Trans. Power Deliv.* **2007**, *22*, 944–950.
16. Li, L.; Yi, J.L.; Zhu, J.L. Parameter estimation of power quality disturbances using modified incomplete S-transform. *Trans. China Electro Tech. Soc.* **2011**, *26*, 187–193.
17. Poisson, O.; Rioual, P.; Meunier, M. Detection and measurement of power quality disturbances using wavelet transform. *IEEE Trans. Power Del.* **2000**, *15*, 1039–1044.
18. Liu, Z.G.; Han, Z.W.; Zhang, Y. Multiwavelet packet entropy and its application in transmission line fault recognition and classification. *IEEE Trans. Neural Netw. Learning Syst.* **2014**, *25*, 2043–2052.
19. Dewal, K.Y.; Lal, M.; Shyam, A.R. Wavelet energy and wavelet entropy based epileptic brain signals classification. *Biomed. Eng. Lett.* **2012**, *2*, 147–157.
20. Chen, J.; Li, G. Tsallis wavelet entropy and its application in power signal analysis. *Entropy* **2014**, *16*, 3009–3025.
21. Liu, Z.G.; Hu Q.L.; Cui, Y.; Zhang, Q. A new detection approach of transient disturbances combining wavelet packet and Tsallis entropy. *Neurocomputing* **2014**, *142*, 393–407.
22. Milan, B.; 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**, *24*, 1071–1083.
23. Dash, P.K.; Mishra, S.; Salama, M.M.A.; Liew, A.C. Classification of power system disturbances using a fuzzy expert system and a Fourier linear combiner. *IEEE Trans. Power Deliv.* **2000**, *15*, 472–477.
24. Tsallis, C. Possible generalization of Boltzmann-Gibbs statistics. *J. Stat. Phys.* **1988**, *52*, 479–487.
25. Chowdhury, H.B. Power quality. *IEEE Potentials* **2001**, *20*, 5–11.
26. Guo, J.W.; Li, K.C.; He, S.F.; Zhang, M. A real time power quality disturbance classification based on improved incomplete S-transform and decision tree. *Power Syst. Protect. Control* **2013**, *22*, 2473–2482.
27. Rodriguez, A.; Aguado, J.; Martin, F.; Muñoz, J.; Medina, M.; Ciumbulea, G. Classification of power quality disturbances using wavelet and artificial neural network. In Proceedings of 2010 International Conference on Power System Technology (POWERCON), Hangzhou, China, 24–28 October 2010; pp. 1–7.
28. Zhang, Y.; Liu, Z.G. A new method for power quality mixed disturbance classification based on time-frequency domain multiple features. *Proc. CSEE* **2012**, *34*, 83–90.