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 nonlinear loads makes power system signals complex and distorted and, thus, increases the difficulty of detecting and classifying PQD signals. To address this issue, this article first proposes an optimized S-transform (OST). It optimizes different window parameters to improve time-frequency resolution using maximum energy concentration. A kernel SVM (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 support vector machine (SVM), nonlinearly mixed power quality disturbance (PQD), optimized S-transform (OST), time-frequency resolution.

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 nonlinear 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 nonstationary 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 disturbances 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], an 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., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features may lead to the loss of critical information, especially for mixed PQDs, e.g., speed [20]. However, the handicraft features ma
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

$$E_{OST}(<1, >2) = \frac{1}{\sum_{m=1}^{M} \sum_{n=1}^{N} \left|OST(m, n)\right|^2}$$

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

The goal of optimization analysis is to maximize energy concentration $E_{OST}$, namely $\arg \max (E_{OST}(<1, >2))$.

The time resolution of the low frequency corresponding to $\lambda_1$ should be sufficiently high to meet the minimum conditions. On the other hand, the time resolution for the high-frequency band corresponding to $\lambda_2$ 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

$$m_1 T_s < \sigma(f) < m_2 T_s$$

where $T_s = 1/f_s$ is the sampling period, and $m_1$ and $m_2$ are the numbers of sampling periods to control the time resolutions. Combining (2) and (5), the constraints of $\lambda_1$ and $\lambda_2$ can be further obtained as

$$\lambda_1 < \frac{1}{f_{\max}(m_1 T_s)^2}, \quad \lambda_2 > \frac{1}{f_{\min}(m_2 T_s)^2}$$

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:

$$\arg \min_{\lambda_1, \lambda_2} \{E_{OST}(\lambda_1, \lambda_2)\}$$

s.t. $\lambda_1 \in (1.5 f_0, 1/f_{\max}(m_1 T_s)^2)$

$$\lambda_2 \in [1/f_{\min}(m_2 T_s)^2, 1.5 f_0).$$

It can be seen that (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 $m_2 T_s = 0.05$. Thus, $\lambda_1$ and $\lambda_2$ can be set to $\lambda_1 \in (75, 1365)$ and $\lambda_2 \in [5, 35, 75]$. 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

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

where $|OST(m, n)|$ and $\phi(m, n)$ are the amplitude and phase angle of $OST(m, n)$, respectively. For simplicity, the $OST_A(m, n) = |OST(m, n)|$ 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.

The effect of different methods is 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 $\lambda_1$ and $\lambda_2$ 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 SUPPORT VECTOR MACHINE

#### A. Proposed KSVM

Based on the extracted time-frequency information, a 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 $i = 1, 2, \ldots, n$, $x_i$ 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. 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 defined as

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

where $\delta > 0$ denotes the width parameter, and $\delta$ controls 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 $OST(m, n)$ are used as input features of KSVM. The $F$ and $T$ are set as the maximum time and frequency axes of $OST_{A(m, n)}$, respectively, i.e., $F = \max\{OST_{A(m, n)}\}$ and $T = \max\{OST_{A(m, n)}\}$.

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

$$K_l(x_i, x_j) = u_1K_f(f_i, f_j) + u_2K_t(t_i, t_j) + u_3K_x(x_{t1}, x_{tj})$$

s.t. $u_1 + u_2 + u_3 = 1 \quad (10)$

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:

$$\min_{w,b,\xi} \left\{ \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{n} \xi_i \right\}$$

s.t. $y_i(w^T K_l(x_i, x_j) + b) \geq 1 - \xi_i$

$$\xi_i \geq 0, i = 1, \ldots, n \quad (11)$$

where $w$ and $b$ denote the weight vector and bias term of the decision plane, respectively, $\xi_i$ is the slack variable, and $C$ 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

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

#### B. Analysis of Kernel Functions and Parameters

To verify the performance of KSVM, 24 types of PQD signals are used. As listed in Table I, they contain nine single disturbances and 15 mixed disturbances. All disturbances are generated according to [29] and IEEE Standard 1159 [30], and known to be very close to the real dataset. The numerical model of multiple PQDs in Table I can be described as

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

where $V_{\text{normal}}(t)$ is the normal signal, $V_{\text{add}}(t)$ denotes the additive disturbance type, such as harmonics, transient, notch, and noise; $V_{\text{multiply}}(t)$ 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 ranges from 0.1 to 0.9 p.u., and the depth of the swell ranges from 1.1 to 1.8 p.u. The energy ratio of the transient is set to 0 to 0.1 p.u. The duration of the sag, swell, and interruption is set to one to nine times the fundamental period. The amplitude of the harmonic is set to 0.05 to 0.3 p.u., and the harmonic order mainly includes the third, fifth, and seventh harmonics. The total harmonics distortion is less than 5%. The duration of the transient ranges from 0.5 to 3 times of the fundamental period. The amplitude range of the spike is 1.1 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 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.

The kernel function directly affects the performance of KSVM because different kernel functions have different data mapping.

<table>
<thead>
<tr>
<th>Class</th>
<th>PQ Disturbance</th>
<th>Class</th>
<th>PQ Disturbance</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>Normal</td>
<td>C13</td>
<td>Interrupt + harmonics</td>
</tr>
<tr>
<td>C2</td>
<td>Swell</td>
<td>C14</td>
<td>Swell + transient</td>
</tr>
<tr>
<td>C3</td>
<td>Sag</td>
<td>C15</td>
<td>Sag + transient</td>
</tr>
<tr>
<td>C4</td>
<td>Interruption</td>
<td>C16</td>
<td>Spike + transient</td>
</tr>
<tr>
<td>C5</td>
<td>Harmonics</td>
<td>C17</td>
<td>Transient + harmonics + sag</td>
</tr>
<tr>
<td>C6</td>
<td>Swell + harmonics</td>
<td>C18</td>
<td>Transient + harmonics + swell</td>
</tr>
<tr>
<td>C7</td>
<td>Sag + harmonics</td>
<td>C19</td>
<td>Transient + harmonics + interrupt</td>
</tr>
<tr>
<td>C8</td>
<td>Transient</td>
<td>C20</td>
<td>Transient + harmonics + flicker</td>
</tr>
<tr>
<td>C9</td>
<td>Flicker</td>
<td>C21</td>
<td>Flicker + harmonics + interrupt</td>
</tr>
<tr>
<td>C10</td>
<td>Flicker + harmonics</td>
<td>C22</td>
<td>Flicker + harmonics + sag</td>
</tr>
<tr>
<td>C11</td>
<td>Notch</td>
<td>C23</td>
<td>Flicker + harmonics + swell</td>
</tr>
<tr>
<td>C12</td>
<td>Spike</td>
<td>C24</td>
<td>Spike + transient + swell</td>
</tr>
</tbody>
</table>
TABLE II

PERFORMANCE UNDER DIFFERENT COMBINATIONS OF KERNEL FUNCTIONS

<table>
<thead>
<tr>
<th>Kernel functions</th>
<th>$u_1$</th>
<th>$u_2$</th>
<th>$u_3$</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear + Polynomial</td>
<td>0.5</td>
<td>0.5</td>
<td>0</td>
<td>79.60</td>
</tr>
<tr>
<td>Linear + Gaussian</td>
<td>0.5</td>
<td>0.5</td>
<td>0</td>
<td>93.06</td>
</tr>
<tr>
<td>Gaussian + Linear</td>
<td>0.5</td>
<td>0.5</td>
<td>0</td>
<td>96.94</td>
</tr>
<tr>
<td>Gaussian + Polynomial</td>
<td>0.5</td>
<td>0.5</td>
<td>0</td>
<td>84.58</td>
</tr>
<tr>
<td>Linear kernel</td>
<td>0.5</td>
<td>0.3</td>
<td>0.2</td>
<td>95.64</td>
</tr>
<tr>
<td>Polynomial kernel</td>
<td>0.5</td>
<td>0.3</td>
<td>0.2</td>
<td>90.44</td>
</tr>
<tr>
<td>Gaussian kernel</td>
<td>0.8</td>
<td>0.1</td>
<td>0.1</td>
<td>98.12</td>
</tr>
<tr>
<td>Gaussian kernel</td>
<td>0.9</td>
<td>0.05</td>
<td>0.05</td>
<td>98.82</td>
</tr>
</tbody>
</table>

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 $\delta$ 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. 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 article. 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 $\delta_t$ when the parameter value ranges from 1 to 10. On the contrary, the parameter $\delta_f$ is negatively correlated with a classification accuracy when $\delta_f$ 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 $OST_A(m, n)$ for multiple disturbances via maximizing energy concentration. The features $F, T$ are obtained from $OST_A(m, n)$.
TABLE III

<table>
<thead>
<tr>
<th>Kernel functions</th>
<th>Accuracy (%)</th>
<th>Test time per sample (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ST-KSVM</td>
<td>94.72</td>
<td>10 87.32 97.82 98.97 27</td>
</tr>
<tr>
<td>OST-KSVM</td>
<td>99.59</td>
<td>88 88.22 98.82 99.51 88</td>
</tr>
</tbody>
</table>

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.

From Table III, the accuracy rises with the noise level and 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.

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 are 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 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 two 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], [16], and [22] 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.

### 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 6105 A. After being preheated for an hour, the instrument randomly generates different signals includes Normal (C1), Swell (C2), Sag (C3), Interrupt (C4),
TABLE VIII
PERFORMANCE UNDER THE EXPERIMENTAL SIGNAL

<table>
<thead>
<tr>
<th>Class</th>
<th>Accuracy (%)</th>
<th>Class</th>
<th>Accuracy (%)</th>
<th>Average accuracy (%)</th>
<th>Test time per sample (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>100</td>
<td>C5</td>
<td>98.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C2</td>
<td>96.67</td>
<td>C6</td>
<td>95.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C3</td>
<td>98.33</td>
<td>C7</td>
<td>95.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C4</td>
<td>96.67</td>
<td>C9</td>
<td>96.67</td>
<td>97.08</td>
<td>92</td>
</tr>
</tbody>
</table>

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 floating 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 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 article, 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 were 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 were 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


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

She has been an Associate 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 the Hubei University of Technology, Wuhan, China, in 2015, and the M.Sc. degree in electrical engineering in 2017 from Hunan University, Changsha, China, where he is currently working toward the Ph.D. degree in electrical engineering.

He is also a Joint Doctoral Student with 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 the B.S. degree from Hunan University, Changsha, China, and the M.S. and Ph.D. degrees from Tufts University, Medford, MA, USA, all in electrical engineering.

He is an Associate Professor and Director of the Vision and Image Processing Laboratory, Department of Computer and Information Science, University of Macau, Macau, China. 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. 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 the IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, and four other journals.