This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIE.2019.2952823, IEEE Transactions on Industrial Electronics

Classification of Complex Power Quality Disturbances Using Optimized S-Transform and Kernel SVM

Qiu Tang, Wei Qiu, *Student Member, IEEE*, and Yicong Zhou, *Senior Member, IEEE*

***Abstract*—Accurate power quality disturbance (PQD) classification is significantly important for power grid pollution control. However, the use of non-linear loads makes power system signals complex and distorted, and thus increases the difficulty of detecting and classifying PQD signals. To address this issue, this paper first proposes an optimized S-Transform (OST). It optimizes different window parameters to improve time-frequency resolution using maximum energy concentration. A kernel support vector machine (KSVM) classifier is proposed to classify multiple features using a combination of kernels. Integrating OST and KSVM, a classification framework is further proposed to detect and classify various PQD signals. Extensive experiments on computer simulations and experimental signals demonstrate that the proposed classification framework shows better performance than several state-of-art methods in classifying not only single and multiple PQD signals but also PQD signals with different noise levels. More importantly, our framework has superior performance in detecting nonlinearly mixed PQD signals.**

***Index Terms*—Kernel SVM, optimized S-Transform (OST), power quality disturbance (PQD), time-frequency resolution, nonlinearly mixed PQD**

I. INTRODUCTION

IN recent years, power quality issues have received widespread attention due to the fact that power networks commonly have a large number of non-linear loads such as automotive charging piles, power transfer switches, power electronics devices, and many others [1]. Additionally, the development of renewable energy like wind and geothermal energy has also a certain impact on the grid signals. The use of multiple loads and energy resources will generate different power quality disturbances (PQDs), such as swell, sag, transient and spike [2]. Meanwhile, multiple complex PQD signals are also generated from these single disturbances [3]. To establish a reliable and safe power supply system, accurate detection and classification of these PQD signals are important to deal with disturbance pollution problems. For example, real-time monitoring can be used for protection while offline detection is used to analyze signal components and compensation equipment [4]. To improve the classification performance of complex single and mixed PQD signals, effective signal detection and classification technologies are in demand.

Many detection frameworks have been developed to analyze PQD signals. These frameworks consist of two parts: time-frequency analysis and PQD classification. Since the features of PQD signals are inconspicuous, the time-frequency analysis methods are used to extract the characteristics of PQD signals in the time domain or frequency domain. In [5], both the time and frequency scales of PQD signals are decomposed using Gabor-Wigner Transform. Other examples include ensemble Empirical Mode Decomposition (EMD) [6], Short-Time Fourier Transform (STFT) [7] and Stockwell Transform (ST) [8]. The orthogonality between different intrinsic mode functions (IMFs) was improved using the orthogonal EMD [9]. However, the endpoint effect at the IMF component boundary is still in presence. Thereafter, because the fixed window parameters of STFT are not adaptive to the non-stationary signals, the discrete wavelet transform (DWT) was then chosen to analyze the stationary and transient components of PQD signals [10]. However, DWT is sensitive to noise. As an extension of DWT, various window parameter setting methods were proposed for ST, including adjustable window width and optimally concentrated discrete window [11], [12]. For example, a nonlinear Gaussian window standard deviation was integrated into the fast discrete ST (FDST) in [13]. The resolution of FDST is better than traditional ST, but the window parameters under complex PQD still need to be determined empirically. Moreover, one problem is that the resolution interval of time-frequency analysis between high and low frequencies is difficult to increase. Therefore, effective methods are needed to further improve the time-frequency accuracy of complex PQD signals.

In the PQD classification step, different statistical features, such as skewness, kurtosis and instantaneous harmonic distortion, are first extracted from the time-frequency information and then fed to the classifier [14]. It is noted that different kinds of features have a great impact on the classifier. Thereafter, various classification methods and their improved algorithms are proposed for accurate PQD classification. The Decision Tree (DT) was used to distinguish 13 commonly dis-

This work was supported in part by The Science and Technology Development Fund, Macau SAR (File no. 189/2017/A3), and by the Research Committee at University of Macau under Grants MYRG2016-00123-FST and MYRG2018-00136-FST. *(Corresponding author: Yicong Zhou)*.

Q. Tang and W. Qiu are with the College of Electrical and Information Engineering, Hunan University, Changsha 410082, China (e-mail: tangqiu@hnu.edu.cn; qiuwei@hnu.edu.cn).

Y. Zhou is with the Department of Computer and Information Science, University of Macau, Macau 999078, China (e-mail: yicongzhou@um.edu.mo).

turbances based on four features [4]. In addition, an improved weighted bidirectional Extreme Learning Machine was used to detect multiple disturbances [15]. However, this method has insufficient noise resistance due to limited feature information.

Recently, the Support Vector Machine (SVM) based methods were proposed to provide a reliable solution for a large number of features. A dual multiclass SVM with 58 binary models was proposed to classify 14 types of PQD [16]. It is easy to find that the number of binary models will increase exponentially as the category of disturbances increases. In [10], a SVM method was used to classify the global disturbance ratio index. Nevertheless, the classification accuracy is easily affected by features because only two types of features were used. On the other hand, other SVM methods were also used in PQD detection due to its rigorous theoretical fundamental, including rank wavelet SVM and directed acyclic graph SVM [6], [17]. But their performance of SVM is constrained by a single kernel function that has limited data mapping capabilities.

Combining time-frequency analysis and classifier, different classification frameworks were designed to identify PQDs, such as the framework based on DWT-Fast Fourier transform [18] and hyperbolic ST with DT [19]. It is worth noting that most of their features are statistical features that can guarantee the calculation speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., irrelevant features. In [21], the energy distribution feature was employed as input to the classifier. It ultimately leads to an unsatisfactory classification result. To address this, the feature screening methods were used to select an appropriate feature set based on the optimization strategies [22]–[24]. However, the complexity increases due to a large number of features being designed and selected. A simple yet efficient feature selection method helps to simplify the classification model.

This paper aims to increase the accuracy under multiple complex disturbances. Our contributions are listed as follows:

1) To improve the resolution of time-frequency analysis, an optimized ST (OST) is proposed to decompose PQD signals. Integrating frequency segmentation with maximum energy concentration, OST obtains high frequency resolution at high frequencies while maintaining high time resolution at low frequencies.
2) To enhance the ability of classification, a composite kernel SVM is proposed to form a new kernel SVM (KSVM) for automatically classifing multiple features. Particularly, KSVM improves discriminative information that benefits to classify multi-source features.
3) A multiple PQD classification framework is also proposed based on OST and KSVM. Instead of designing features manually, multi-source feature information is used to enhance classification performance under multiple disturbances.
4) Extensive experiments have been carried out to verify the proposed framework. In addition to testing our framework on single and multiple PQD signals, we are the first to classify the nonlinearly mixed PQD signals. The evaluation results show that our framework has superior performance compared with several state-of-the-arts.

The remaining part of this paper is organized as follows. Section II introduces the proposed OST. In Section III, KSVM is presented to classify the signals based on multi-source information. The framework OST-KSVM of the PQD classification is proposed in Section IV. Then, experiments are conducted in Section V. Finally, the entire paper is concluded in Section VI.

II. Optimized S-Transform

A. Motivation

The time and frequency information is important for accurate detection of multiple PQD signals. However, the resolution of the traditional S-Transform is difficult to meet the requirements of the classifier, especially for mixed disturbances. An S-Transform with two resolutions was selected as the time-frequency analysis of PQD signals [17]. The ST with two resolutions can be described as

$$\mathrm{ST}(\tau, f) = \int_{-\infty}^{\infty} x(t) \frac{\sqrt{\lambda_{1,2}|f|}}{\sqrt{2\pi}} e^{-\frac{(\tau-t)^2 \lambda_{1,2}|f|}{2}} e^{-j2\pi f t} dt \quad (1)$$

where $x(t)$ denotes the PQD signals, f denotes the signal frequency. The Gaussian window function of S-Transform is $g(f) = \frac{\sqrt{\lambda_{1,2}|f|}}{\sqrt{2\pi}} e^{-\frac{(\tau-t)^2 \lambda_{1,2}|f|}{2}}$, and $\lambda_{1,2} = \{\lambda_1, \lambda_2\}$ denotes the window parameters. To obtain better time resolution in low frequency part and better frequency resolution in high frequency part, the frequency intervals are set to $f \leqslant 1.5f_0,\ \lambda_1 > |f|$ and $f > 1.5f_0,\ 0 < \lambda_2 < |f|$, where f_0 is the fundamental frequency, λ_1 and λ_2 are used to control the resolution of low frequency and high frequency bands, respectively. The boundary frequency is determined by the intermediate value of the second harmonic frequency and the fundamental frequency. Therefore, the standard deviation $\sigma(f)$ can be obtained as

$$\sigma(f) = \frac{1}{\sqrt{\lambda_{1,2}|f|}} \quad (2)$$

To achieve high resolutions in both high and low frequencies, ST divides the frequency component into two parts using the boundary $1.5f_0$ and fixed values of $\lambda_{1,2}$. The double resolution works perfectly for single PQD signals. However, it is quite difficult to achieve the satisfactory resolution for mixed PQDs using fixed $\lambda_{1,2}$ values. To address this, an Optimized S-Transform (OST) is proposed to dynamically adjust parameter $\lambda_{1,2}$.

B. Proposed OST

Here, an OST using the energy concentration is proposed to dynamically adjust parameter $\lambda_{1,2}$. Our OST has the same definition as shown in Equation (1). The sampling rate is set to f_s, $\tau = m/f_s$ and $f = nf_s/N$. When $n > 0$, the discrete OST of $x(n)$ can be expressed as

$$\mathrm{OST}(m, n) = \sum_{r=0}^{N-1} X(n+r) G(n) e^{\frac{2\pi r m j}{N}} \quad (3)$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIE.2019.2952823, IEEE Transactions on Industrial Electronics

where $X(n+r)$ and $G(n)$ are the Fourier transform results of discrete $x(n)$ and Gaussian window $g(f)$, respectively. Motivated by [25], the energy concentration of the discrete OST can be calculated by

$$E_{\text{OST}}(\lambda_1,\lambda_2)=\frac{1}{\sum\limits_{m=1}^{M}\sum\limits_{n=1}^{N}\left|\frac{\text{OST}(m,n)}{\sqrt{\sum\sum|\text{OST}(m,n)|^2}}\right|} \tag{4}$$

where matrix $\text{OST}(m,n)$ is of size $M\times N$.

The goal of optimization analysis is to maximize energy concentration E_{OST}, namely $\arg\max(E_{\text{OST}}(\lambda_1,\lambda_2))$.

The time resolution of the low frequency corresponding to λ_1 should be sufficiently high to meet the minimum conditions. On the other hand, the time resolution for the high frequency band corresponding to λ_2 should have a maximum limit to satisfy a high frequency resolution. To quickly obtain the proper window parameters, the parameter constraints should be considered. Thus, the time resolution is set to satisfy

$$m_1T_s<\sigma(f)<m_2T_s \tag{5}$$

where $T_s=1/f_s$ is the sampling period, m_1 and m_2 are the numbers of sampling periods to control the time resolutions. Combining Equations (2) and (5), the constraints of λ_1 and λ_2 can be further obtained as

$$\lambda_1<\frac{1}{f_{max}(m_1T_s)^2},\quad \text{and}\quad \lambda_2>\frac{1}{f_{min}(m_2T_s)^2} \tag{6}$$

where f_{min} and f_{max} are the maximum and minimum frequencies in different frequency bands. Finally, the optimization problem can be summarized as follows

$$\begin{aligned} &\arg\min_{\lambda_1,\lambda_2}\{E_{\text{OST}}(\lambda_1,\lambda_2)\}\\ \text{s.t.}\quad &\lambda_1\in(1.5f_0,\ 1/f_{max}(m_1T_s)^2)\\ &\lambda_2\in[1/f_{min}(m_2T_s)^2,\ 1.5f_0)\end{aligned} \tag{7}$$

It can be seen that Equation (7) is a nonlinear optimization problem. An interior point method is used to solve this problem [26]. In this work, the sampling frequency $f_s=3200$ Hz, $m_1=10$ and $m_2T_s=0.05$. Thus, λ_1 and λ_2 can be set to $\lambda_1\in(75,1365)$ and $\lambda_2\in[5.33,75)$. To speed up the optimization process, the parameter space narrows down to $\lambda_1\in(75,200)$.

In fact, OST is a complex-valued matrix, and can be expressed as

$$\text{OST}(m,n)=|\text{OST}(m,n)|e^{j\phi(m,n)} \tag{8}$$

where $|\text{OST}(m,n)|$ and $\phi(m,n)$ are the amplitude and phase angle of $\text{OST}(m,n)$, respectively. For simplicity, the $\text{OST}_A(m,n)=|\text{OST}(m,n)|$ is set as the amplitude of the OST in time-frequency analysis.

C. Signal Analysis Using OST

To verify the effectiveness of the proposed OST, OST is compared with different algorithms including traditional ST and DRST [17], [27], for analyzing complex signals.

Fig. 1 illustrates the comparison results of different time-frequency analysis methods. Three types of PQD signals occur simultaneously within 0.2 ms. It is worthy to notice that only key frequency points are selected to reduce the calculation complexity of DRST and OST. Fig. 1(a) shows that the traditional ST is insensitive to sag and spike signals. The time interval of DRST is wider than that of OST. This means that OST has a higher time resolution. In addition, the spike signal of OST in Fig. 1(d) shows more details. This indicates that OST has a better energy concentration than the traditional ST and DRST.

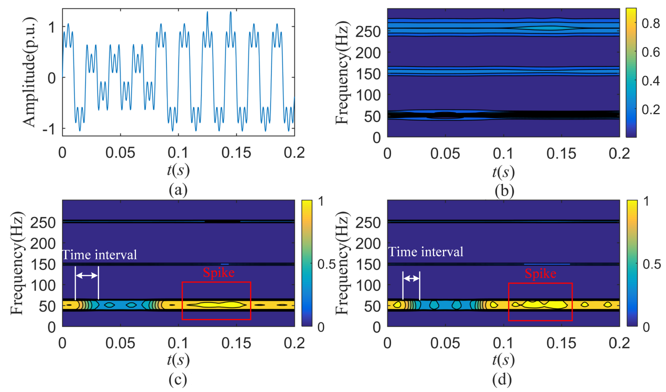

Fig. 1. Comparison results on voltage sag with harmonics and spikes. The start and end time of the sag is 0.02 s and 0.08 s. The magnitude of the sag is 0.4 p.u.. The 2nd and 3rd harmonics are used, and the amplitudes of the 2nd and 3rd harmonics are 0.23 p.u. and 0.28 p.u. respectively. The amplitude range of the spike is 1.20 p.u.. (a) Input signal. (b) Traditional ST. $E_{\text{ST}}=100.35$. (c) DRST. $E_{\text{DRST}}=54.62$. (d) Proposed OST. The optimized parameters are $\lambda_1=199, \lambda_2=5.3$, and $E_{\text{OST}}=53.73$.

The effect of different methods are further tested on time-varying and transient signals. The time-frequency results of the voltage flicker with transient and harmonics are depicted in Fig. 2. The traditional ST cannot detect the flicker changes and has a poor transient detection capability. Both transient and flicker signals can be detected using DRST and OST. However, in the OST detection results, the flicker location contains more detail information, and the reaction of flicker on harmonics is clear. Specifically, the energy of OST is more concentrated than that of DRST. The optimal values of λ_1 and λ_2 are 193 and 5.3, respectively.

After extracting the time-frequency information of PQD signals, a classifier is needed to further detect and identify multiple complex disturbances. Next, a new kernel SVM will be introduced as a classifier for PQD classification.

III. Kernel SVM

A. Proposed Kernel SVM

Based on the extracted time-frequency information, a kernel SVM (KSVM) classification method is proposed for PQD classification. Different from traditional methods, it is unnecessary to manually design statistical features for different types of disturbances.

For KSVM, given a set of PQD feature samples $D=\{x_i,y_i\}_n$, where i =1,2,...n, x_i is the element of $x(n)$, y_i is the label of x_i. Traditionally, only one kernel function is used to map the feature data D to a high-dimensional space.

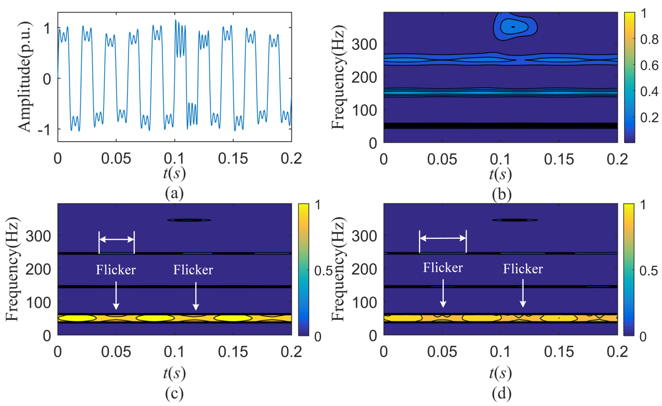

Fig. 2. Comparison results on voltage flicker with transient and harmonics. The 2nd and 3rd harmonics are used, and the amplitudes of the 2nd and 3rd harmonics are 0.30 p.u. and 0.21 p.u. respectively. The amplitude of the flicker is 1.1 p.u.. The frequency of the flicker is 15 Hz. The transient occurs from is 0.1 to 0.13 s. (a) Input signal. (b) Traditional ST. $E_{\text{ST}} = 93.36$. (c) DRST. $E_{\text{DRST}} = 56.14$. (d) Proposed OST. The optimized parameters are $\lambda_1 = 193, \lambda_2 = 5.3$, and $E_{\text{OST}} = 55.56$.

The Gaussian kernel, one of the most commonly used kernel functions, is used to convert data in SVM [28]. For the two points $x_i, x_j \in D$, the Gaussian kernel is deifned as

$$K(x_i, x_j) = exp(-\frac{\|x_i - x_j\|^2}{2\delta^2}) \tag{9}$$

where $\delta > 0$ denotes the width parameter, and δ controlls the mapping results.

For PQD classification, more information contributes to identify signals, especially for the mixed disturbances consisting of three single components. Thus, the raw data and the time-frequency data are combined as the input of the classifier. To reduce the time of feature design and model computation, the maximum values of the time and frequency axes of the time-frequency matrix $\text{OST}(m, n)$ are used as input features of KSVM. The F and T are set as the maximum time and frequency axes of $\text{OST}_A(m, n)$ respectively, i.e., $F = \max\{\text{OST}_A(m, n)_{row}\}$ and $T = \max\{\text{OST}_A(m, n)_{column}\}$.

To deal with these features, a weighted linear combination kernel is proposed based on the F, T and raw PQD signals $x(n)$

$$\begin{gathered} K_l(x_i, x_j) = u_1 K_f(f_i, f_j) + u_2 K_t(t_i, t_j) + u_3 K_x(x_{ti}, x_{tj}) \\ s.t. \quad u_1 + u_2 + u_3 = 1 \end{gathered} \tag{10}$$

where K_f, K_t and K_x denote the kernel functions of F, T and $x(n)$ respectively, u_1, u_2 and u_3 are the weight coefficients.

To learn the decision plane for classification, the optimization framework of the proposed KSVM is introduced as follows

$$\begin{gathered} \min_{\boldsymbol{w}, b, \xi_i} \left\{ \frac{\|\boldsymbol{w}\|^2}{2} + C \sum_{i=1}^{n} \xi_i^2 \right\} \\ s.t.\ y_i(\boldsymbol{w}^T K_l(x_i, x_j) + b) \geq 1 - \xi_i \\ \xi_i \geq 0, i = 1, ..., n \end{gathered} \tag{11}$$

TABLE I
24 TYPES OF PQD SIGNALS

Class	PQ diaturbance	Class	PQ diaturbance
C1	Normal	C13	Interrupt + harmonics
C2	Swell	C14	Swell + transient
C3	Sag	C15	Sag + transient
C4	Interruption	C16	Spike + transient
C5	Harmonics	C17	Transient + harmonics + sag
C6	Swell + harmonics	C18	Transient + harmonics + swell
C7	Sag + harmonics	C19	Transient + harmonics + interrupt
C8	Transient	C20	Transient + harmonics + flicker
C9	Flicker	C21	Flicker + harmonics + interrupt
C10	Flicker + harmonics	C22	Flicker + harmonics + sag
C11	Notch	C23	Flicker + harmonics + swell
C12	Spike	C24	Spike + transient + swell

where $\boldsymbol{w}$ and b denote the weight vector and bias term of the decision plane respectively, ξ_i is the slack variable, C is the penalty coefficient which represents the tolerance of model error. Introducing dual Lagrange function, the parameters of the proposed KSVM model can be obtained via partial derivatives.

To detect the category of new PQD data z, the decision function is calculated as

$$\hat{y} = \text{sign}(\boldsymbol{w}^T K_l(x_i, z) + b) \tag{12}$$

B. Analysis of Kernel Functions and Parameters

To verify the performance of KSVM, twenty four types of PQD signals are used. As listed in Table I, they contain nine single disturbances and fifteen mixed disturbances. All disturbances are generated according to [29] and IEEE standard 1159 [30], and known to be very close to the real data set. The numerical model of multiple PQDs in Table I can be described as

$$x(t) = [V_{normal}(t) + V_{add}(t)]V_{multiply}(t) \tag{13}$$

where $V_{normal}(t)$ is the normal signal, $V_{add}(t)$ denotes the additive disturbance type, such as harmonics, transient, notch and noise; $V_{multiply}(t)$ denotes the disturbance components, including swell, sag, interrupt and flicker.

The parameters under a single disturbance are set as follows. The depth of the sag ranging from 0.1 p.u. to 0.9 p.u., the depth of the swell ranging from 1.1 p.u. to 1.8 p.u.. The depth range of interruption is set to 0 p.u. to 0.1 p.u.. The duration of the sag, swell and interruption is set to 1 to 9 times the fundamental period. The amplitude of the harmonic is set to 0.05 p.u. to 0.3 p.u., and the harmonic order mainly includes the 3rd, 5th, and 7th harmonics. The total harmonics distortion is less than 5%. The duration of the transient ranging from 0.5 to 3 times of the fundamental period. The amplitude range of the spike is 1.1 p.u. to 1.4 p.u., where the amplitude of the notch is opposite to the spike. The durations of the spike and notch are 0.01 to 0.05 times of the fundamental frequency period. The amplitude range of the flicker is 1.1 p.u. to 1.2 p.u.. The frequency of the flicker component varies from 5 to 25 Hz. The mixed PQD signals are randomly generated based on single perturbed parameters.

TABLE II
PERFORMANCE UNDER DIFFERENT COMBINATIONS OF KERNEL FUNCTIONS

Kernel functions	Weight coefficients			Accuracy (%)
	u_1	u_2	u_3	
Linear + Polynomial	0.5	0.5	0	79.60
Linear + Gaussian	0.5	0.5	0	93.06
Gaussian + Linear	0.5	0.5	0	96.94
Gaussian + Polynomial	0.5	0.5	0	84.58
Linear kernel	0.5	0.3	0.2	95.64
Polynomial kernel	0.5	0.3	0.2	90.44
Gaussian kernel	0.5	0.3	0.2	97.31
Gaussian kernel	0.8	0.1	0.1	98.12
Gaussian kernel	0.9	0.05	0.05	98.82

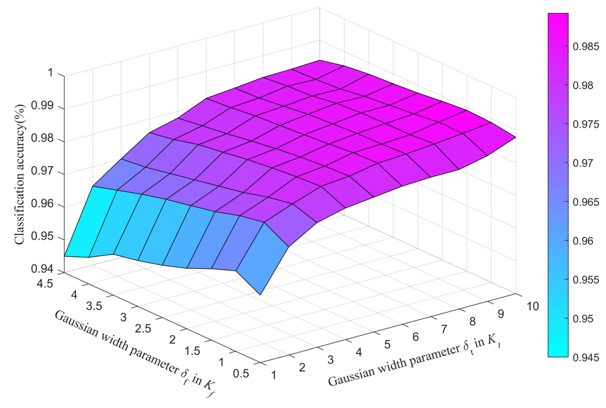

Fig. 3. Relationship between different kernel parameters and classification accuracy.

The kernel function directly affects the performance of KSVM because different kernel functions have different data mapping effects. Therefore, it is important to choose a suitable kernel function.

To select a proper kernel function, three commonly used kernel functions are selected to validate the model, including linear kernel, polynomial kernel and Gaussian kernel. For a fair comparison, the coefficients δ of the kernel function are set to be the same. The performance under different combinations of kernel functions is listed in Table II.

It is observed from the results in Table II that the Gaussian and linear kernels outperform the polynomial kernel. In addition, the linear kernel does not perform well as Gaussian kernel. For example, the accuracy rates are 95.64% and 97.31% when all kernels are set as linear kernel and Gaussian kernel, respectively. Additionally, the weight coefficient frequency domain component u_1 should have a higher value.

For simplicity, all kernel functions of the combination kernel are set to Gaussian kernel in the rest of this paper. As for the coefficients of the kernel functions, the selection of parameters can be divided into two steps: they are first empirically determined in a suitable range of values, and then the optimal parameters are selected in conjunction with the grid search method. In the first step, the parameter values are searched by a fixed step size based on the model performance. For example, the 0.5 and 0.8 are first specified to the parameter u_1, and the result shows that KSVM performs the best when u_1 = 0.8. Then the scope of u_1 can be set to [0.5, 0.99]. The grid search is used to search for the best u_1 value when the step size is set to 0.01. Finally, the coefficients of the kernel functions are set to $u_1 = 0.9$, $u_2 = 0.05$ and $u_3 = 0.05$ respectively.

After the kernel function and the corresponding coefficient are specified, the parameters of the Gaussian kernel need to be further explored. The performance of KSVM under different kernel parameters is shown in Fig. 3 using the grid search method. According to the experiment results, the parameter of Gaussian kernel K_x is set to a constant 3.2. The penalty coefficient C is set to 2000.

As can be seen, the accuracy increases with the increase of the parameter δ_t when the parameter value ranges from 1 to 10. On the contrary, the parameter δ_f is negatively correlated with classification accuracy when δ_f is higher than 1. The highest accuracy is obtained at $\delta_t = 10$ and $\delta_f = 1.5$. Based on Fig. 3, the parameters are further fine-tuned, $\delta_f = 1.3$ and $\delta_t = 10.7$ are then selected for the kernel function.

IV. PQD CLASSIFICATION FRAMEWORK

Using the proposed OST and KSVM, this section proposes a framework called OST-KSVM for PQD classification. Its flowchart is shown in Fig. 4. The framework OST-KSVM can be divided into two parts:

1) Time-frequency analysis: Calculation of time frequency amplitude matrix $\text{OST}_A(m,n)$ for multiple disturbances via maximizing energy concentration. The features F, T are obtained from $\text{OST}_A(m,n)$.
2) Automatic classification of PQD signals: Multiple sources of information, including F, T and $x(t)$, are obtained to compute the weighted linear combination kernel. Then the PQD signals can be classified by KSVM.

Thereby, the proposed framework OST-KSVM is further evaluated by different experiments.

V. EXPERIMENTS AND EVALUATIONS

To verify the effectiveness of the proposed OST-KSVM, experiments are conducted under noisy and noise-free conditions. In our experiments, 2000 samples per disturbance are generated in MATLAB. The cross-validation is used to verify the effect of the model [31], with 800 groups being used for training, 600 samples for testing and 600 for verification. The fundamental frequency f_0 is 50 Hz, and the window size is ten times of the fundamental period, namely 640 points for each signal sample.

A. Performance under different noise levels

The features extracted by time-frequency analysis are directly related to the validity of classification. In order to verify the effect of the proposed OST-KSVM under different noise levels, various noise conditions are verified, as shown in Table III.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIE.2019.2952823, IEEE Transactions on Industrial Electronics

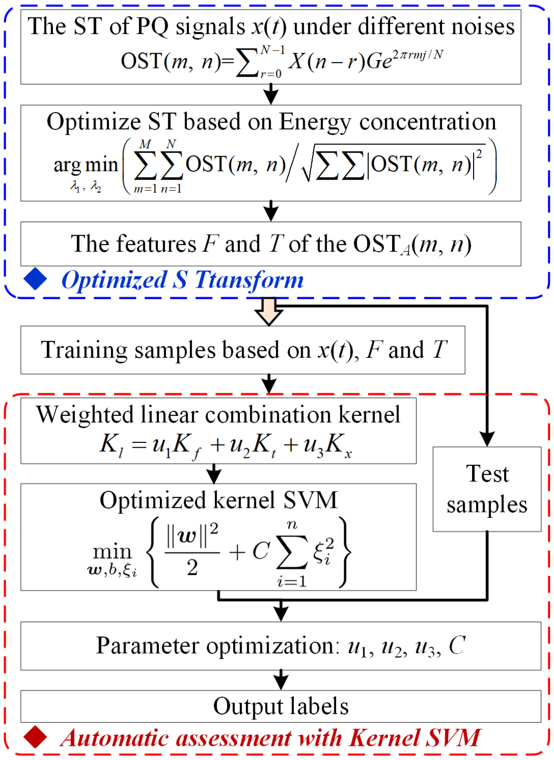

Fig. 4. Automatic classification framework of PQDs based on OST-KSVM

TABLE III
PERFORMANCE UNDER DIFFERENT NOISE LEVELS

Kernel functions	Accuracy (%)				Test time per sample (ms)
	Clean	10 dB	20 dB	40 dB	
ST-KSVM	94.72	87.32	97.82	98.97	27
OST-KSVM	99.59	88.22	98.82	99.51	88

From Table III, the accuracy rises with the noise level decreases for KSVM with traditional ST and OST. The performance of KSVM with ST is not higher than 99% under all the noise levels. However, the accuracy of OST-KSVM exceeds 99% when the noise level is over 30 dB. The test time of OST-KSVM is higher than traditional KSVM with ST due to the iterative optimization process. However, the sample period is ten times of the fundamental period. It means that the detection time must be lower than 200 ms per sample to meet real-time requirements. Table III shows that the proposed method still meets the real-time requirements. Meanwhile, its detection accuracy is higher than other methods.

The performance of each PQD under different noise conditions is listed in Table IV. Obviously, the classification accuracy significantly increases at a low noise level. Concretely, the accuracy of single disturbance is higher than that of mixed disturbances, indicating that a single signal is easy to classify. Meanwhile, for mixed disturbances, especially for three mixed disturbances, the minimum accuracy of the proposed OST-KSVM is 96.33% even with noise level of 20 dB. The proposed OST-KSVM has good performance for mixed disturbances.

TABLE IV
DETAILED PERFORMANCE OF OST-KSVM

Class	PQD	Accuracy (%)		
		Clean	40dB	20dB
C1	Normal	100	100	99.83
C2	Swell	100	99.83	100
C3	Sag	99.00	97.17	95.67
C4	Interrupt	99.17	98.50	100
C5	Harmonics	100	100	100
C6	Swell + harmonics	99.33	99.83	100
C7	Sag + harmonics	100	100	98.25
C8	Transient	100	100	99.17
C9	Flicker	100	100	100
C10	Flicker + harmonics	100	100	98.67
C11	Notch	100	100	100
C12	Spike	100	100	97.83
C13	Interrupt + harmonics	100	98.83	99.67
C14	Swell + transient	99.00	100	98.67
C15	Sag + transient	99.33	99.00	98.00
C16	Spike + transient	99.33	100	97.17
C17	Transient + harmonics + sag	99.67	99.00	96.83
C18	Transient + harmonics + swell	99.00	99.00	99.00
C19	Transient + harmonics + interrupt	96.67	98.33	96.33
C20	Transient + harmonics + flicker	100	100	97.17
C21	Flicker + harmonics + interrupt	100	99.50	99.50
C22	Flicker + harmonics + sag	99.67	99.67	98.33
C23	Notch + transient + swell	100	100	99.50
C24	Spike + transient + swell	100	100	100

B. Performance under multiple proportions of training data

To verify performance under different proportions of training data, the proposed method is compared with different ST and SVM methods. Meanwhile, as a high spectral resolution method, the MUSIC method is used to replace the OST method and combined with kernel SVM (namely MUSIC-KSVM) [32]. MUSIC-KSVM is then compared with the proposed method. The experiment results are listed in Table V.

In MUSIC-KSVM, the inputs of KSVM are the raw PQD signals and power spectrum signals. Two corresponding Gaussian kernel functions are used. The weight coefficients of the raw PQD signal and spectral signal are set to 0.4 and 0.6 in the optimized MUSIC-KSVM$_1$ respectively. Table V shows that the proposed method outperforms MUSIC-KSVM, indicating that the information provided by the MUSIC method is insufficient to achieve higher accuracy. It also shows that the accuracy of different methods increases as the training data increases. It is worth noting that the detection accuracy of the proposed method is still above 97% even if the training data is limited to 10%. Table V shows that the proposed method is more robust and adaptable by comparison with Quad-SVM.

C. Classification of Nonlinearly Mixed PQDs

For the aforementioned experiments, components of mixed disturbances are linearly combined together. This, however, ignores the complexity of nonlinear loads in the real power systems. To further verify the effectiveness of the proposed

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIE.2019.2952823, IEEE Transactions on Industrial Electronics

TABLE V
TEST ACCURACY UNDER DIFFERENT TRAINING DATA IN 20 DB NOISE

Method	Training data (%)				
	10%	20%	30%	40%	50%
MUSIC-KSVM$_1$	83.58	86.66	88.08	89.59	90.09
MOFDST and Quad-SVM [33]	–	92.02	–	–	94.15
RST and Quad-SVM [34]	88.03	93.31	95.05	96.02	96.38
OST-KSVM	97.33	98.23	98.66	98.82	98.96

KSVM$_1$: the kernel parameters are optimized, –: it is not reported.

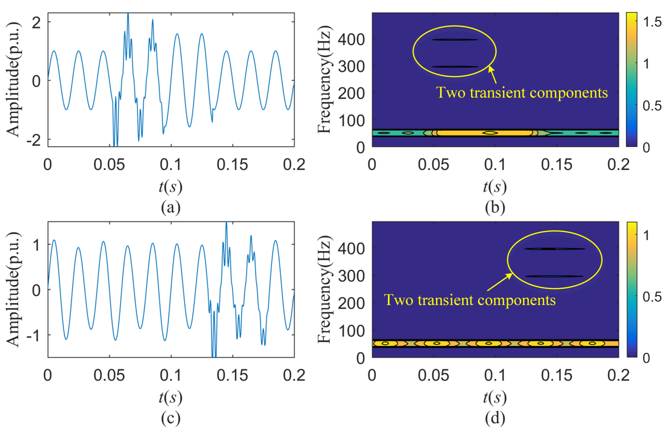

Fig. 5. Results of nolinearly mixed disturbances. (a) Voltage transient with swell. The start and end time of the swell is 0.05 s and 0.135 s. The magnitude of the swell is 1.70 p.u.. The transient occurs from is 0.05 to 0.085 s. (b) OST of (a). (c) Voltage flicker with transient. The amplitude and frequency of the flicker are 1.17 p.u. and 22.5 Hz respectively. The transient occurs from is 0.13 to 0.18 s. (d) OST of (c).

method, the nonlinearly mixed PQD signals are explored and simulated for the first time.

Two kinds of nonlinearly mixed PQD signals are considered, including transient with swell and flicker with transient, as shown in Fig. 5. All the single PQD components are multiplied to simulate nonlinear changes. It can be seen that two transient components are detected. There is no significant difference in the fundamental components.

The classification results are listed in Table VI. It is observed that the classification accuracy of OST-KSVM is higher than 97.5%. This means that the proposed OST-KSVM has excellent performance for classifying nonlinearly mixed PQD signals.

D. Performance Comparison

To further verify the performance of the proposed OST-KSVM, OST-KSVM is compared with several recently proposed techniques as listed in Table VII. It can be seen that different methods have different numbers of features. For example, the ST and NSGA-II method [22] has 26 kinds of features and obtains a higher accuracy rate. On the contrary, both the WT and SVM method [10] and the ADLINE and FFNN method [3] have only 2 features and they have lower accuracy rates. In the noise level of 30 dB, the FDST and DT method [13] contains a larger number of features and thus performs better than the VMD and DT methods [4]. This means that more effective features are beneficial for classification. The loss of information can be effectively avoided by automatic feature extraction. In addition, the methods in [10], [22] and [16] fail to consider mixed PQDs especially for mixed PQDs consisting of three single disturbances. In this case, the proposed OST-KSVM achieves an accuracy of 98.82% even with a noise level of 20 dB.

TABLE VI
PERFORMANCE OF NONLINEARLY MIXED PQDS

Nonlinearly mixed disturbances	Accuracy(%)		
	Clean	20 dB	40 dB
Transient with swell	98.67	97.50	98.42
Flicker with transient	100	99.00	99.60

TABLE VII
PERFORMANCE COMPARISON WITH OTHER METHODS

Method	Num. of PQD	Num. of Features	Noise (dB)	Accuracy (%)
DRST and DAG-SVM [17]	9	9	20	97.77
TQWT and MSVM [16]	14	5	20	96.42
WT and SVM [10]	9	2	20	94.22
VMD and DT [4]	14	4	30	96.73
HT and slip-SVDNSA [35]	11	8	20	98.45
ADLINE and FFNN [3]	12	2	20	90.58
HST and DT [19]	13	13	20	96.10
FDST and DT [13]	13	20	30	97.44
DWT and PNN [36]	16	9	20	93.60
ST and NSGA-II [22]	15	26	20	96.43
OST-KSVM	**24**	**Automatic**	**20**	**98.82**

E. Experimental Verification Analysis

Different from simulation signals, experimental signals have high randomness and heterogeneity. To verify the performance of the proposed OST-KSVM under experimental signals, a sampling hardware circuit with a real-time acquisition function is designed to sample the signals. The hardware platform is shown in Fig. 6.

As can be seen from Fig. 6, the PQD signals are generated by disturbance signal source Fluke 6105A. After being preheated for an hour, the instrument randomly generates different signals includes Normal (C1), Swell (C2), Sag (C3), Interrupt (C4), Harmonics (C5), Swell with harmonics (C6), Sag with harmonics (C7) and Flicker (C9) due to the functional limitations of the instrument. Data sampling platform is composed of 16 bit ADS 8556 (ADC) and 32 bit float point TMS320VC6748 (DSP). The clock frequency of DSP is set to 375 MHz. The sampling frequency of ADC is set to 5 kHz. The required frequency is obtained by downsampling. The signals are collected after passing through the voltage transformer. After sampling the signals, DSP transmits the signal data to the computer in real time via the serial interface. Finally, the signals can be analyzed by the proposed OST-KSVM. The labels of the disturbance signals are determined

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIE.2019.2952823, IEEE Transactions on Industrial Electronics

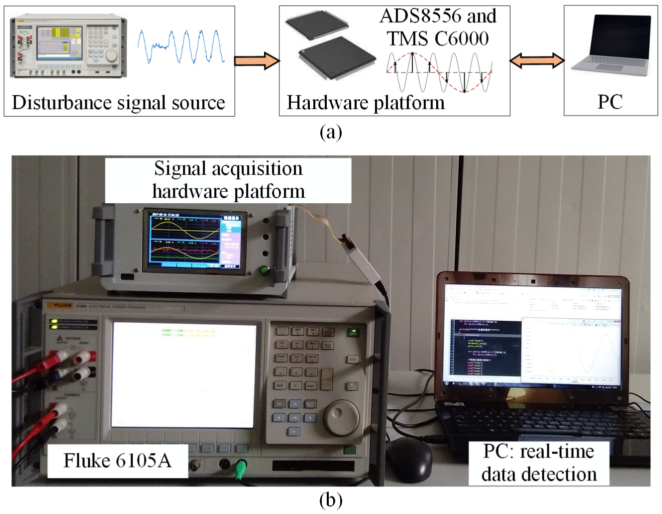

Fig. 6. Power quality sampling hardware platform. (a) Schematic framework of power quality signal sampling, (b) Hardware platform.

TABLE VIII
PERFORMANCE UNDER THE EXPERIMENTAL SIGNAL

Class	Accuracy (%)	Class	Accuracy (%)	Average accuracy (%)	Test time per sample (ms)
C1	100	C5	98.33	97.08	92
C2	96.67	C6	95.00		
C3	98.33	C7	95.00		
C4	96.67	C9	96.67		

by the set parameters, and the accuracy can be calculated by comparing the predicted results with the labels corresponding to the set parameters.

For each type of disturbance signals, 60 samples are generated. The experimental classification results are listed in Table VIII. As can be observed, the average accuracy of the experimental signals is lower than that of the simulated signals. This means that the noise in hardware circuits reduces the accuracy of the experiments, and there is precision loss in the ADC sampling process. The calculation time is less than 200 ms, indicating that it can meet real-time requirements. Hence, the proposed OST-KSVM has satisfactory performance under the experimental signals.

VI. CONCLUSION

In this paper, an optimized S-transform and kernel SVM were proposed to automatic detection and assessment of the single and mixed PQD signals. The time and frequency resolutions have been improved by maximizing the energy concentration in OST. The experimental results showed that OST has higher time resolution at low frequencies and better frequency resolution in high frequency intervals. Thereafter, raw and time-frequency features are integrated and automatically learned by the proposed KSVM. Simulation results of different kernel functions showed that the linear combination of kernels has a stronger separability than a single kernel. Various simulations and experiments were conducted to verify the proposed framework OST-KSVM, and the results showed that OST-KSVM has stronger noise immunity and better performance than several existing methods. In particular, several nonlinearly mixed disturbances were tested and verified for the first time, and OST-KSVM shows satisfactory results. Finally, the performance of the proposed OST-KSVM was tested and verified by the data collected from the experimental platform.

REFERENCES

[1] Z. Oubrahim, V. Choqueuse, Y. Amirat, and M. E. H. Benbouzid, "Disturbances Classification Based on a Model Order Selection Method for Power Quality Monitoring," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 12, pp. 9421–9432, 2017.

[2] Y. Luo, K. Li, Y. Li, D. Cai, C. Zhao, and Q. Meng, "Three-Layer Bayesian Network for Classification of Complex Power Quality Disturbances," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 9, pp. 3997–4006, 2018.

[3] M. Valtierra-Rodriguez, R. de Jesus Romero-Troncoso, R. A. Osornio-Rios, and A. Garcia-Perez, "Detection and Classification of Single and Combined Power Quality Disturbances Using Neural Networks," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 5, pp. 2473–2482, 2014.

[4] P. D. Achlerkar, S. R. Samantaray, and M. S. Manikandan, "Variational Mode Decomposition and Decision Tree Based Detection and Classification of Power Quality Disturbances in Grid-Connected Distributed Generation System," *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 3122–3132, 2018.

[5] S. Cho, G. Jang, and S. Kwon, "Time-Frequency Analysis of Power Quality Disturbances via the Gabor-Wigner Transform," *IEEE Transactions on Power Delivery*, vol. 25, no. 1, pp. 494–499, Jan. 2010.

[6] Z. Liu, Y. Cui, and W. Li, "A Classification Method for Complex Power Quality Disturbances Using EEMD and Rank Wavelet SVM," *IEEE Transactions on Smart Grid*, vol. 6, no. 4, pp. 1678–1685, 2015.

[7] K. Satpathi, Y. M. Yeap, A. Ukil, and N. Geddada, "Short-Time Fourier Transform Based Transient Analysis of VSC Interfaced Point-to-Point DC System," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 5, pp. 4080–4091, 2018.

[8] T. A. Kawady, N. I. Elkalashy, A. E. Ibrahim, and A.-M. I. Taalab, "Arcing fault identification using combined gabor transform-neural network for transmission lines," *International Journal of Electrical Power & Energy Systems*, vol. 61, pp. 248 – 258, 2014.

[9] R. S. H., S. R. Mohanty, N. Kishor, and A. T. K., *IEEE Transactions on Industrial Electronics*.

[10] M. D. Borras, J. C. Bravo, and J. C. Montano, "Disturbance Ratio for Optimal Multi-Event Classification in Power Distribution Networks," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 5, pp. 3117–3124, 2016.

[11] J. Ma and J. Jiang, "Analysis and design of modified window shapes for S-transform to improve time frequency localization," *Mechanical Systems and Signal Processing*, vol. 58-59, pp. 271–284, 2015.

[12] N. Singh and P. M. Pradhan, "Efficient Discrete S-Transform Based on Optimally Concentrated Window," *IEEE Signal Processing Letters*, vol. 26, no. 1, pp. 14–18, 2019.

[13] M. Biswal and P. Dash, "Detection and characterization of multiple power quality disturbances with a fast S-transform and decision tree based classifier," *Digital Signal Processing*, vol. 23, no. 4, pp. 1071 – 1083, 2013.

[14] R. Kumar, B. Singh, D. T. Shahani, A. Chandra, and K. Al-Haddad, "Recognition of Power-Quality Disturbances Using S-Transform-Based ANN Classifier and Rule-Based Decision Tree," *IEEE Transactions on Industry Applications*, vol. 51, no. 2, pp. 1249–1258, Mar. 2015.

[15] M. Sahani and P. K. Dash, "Automatic Power Quality Events Recognition Based on Hilbert Huang Transform and Weighted Bidirectional Extreme Learning Machine," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 9, pp. 3849–3858, 2018.

[16] K. Thirumala, M. S. Prasad, T. Jain, and A. C. Umarikar, "Tunable-Q Wavelet Transform and Dual Multiclass SVM for Online Automatic Detection of Power Quality Disturbances," *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 3018–3028, 2018.

[17] J. Li, Z. Teng, Q. Tang, and J. Song, "Detection and Classification of Power Quality Disturbances Using Double Resolution S-Transform and DAG-SVMs," *IEEE Transactions on Instrumentation and Measurement*, vol. 65, no. 10, pp. 2302–2312, 2016.

[18] S. Deokar and L. Waghmare, "Integrated DWT-FFT approach for detection and classification of power quality disturbances," *International Journal of Electrical Power & Energy Systems*, vol. 61, pp. 594 – 605, 2014.

[19] P. K. Ray, S. R. Mohanty, N. Kishor, and J. P. S. CatalÃ£o, "Optimal Feature and Decision Tree-Based Classification of Power Quality Disturbances in Distributed Generation Systems," *IEEE Transactions on Sustainable Energy*, vol. 5, no. 1, pp. 200–208, 2014.

[20] S. Wang and H. Chen, "A novel deep learning method for the classification of power quality disturbances using deep convolutional neural network," *Applied Energy*, vol. 235, pp. 1126 – 1140, 2019.

[21] C. Zhao, K. Li, Y. Li, L. Wang, Y. Luo, X. Xu, X. Ding, and Q. Meng, "Novel Method Based on Variational Mode Decomposition and a Random Discriminative Projection Extreme Learning Machine for Multiple Power Quality Disturbance Recognition," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 5, pp. 2915–2926, May. 2019.

[22] U. Singh and S. N. Singh, "Optimal Feature Selection via NSGA-II for Power Quality Disturbances Classification," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 7, pp. 2994–3002, 2018.

[23] M. Markovska and D. Taskovski, "The effectiveness of wavelet based features on power quality disturbances classification in noisy environment," in *2018 18th International Conference on Harmonics and Quality of Power (ICHQP)*, pp. 1–6, May. 2018.

[24] C. Lee and Y. Shen, "Optimal Feature Selection for Power-Quality Disturbances Classification," *IEEE Transactions on Power Delivery*, vol. 26, no. 4, pp. 2342–2351, Oct. 2011.

[25] L. Stankovic, "A measure of some time frequency distributions concentration," *Signal Processing*, vol. 81, no. 3, pp. 621 – 631, 2001.

[26] A. Moukadem, Z. Bouguila, D. O. Abdeslam, and A. Dieterlen, "A new optimized Stockwell transform applied on synthetic and real nonstationary signals," *Digital Signal Processing*, vol. 46, pp. 226 – 238, 2015.

[27] R. G. Stockwell, L. Mansinha, and R. P. Lowe, "Localization of the complex spectrum: the S-transform," *IEEE Transactions on Signal Processing*, vol. 44, no. 4, pp. 998–1001, 1996.

[28] B. Pan, W. Chen, C. Xu, and B. Chen, "A Novel Framework for Learning Geometry-Aware Kernels," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 5, pp. 939–951, May. 2016.

[29] R. Tan and V. K. Ramachandaramurthy, "Numerical model framework of power quality events," *European Journal of Scientific Research*, vol. 43, no. 1, pp. 30–47, 06 2010.

[30] "IEEE Power and Energy Society. IEEE Recommended Practice for Monitoring Electric Power Quality," *IEEE Standard 1159-2009*, 2009.

[31] K. Chen, J. Hu, and J. He, "A Framework for Automatically Extracting Overvoltage Features Based on Sparse Autoencoder," *IEEE Transactions on Smart Grid*, vol. 9, no. 2, pp. 594–604, 2018.

[32] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, Mar. 1986.

[33] M. Venkateswara Reddy and R. Sodhi, "Modified S-Transform and Random Forests-Based Power Quality Assessment Framework," *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 1, pp. 78–89, Jan. 2018.

[34] M. V. Reddy and R. Sodhi, "A rule-based S-Transform and AdaBoost based approach for power quality assessment," *Electric Power Systems Research*, vol. 134, pp. 66 – 79, 2016.

[35] Y. Wang, Q. Li, F. Zhou, Y. Zhou, and X. Mu, "A New Method With Hilbert Transform and Slip-SVD-Based Noise-Suppression Algorithm for Noisy Power Quality Monitoring," *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 4, pp. 987–1001, Apr. 2019.

[36] S. Khokhar, A. A. M. Zin, A. P. Memon, and A. S. Mokhtar, "A new optimal feature selection algorithm for classification of power quality disturbances using discrete wavelet transform and probabilistic neural network," *Measurement*, vol. 95, pp. 246 – 259, 2017.

Qiu Tang received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from Hunan University, Changsha, China, in 1992, 1995, and 2010, respectively, and the M.Sc. degree in electrical engineering from the University of Nottingham, Nottingham, U.K., in 2005.

She has been an Assistant Professor with Hunan University since 2006. Her current research interests include power system analysis, digital signal processing, and virtual instruments.

Wei Qiu (S'19) received the B.Sc. degree in electrical engineering from Hubei University of Technology, Wuhan, China, in 2015, and M.Sc. degree in electrical engineering from Hunan University, Changsha, China, in 2017. He is currently pursuing the Ph.D. degree at Hunan University. He is also a joint doctoral student at the University of Tennessee from 2019.

His current research interests include power system analysis, power quality measurement, and reliability analysis of power equipment.

Yicong Zhou (M'07-SM'14) received his B.S. degree from Hunan University, Changsha, China, and his M.S. and Ph.D. degrees from Tufts University, Massachusetts, USA, all in electrical engineering.

He is an Associate Professor and Director of the Vision and Image Processing Laboratory in the Department of Computer and Information Science at University of Macau. His research interests include image processing and understanding, computer vision, machine learning, and multimedia security.

Dr. Zhou is a senior member of the International Society for Optical Engineering (SPIE). He was a recipient of the Third Price of Macau Natural Science Award in 2014. He is a Co-Chair of Technical Committee on Cognitive Computing in the IEEE Systems, Man, and Cybernetics Society. He serves as an Associate Editor for *IEEE Transactions on Circuits and Systems for Video Technology*, *IEEE Transactions on Geoscience and Remote Sensing*, and four other journals.