

ORIGINAL RESEARCH

Open Access



# Power quality disturbance classification based on time-frequency domain multi-feature and decision tree

Wenjing Zhao<sup>1\*</sup> , Liqun Shang<sup>1</sup> and Jinfan Sun<sup>2</sup>

## Abstract

Accurate classification of power quality disturbance is the premise and basis for improving and governing power quality. A method for power quality disturbance classification based on time-frequency domain multi-feature and decision tree is presented. Wavelet transform and S-transform are used to extract the feature quantity of each power quality disturbance signal, and a decision tree with classification rules is then constructed for classification and recognition based on the extracted feature quantity. The classification rules and decision tree classifier are established by combining the energy spectrum feature quantity extracted by wavelet transform and other seven time-frequency domain feature quantities extracted by S-transform. Simulation results show that the proposed method can effectively identify six types of common single disturbance signals and two mixed disturbance signals, with fast classification speed and adequate noise resistance. Its classification accuracy is also higher than those of support vector machine (SVM) and k-nearest neighbor (KNN) algorithms. Compared with the method that only uses S-transform, the proposed feature extraction method has more abundant features and higher classification accuracy for power quality disturbance.

**Keywords:** Power quality, Disturbance classification; wavelet transform, S-transform, Decision tree, Classification rules

## 1 Introduction

With the development of grid interconnection, grid-connection of new energy generation, extensive application of power electronic equipment and access of impact load, the problem of power quality disturbance has attracted more and more attention [1]. In-depth study of the power quality influencing factors, accurate extraction of feature quantities, and accurate classification of power quality disturbance are required for improving and controlling power quality [2].

The processes of power quality disturbance classification consist of feature extraction and classification recognition. The methods of feature extraction mainly include Fast Fourier transform (FFT), Short-time Fourier transform (STFT), wavelet transform, S-transform, Hilbert yellow transform (HHT), etc. FFT is a conversion

from the time domain to the frequency domain, and has orthogonal and complete features. The frequency analysis of a signal is considered from the perspective of the overall composition of frequency, but the local frequency characteristics of the signal cannot be analyzed. Thus, it is only suitable for the analysis of steady-state disturbance [3]. STFT has fast computation speed and the algorithm is easy to implement. It can detect and analyze the signal's local spectrum features, but its window function is fixed with no ability of self-adaptation [4, 5]. Wavelet transform has the ability of multi-scale time-frequency resolution, which can be used for local analysis of signals, but signal analysis can be easily influenced by wavelet base and decomposition layer [6, 7]. S-transform is developed on the basis of wavelet transform and STFT. It not only overcomes their shortcomings, but also enables the analysis of amplitude change with time of a certain frequency component of the signal. Its window function changes with frequency, resulting in higher frequency resolution but also large amount of calculation [8, 9]. HHT is suitable for time-frequency analysis detecting methods of non-

\* Correspondence: [17782897852@163.com](mailto:17782897852@163.com)

<sup>1</sup>School of Electrical and Control Engineering, Engineering, Xi'an University of Science and Technology, No.58, Yanta Road, Bei Lin District, Xi'an, Shaanxi Province, China

Full list of author information is available at the end of the article

stationary and nonlinear signals, but is easy to generate modal aliasing during analysis [10]. At present, the main methods for power quality disturbance classification are artificial neural network, support vector machine (SVM), decision tree, K-neighbor (KNN), etc. Artificial neural network has long training time and is easy to fall into local optimal solution [11, 12] whereas SVM is susceptible to kernel function and cannot take into account both learning ability and generalization ability [13]. KNN classification requires large amount of computation and large memory [14], while the decision tree has the advantages of simple structure, convenient expansion and fast classification [15, 16].

Based on this, this paper mainly analyzes 6 types of single power quality disturbances, and the compound disturbances of swell + harmonic and sag + harmonic. Wavelet transform and S-transform are combined to extract more abundant feature quantities in time and frequency domains. According to the extracted feature quantities, the classification rules suitable for the 8 disturbance signals are established, and thus, accurate classification effect of disturbance signals can be obtained quickly by constructing decision tree classifier. Due to noise interference in actual power systems, the noise resistance of the proposed method is verified by adding gaussian white noise. The classification speed and accuracy of the proposed method are verified by simulation comparison.

## 2 Methods

This paper classifies 8 different disturbances, including the standard voltage( $C_0$ ), voltage swell( $C_1$ ), voltage sag( $C_2$ ), voltage interruption( $C_3$ ), transient oscillation( $C_4$ ), flicker( $C_5$ ), harmonic( $C_6$ ), swell + harmonic ( $C_7$ ) and sag + harmonic( $C_8$ ).

### 2.1 Feature extraction based on wavelet transform and S-transform

#### 2.1.1 Wavelet transform extracts energy feature quantity

Wavelet transform is used for multi-scale decomposition of power quality disturbance signals. The obtained wavelet coefficients reflect the distribution of signals on different decomposition scales, and their differences after decomposition of different disturbance signals can be used to represent the signals' feature quantities. Due to the large amount of data of the wavelet coefficients as the feature quantities, the wavelet energy of different decomposition scales can be calculated through the wavelet coefficients to significantly reduce the feature dimension.

Power quality disturbances are nonlinear mutation signals. In order to analyze the disturbances such as transient oscillations and harmonics, which mainly cause high frequency band mutation, and improve the operation speed of the wavelet transform, the wavelet basis function needs

to satisfy the requirement of tight support, orthogonality, higher vanishing moments and calculation speed. The Daubechies (dbN) series wavelets in the basic wavelet have the above characteristics and are most commonly used in the Mallat algorithm. With the increase of N (the wavelet order), dbN wavelet in time domain increases the support interval of wavelet, while reduces the overlap of windows between different scales and the spectrum leakage between frequency bands. Higher-order db wavelets have higher vanishing moments, though larger N is not necessary better as the vanishing moments are opposite to the characteristics of tight support. Therefore, combined with the characteristics of power quality disturbance signals, db4 wavelet basis is selected in this paper for wavelet analysis. In order to reduce the spectrum leakage of wavelet transform, the main frequency components of transient components are distributed as far as possible in the center of the wavelet frequency band, and the disturbance signals are decomposed by 10 layers. The wavelet energy of each decomposition layer [17] is:

$$E_{D_j} = \sum_{k=1}^N cd_j(k)^2 \quad (1)$$

where,  $cd_j$  is the detail coefficient of layer j, and  $N$  is the number of detail coefficient of layer j. The normalized value distribution of wavelet energy for the 8 disturbances is shown in Fig. 1.

It can be observed from Fig. 1 that the wavelet energy of each power quality disturbance signal is mainly concentrated in the 6th and 7th layers, while the 7th layer wavelet energy of harmonic, swell + harmonic and sag + harmonic is noticeably lower compared to the others. To avoid contingency, each disturbance randomly generates 300 samples that are superimposed by 30 dB of noise for testing. Finally, the 7th layer energy is set to the feature  $F_j$ , and when the threshold value is 0.56, the harmonic and harmonic-containing disturbances are recognized with high precision.

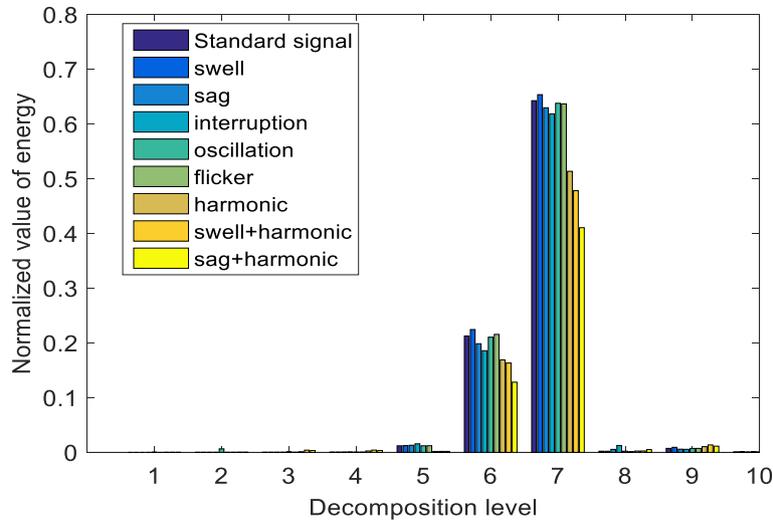
#### 2.2 S-transform extracts feature quantity

The S-transform is a reversible time-frequency analysis method proposed by Stockwell. Because the height and width of window function vary with frequency, it has the advantages of both the WT and STFT, and thus is widely used. Continuous S-transformation is defined as:

$$S(\tau, f) = \int_{-\infty}^{\infty} h(t)g(\tau-t, f)e^{-j2\pi ft} dt \quad (2)$$

$$g(\tau-t, f) = \frac{|f|}{\sqrt{2\pi}} e^{-\frac{f^2(\tau-t)^2}{2}} \quad (3)$$

where  $h(t)$  is the disturbance signal, and  $g(\tau - t, f)$  is the Gaussian window. In practical applications, the signal is obtained by sampling. Let the sampled signal be  $h[kT]$ ( $k =$



**Fig. 1** Comparison of wavelet energy distribution of different power quality disturbance signals

0,1,2...N - 1), where N is the number of sampling points and T is the sampling period, then the expression of discrete S-transformation is given as:

$$\begin{cases} S[kT, \frac{n}{NT}] = \sum_{m=0}^{N-1} H(\frac{m+n}{NT}) G(m, n) e^{\frac{j2\pi mnk}{N}}, n \neq 0 \\ S[kT, 0] = \frac{1}{N} \sum_{m=0}^{N-1} H(\frac{m}{NT}), n=0 \end{cases} \quad (4)$$

where j, m, n = 0,1,2... N - 1, H(n/NT) and G(m, n) are the FFT of signal H[kT] and Gaussian window respectively, And are given as:

$$\begin{cases} H[kT] = \frac{1}{N} \sum_{k=0}^{N-1} h[kT] e^{-\frac{j2\pi mk}{N}} \\ G(m, n) = e^{-\frac{2\pi^2 m^2}{n^2}} \end{cases} \quad (5)$$

The result of S-transformation is a two-dimensional complex matrix, which is modeled to obtain the modulus matrix. Its row vectors represent the change of amplitude of a certain frequency with time, and the column vectors represent the change of amplitude with frequency at a sampling moment. It reflects the time-frequency characteristics of the signal. If there is disturbance, it must be shown in the modulus matrix. According to the time-frequency matrix, the amplitude and frequency mutation of the disturbance can be detected.

According to IEEE’s relevant standards for power quality disturbances and the principles of their generation, it is concluded that swell, sag, interruption and flicker occur mainly in amplitude mutation, with the amplitude of flicker changing periodically. Harmonic and transient oscillations occur mainly in frequency mutation. S-transform has been carried out on standard signals and the 8 disturbances. Based on the differences in time, amplitude and frequency of each disturbance signal, the following characteristics are extracted.

(1) The maximum value  $F_2$ , minimum value  $F_3$  and standard deviation  $F_4$  of the maximum amplitude vector of time can be calculated as:

$$F_2 = \max[V_{1t-A}(kT)] \quad (6)$$

$$F_3 = \min[V_{1t-A}(kT)] \quad (7)$$

$$F_4 = \sqrt{\frac{1}{N} \sum_{k=0}^{N-1} [V_{1t-A}(kT) - \bar{V}_1]^2} \quad (8)$$

where,  $V_{1t-A}$  is the largest amplitude vector of time,  $k = 0, 1, 2, \dots, N - 1$ .

(2) Standard deviation  $F_5$  of the maximum frequency amplitude vector of 100 - 600 Hz frequency band is calculated as:

$$F_5 = \sqrt{\frac{1}{N-1} \sum_{f=100}^{600} [V_{f-maxA} - \bar{V}_m]^2} \quad (9)$$

where  $V_{f-maxA}$  is the maximum frequency amplitude vector, and  $\bar{V}_m$  is the average value of the maximum amplitude vector of the frequency band of 100 - 600 Hz. N is a sampling point in the range of 100 to 600 Hz.

(3) Mean value  $F_6$  and standard deviation  $F_7$  of the maximum frequency amplitude vector in the 700 - 1500 Hz frequency band are calculated as:

$$F_6 = \frac{f_0}{\Delta f} \sum_{f=700}^{1500} [V_{f-maxA}(f)] \quad (10)$$

$$F_7 = \sqrt{\frac{1}{N-1} \sum_{f=700}^{1500} [V_{f-maxA} - \bar{V}_h]^2} \quad (11)$$

where  $f_0$  refers to the frequency resolution of 5 Hz. The frequency range  $\Delta f$  is 700 - 1500 Hz, and  $\bar{V}_h$  is the average value of the maximum amplitude vector of the

**Table 1** Analytical comparison of the disturbances and feature quantities

Disturbance types	F <sub>1</sub>	F <sub>2</sub>	F <sub>3</sub>	F <sub>4</sub>	F <sub>5</sub>	F <sub>6</sub>	F <sub>7</sub>	F <sub>8</sub>
C <sub>0</sub>	> 0.56	≈1	≈1	≈1	< 0.0009	≈0	≈0	0
C <sub>1</sub>	> 0.56	> 1.1	≈1	> 0.018, < 0.29	< 0.01	< 0.1	< 0.009	1
C <sub>2</sub>	> 0.56	≈1	> 0.1, < 0.9	> 0.017, < 0.32	< 0.01	< 0.12	< 0.01	1
C <sub>3</sub>	> 0.56	≈1	< 0.1	> 0.12, < 0.36	< 0.01	< 0.13	< 0.02	1
C <sub>4</sub>	> 0.56	≈1	≈1	< 0.002	< 0.004	> 0.045, < 0.3	> 0.02, < 0.13	0
C <sub>5</sub>	> 0.56	> 1.1	> 0.1, < 0.9	> 0.002	≈0	< 0.04	< 0.002	≥2
C <sub>6</sub>	≤0.56	≈1	≈1	≈1	> 0.01, < 0.06	< 0.006	< 0.012	0
C <sub>7</sub>	≤0.56	> 1.1	≈1	> 0.015, < 0.29	> 0.01, < 0.06	< 0.006	< 0.012	0
C <sub>8</sub>	≤0.56	≈1	> 0.1, < 0.9	> 0.02, < 0.33	> 0.01, < 0.06	< 0.006	< 0.012	0

frequency band of 700 – 1500 Hz, N is the sampling point in the 700-1500 Hz frequency band.

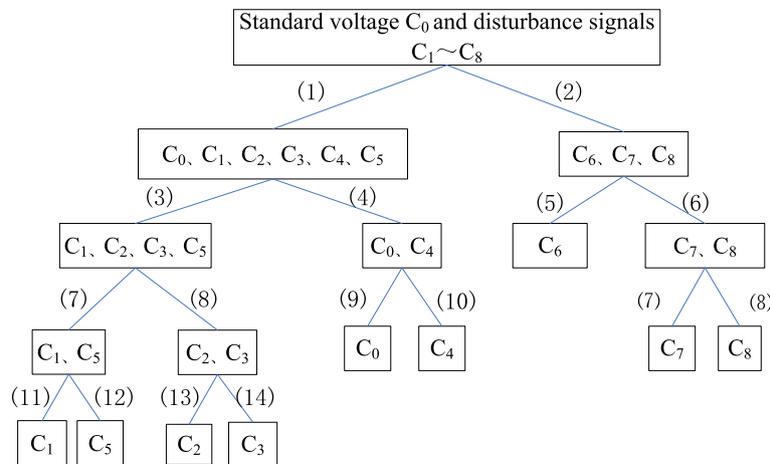
(4) The fluctuation time of the maximum amplitude curve is F<sub>8</sub>. Amplitude varying from small to large or from large to small is regarded as a fluctuation.

**2.3 The establish of decision tree classifier model**

Decision tree is a supervised learning algorithm, a kind of classifier similar to tree structure. The decision tree has the advantages of simple structure, convenient expansion and fast classification speed. It overcomes the disadvantages of SVM which is affected by kernel function, and is unable to give consideration to learning ability and generalization ability, and large computation and memory demand in KNN classification. Through classification rules to build the classification decision tree model. The structure of decision tree plays an important role in the accuracy of classification. In order to reduce the selection requirement of classification threshold and improve the classification accuracy, binary tree structure is adopted for classification.

Decision tree is usually used to recursively select the best feature and segment of the training data according to the feature, so as to optimize the classification process of each sub-data set. This process corresponds to the division of feature space and the construction of decision tree classification rules. According to the feature extraction method in Section 2.1, 300 groups of random disturbance samples are generated and 30 dB noise superimposed. The feature quantities F<sub>1</sub>~F<sub>8</sub> are calculated and statistically analyzed. The different disturbance types and feature quantities are compared in Table 1.

According to Table 1, the optimal feature is selected recursively and the following 14 classification rules are established: (1) F<sub>1</sub> > 0.56 and F<sub>5</sub> < 0.01; (2) F<sub>1</sub> < 0.56 and 0.01 < F<sub>5</sub> < 0.06; (3) F<sub>4</sub> > 0.002 and F<sub>7</sub> < 0.02; (4) F<sub>4</sub> < 0.002; (5) F<sub>4</sub> < 0.002 and F<sub>7</sub> < 0.02; (6) F<sub>4</sub> > 0.002 and F<sub>8</sub> < 2; (7) F<sub>2</sub> > 1.1; (8) F<sub>2</sub> < 1.1; (9) F<sub>6</sub> < 0.045 or F<sub>7</sub> < 0.02; (10) F<sub>6</sub> > 0.045 and F<sub>7</sub> > 0.02; (11) F<sub>8</sub> < 2 and F<sub>3</sub> > 0.9; (12) F<sub>8</sub> ≥ 2 and F<sub>3</sub> < 0.9; (13) 0.1 < F<sub>3</sub> < 0.9; (14) F<sub>3</sub> < 0.1. The corresponding power quality disturbance classification decision tree is constructed as shown in Fig. 2.



**Fig. 2** Decision tree of power quality disturbance classification

**Table 2** Comparison of recognition results of different feature extraction methods

Methods	Noise/dB				
	0	50	40	30	20
ST	99.3%	98.7%	98.6%	97.3%	92.6%
WT + ST	99.8%	99.68%	99.46%	98.8%	97.1%

**2.4 Simulation experiment results and discussion**

The mathematical model for power quality disturbances described in [9] is considered. Random noise-free disturbance samples with parameters are generated through MATLAB simulation for testing (300 samples of  $C_0 \sim C_8$  each). The sampling frequency is 6.4 kHz, the fundamental frequency is 50 Hz, the weekly wave sampling points are 128, and the data length is 1280 points. Wavelet transform is used to extract the normalized wavelet energy of the 7th layer, and S-transform is used to extract the feature quantities of the disturbances in time and frequency domains. The extracted feature variables are inputted into the constructed decision tree classifier to realize the recognition of the power quality disturbance signals. Considering that actual power system is affected by noise, 20 dB, 30 dB, 40 dB, 50 dB Gaussian white noises are superimposed onto the samples respectively to generate a total of 13,500 samples.

To verify the effectiveness of the power quality disturbance classification method based on time-frequency domain multi-feature and decision tree, Table 2 compares the classification accuracy of the disturbance signals by using only S-transform (method 1) and the combination of wavelet transform and S-transform (method 2) to extract feature quantities under different noise conditions. Table 3 shows the classification effect of decision tree, SVM and KNN on disturbance signals under different noise conditions, whereas Table 4 shows the time required for classification of each detection algorithm under the condition of SNR = 30 dB.

It can be seen from Table 2 that the classification effect of method 2 is better than that of method 1. The classification accuracies of both feature extraction methods decrease with the reduction of SNR, though the reduction of classification accuracy of method 1 is more significant than that of method 2. When SNR = 20 dB, the accuracy of method 2 is 97.1%, which is 4.5%

**Table 3** Comparison of recognition accuracy of different detection algorithms

Detection algorithms	Noise/dB				
	0	50	40	30	20
DT	99.78%	99.59%	99.59%	99.11%	97.81%
SVM	98.22%	97.67%	97.33%	96.78%	92.52%
KNN	98.93%	98.56%	98.48%	98.0%	96.78%

**Table 4** Classification time comparison of different detection algorithms

Classification of time	Detection algorithms		
	DT	SVM	KNN
t/s	1.14	3.42	2.53

higher than that of method 1. It indicates that the feature extraction method of wavelet transform + S-transform has better noise resistance and richer feature quantity than those of S-transform.

As shown in Table 3, the classification accuracies of DT, SVM and KNN algorithms decrease with the increase of noise intensity. When SNR = 20 dB, the accuracy of DT is 5.29% higher than that of SVM and 1.03% higher than that of KNN. In addition, as can be seen from Table 4, DT classification is faster than the other two methods.

**3 Conclusion**

For the various types of power quality disturbance, this paper proposes a power quality disturbance classification method based on time-frequency domain multi-feature and decision tree, for power quality improvement and governance. By combining the advantages of wavelet and S-transform, 8 time-frequency domain eigenvalues are extracted from 6 single disturbances and 2 compound disturbances. According to the extracted feature quantities, the classification rules of decision tree are established, and the decision tree model for classification is constructed. Simulation results show that the method is effective, and the extracted feature quantities can be effectively used for the classification, and classification of decision tree. Compared with only using S-transform, the proposed feature extraction method has richer feature quantities, higher classification accuracy and robustness to noise. For the feature quantities extracted in this paper, the classification accuracy of decision tree classifier is higher and the calculation speed faster than those of SVM and KNN.

The example in this paper is based on MATLAB simulation platform. Further research will try to apply the proposed method to practical power quality disturbance classification. Additional types of power quality disturbance will be included and the classification method will be made more universal.

**Acknowledgements**

The authors would like to thank Natural Science Basic Research Plan in Shaanxi Province of China for supporting the project.

**Authors' contributions**

Zhao WJ and Sun JF performed the simulation examination, analyzed and interpreted the simulation results. Shang LQ designed and supervised the experiment, prepared and revised the manuscript. All authors read and approved the final manuscript.

**Funding**

The Project is supported by Natural Science Basic Research Plan in Shaanxi Province of China (Program No. 2019JM-544).

**Availability of data and materials**

The power quality disturbance samples in this paper are generated by MATLAB using the power quality disturbance mathematical model.

**Competing interests**

The authors declare that they have no competing interests.

**Author details**

<sup>1</sup>School of Electrical and Control Engineering, Engineering, Xi'an University of Science and Technology, No.58, Yanta Road, Bei Lin District, Xi'an, Shaanxi Province, China. <sup>2</sup>Shaanxi Power Generation co., LTD. Weihe Thermal Power Plant, Xianyang 712000, China.

Received: 13 July 2019 Accepted: 31 October 2019

Published online: 24 December 2019

**References**

- Xiao, X.N. (2010). Analysis and control of power quality [M]. *Beijing: China Electric Power Press*, 124–128.
- Zhang, Y., & Liu, Z.G. (2012). A new power quality hybrid disturbance classification method based on time-frequency domain multi-characteristic quantities [J]. *Proceedings of the CSEE*, 32(34), 83–90.
- Zhang, B. (2010). Power quality analysis method based on Mallat algorithm and fast Fourier transform [C]. *Power Quality Seminar*, 35–40.
- Jurado, F., & Saenz, J. R. (2002). Comparison between discrete STFT and wavelets for the analysis of power quality events[J]. *Electr Power Syst Res*, 62(3), 183–190.
- Huang, J.M, Qu, H.Z, & Li, X.M. (2016). Classification of mixed disturbance of power quality based on short-time Fourier transform and spectral kurdiness [J]. *Power System Technology*, 40(10), 3184–3191.
- Qu, H.Z, Liu, H, Li, X.M, et al. (2017). A feature combination optimization method for multi-disturbance classification of power quality [J]. *Electric Power Automation Equipment*, 37(3), 146–152.
- Luciano C.M. Andrade, Mário Oleskovicz, & Ricardo A.S. Fernandes. (2016). Adaptive threshold based on wavelet transform applied to the segmentation of single and combined power quality disturbances[J]. *Neurocomputing*, 38, 967–977.
- Wu, Y., Tang, Q., Teng Z.S, et al. (2016). Power quality disturbance signal feature extraction method based on improved S transform [J]. *Proceedings of the CSEE*, 36(10), 2682–2689.
- Huang, N.T, Peng, H., Cai, G.W, et al. (2017). Composite disturbance feature selection and optimal decision tree construction of power quality [J]. *Proceedings of the CSEE*, 37(3), 776–786.
- Li, X.N. (2017). Power quality disturbance detection and recognition based on hilbert-huang transformation [D], China University of Mining and Technology.
- Anamika Yadav, Yajnaseni Dash & V. Ashok. (2016). ANN based directional relaying scheme for protection of Korba-Bhilai transmission line of Chhattisgarh state[J]. *Protection and Control of Modern Power Systems*, 1(1), 15.
- He, J.L, Wang, G.P, Liu, D., et al. (2017). Location and identification of power quality disturbance in distribution network system based on lifting wavelet and improving BP neural network [J]. *Power System Protection and Control*, 45(10), 69–76.
- Ren, Z.H, & Wang, Q. (2008). Power quality disturbance identification based on optimal DDAGSVM multi-class classification strategy [J]. *Power System Protection and Control*, 46(5), 82–88.
- Panigrahi B.K, Pandi V.R. optimal feature selection for classification of power quality disturbances using wavelet packet-based fuzzy k-nearest neighbour algorithm[J]. *Neurocomputing*, 2009, 3(3): 296–306.
- Biswal, M., & Dash, P.K. (2013). Measurement and classification of simultaneous power signal patterns with an S-transform variant and fuzzy decision tree[J]. *IEEE Transactions on Industrial Informatics*, 9(4), 1819–1827.
- Zhou, Z.N. (2017). Research on power quality disturbance identification algorithm based on S transform [D]. Harbin Institute of Technology.
- Han, G., Zhao, J.W, Zhu, X., et al. (2015). Power quality disturbance identification based on multi-feature combination [J]. *Proceedings of the CSU-EPSSA*, 27(8), 71–77.

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)