Received March 30, 2019, accepted May 16, 2019, date of publication May 20, 2019, date of current version June 6, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2917886

Power Quality Disturbance Feature Selection and Pattern Recognition Based on Image Enhancement Techniques

LIN LIN[1], DA WANG[2], SHUYE ZHAO[3], LINGLING CHEN[1], AND NANTIAN HUANG[4]

[1]College of Information and Control Engineering, Jilin Institute of Chemical Technology, Jilin 132022, China
[2]Dezhou Power Supply Company, State Grid Shandong Electric Power Co., Ltd., Dezhou 253000, China
[3]Economic and Technical Research Institute, State Grid Inner Mongolia Eastern Electric Power Co., Ltd., Mongolia 010010, China
[4]Key Laboratory of Modern Power System Simulation and Control and Renewable Energy Technology, Ministry of Education, Northeast Electric Power University, Jilin 132012, China

Corresponding author: Lin Lin (jllinlin@126.com)

This work was supported in part by the National Nature Science Foundation of China under Grant 51307020, in part by the Science and Technology Development Project of Jilin Province under Grant 20160411003XH, in part by the Science and Technology Project of Jilin Province Education Department under Grant JJKH20170219KJ, in part by the Major Science and Technology Projects of Jilin Institute of Chemical Technology under Grant 2018021, in part by the Science and Technology Innovation Development Plan Project of Jilin City under Grant 201750239, in part by the Science and Technology Program Project of Jilin Provincial Science and Technology Department under Grant 20180101336JC.

ABSTRACT In the existing research of power quality disturbance (PQD) identification, the efficiency of signal processing is low and cannot meet the needs of practical application analysis. Furthermore, due to the lack of effective analysis of features, the complexity of classifiers is increased, and the efficiency of classification reduced by the redundant features. In this paper, in order to overcome these shortcomings, a PQD recognition method based on image enhancement techniques and feature importance analysis is proposed. First, PQD signals are converted into gray images, and three image enhancement techniques include gamma correction, edge detection, and peaks and valley detection are used to enhance the disturbance features. Then, the disturbance features are extracted from the binary images, and the original feature set is constructed, the classification ability of each feature is measured by Gini importance. Based on the descending order of the Gini importance, the sequence forward search (SFS) method is used for feature selection to determine the optimal feature subset. Finally, random forest (RF) classifier is constructed by the optimal feature subset to identify the PQD signals. The results of the simulation and contrast experiments show that the new method can determine the optimal classification subset, which recognizes the PQD signals effectively in different noise environments. Furthermore, the new method has higher signal processing efficiency compared with the EMD and ST methods.

INDEX TERMS Power quality, disturbance classification, image processing, feature selection, Gini importance, random forest.

I. INTRODUCTION

The matter of the PQD is more serious because of the application of power electronic devices, nonlinear loads and solid-state switching devices in the power system. In addition, the distributed photovoltaic, wind power and other new energy are connected to the grid, which further deteriorates the power quality [1]–[3]. The problems of power quality have brought serious national economic losses, and industrial production and resident life are affected. Therefore, it is urgent to improve the power quality [4]. The classes of PQD signals are complex, and the accurate PQD detection so as to carry out targeted governances is necessary for improving of the power quality and the stability of the power system [5]. In addition, the high sampling rate PQD signals collected by sensor devices have higher requirements for the complexity of the disturbance recognition methods.

The recognition of PQD signals includes three steps: signal processing, feature selection and pattern recognition. The shortcomings in the research process are as follows: (1) The analysis ability of signal processing methods is affected by the setting of parameters. Furthermore, the efficiency of the

The associate editor coordinating the review of this manuscript and approving it for publication was Vigna K. Ramachandaramurthy.

signal processing method is low, which cannot meet the demand of real-time analysis of high sampling rate PQD signals. (2) After obtaining the original feature set, there is no effective analysis of the classification ability of features. The huge feature dimension in the original feature set reduces the efficiency of feature extraction, and the redundant features increase the complexity of the classifier. (3) The classification effect of common single classifiers can be further improved.

S-Transform (ST) [6]–[8], Empirical Mode Decomposition (EMD) [9], Ensemble Empirical Mode Decomposition (EEMD) [10] and Wavelet Transform (WT) [11] are widely used in signal processing. WT carry out the multi-scale transform on disturbance signals, and extract features from the wavelet coefficients of each wavelet layer as well as the approximate coefficients. WT has achieved high recognition accuracy in the existing PQD identification studies. However, the analysis ability of WT is affected by the selection of wavelet basis and the decomposition layers of signals. EMD and EEMD methods decompose the disturbance signals into multi-layers, and extract features from each layer of signals. However, the performance of the algorithm is also affected by the number of decomposition layers. ST carries out time-frequency analysis on signals and obtains the time-frequency matrix, which is used to extract the required time-frequency features. ST has been widely used in the field of PQD identification and has achieved good analysis results. However, the analysis ability of ST is affected by parameters such as window width factor, and the computation of the algorithm is large. As we all know, the huge amount of computation of ST is an important reason that limits its practical application [4]. Therefore, the huge amount of computation and running time in the existing signal processing process cannot meet the needs of real-time analysis. Under the current situation of large data analysis of power quality, it is of great practical significance to improve the efficiency of feature extraction on the premise of ensuring high classification accuracy.

In literature [12], the power quality signal in the form of original time series are converted into gray image, and the disturbance features are enhanced through different image enhancement techniques. Compared with traditional signal processing methods, the feature extraction efficiency of literature [12] is improved significantly. Therefore, the study has made a useful exploration for the application of image enhancement techniques in PQD recognition. Different image enhancement techniques are used for different PQD signals in literature [12], However, how to determine which image enhancement technique and which feature to use for pattern recognition are not analyzed. In literature [12] and [13], the methods of gamma correction, edge detection and peaks and valley detection are used for different classes of disturbances to enhance the gray image features. On this basis, morphological features are extracted from the binary images for disturbance analysis. However, from the point of view of feature selection, there is no analysis on the ability of feature classification for different feature combinations.

Some studies lack the analysis of feature selection. The excessive feature dimension not only reduces the efficiency of feature extraction, but also increases the complexity of classifier and affects the classification ability [14]. Therefore, it is of great practical significance to remove the redundant features which have little contribution to classification, and determine the optimal feature subset. The reasonable feature selection can improve the efficiency of feature extraction, simplify the structure of classifier and improve the classification effect. The existing methods of feature selection mainly include the filter method and the wrapper method. The filter method is carried out based on statistical characteristics of features, which made it difficult analyze the classification ability of different feature combinations [15]. The wrapper method is carried out by combining with genetic algorithm or other intelligent algorithms, but the efficiency of the algorithm is low [16]. Therefore, it is of great practical significance to design a feature selection method for analyzing feature combination classification ability and optimization efficiency.

In the classifier design, various classifiers such as neural network (NN) [17], support vector machine (SVM) [18], extreme learning machine (ELM) [19] and decision tree (DT) [20] have achieved good results in PQD recognition. However, there are many parameters to be set in NN and SVM, and they are prone to overfitting in the classification process. DT has simple structure and good classification effect compared with SVM and NN, the accuracy of DT is better than that of SVM under the same set of training and testing samples. However, the generalization ability of DT is poor and the setting of the classification threshold depends on the training samples. RF [21] is an excellent integrated classification algorithm. RF has many advantages, such as better noise immunity, few parameters and little influenced by overfitting problems.

Furthermore, RF is an effective integrated feature selection method. According to the classification effect of each RF node in the training process, the Gini importance ranking [22], [23] of features can be obtained. The Gini importance of features can be used as a reference index to determine the optimal feature subset, the analysis process is similar to that of Filter method. Gini index has been used as an important index to analyze the ability of feature classification by many classification problems, and good results have been achieved. The literature [24] ranks all features based on Gini index and features with the highest Gini importance are used to train SVM classifiers. The result of experiment shows that the selected features have good classification ability. The literature [25] also regards Gini index as the basis to measure the ability of feature classification. The result of experiment shows that at the time that the number of features is large, the feature selection method based on Gini index analysis can achieve the best classification effect. The above researches prove that Gini index is effective for feature classification ability analysis. Gini importance obtained during RF training is used by the proposed method to analyze the classification

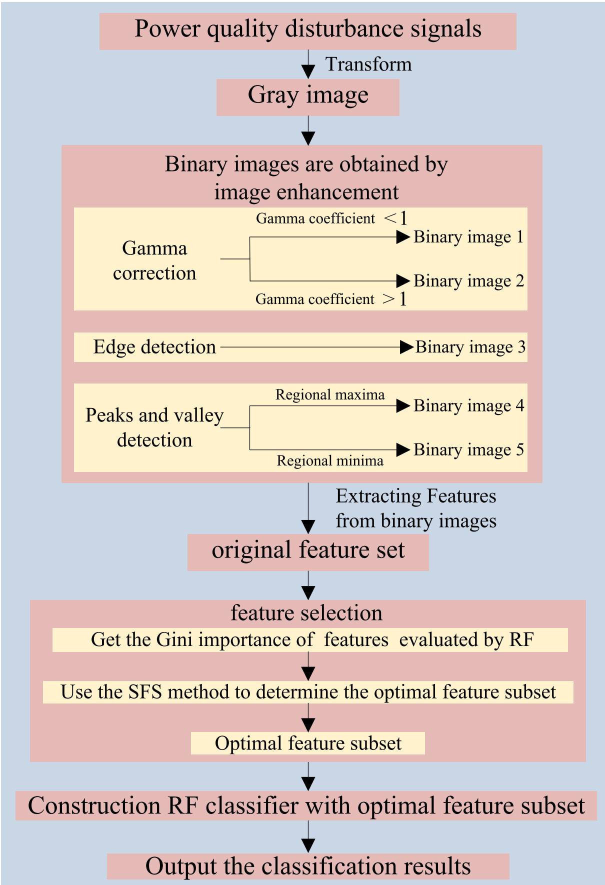


FIGURE 1. Process of the research.

ability of features, on this basis, feature selection is carried out by combining sequence forward search method. The proposed method not only can analyze the classification ability of different feature combinations, but also has higher efficiency than Wrapper method. In order to extract disturbance features of signals comprehensively, it is necessary to establish an original feature set with a large number of features and complex feature types. The proposed method has good feature analysis ability and feature selection efficiency, so it can meet the analysis needs of complex original feature sets.

Research process of the new method is as follows:

On the basis of feature selection based on the SFS method and RF, a new method of PQD recognition based on image enhancement techniques is proposed in this paper. Firstly, nine kinds of common PQD signals including standard signals are generated by simulation, and the PQD signals in time series form are converted into gray images. Secondly, features of the gray images are enhanced by three image enhancement techniques, including gamma correction, edge detection and peaks and valley detection. On this basis, the original feature set is constructed. After that, Gini importance of all features are evaluated by RF, and the SFS method was adopted for feature selection. Finally, the optimal feature subset obtained by feature selection is used to reconstruct the RF classifier, and the optimal RF classifier is used to classify PQ signals.

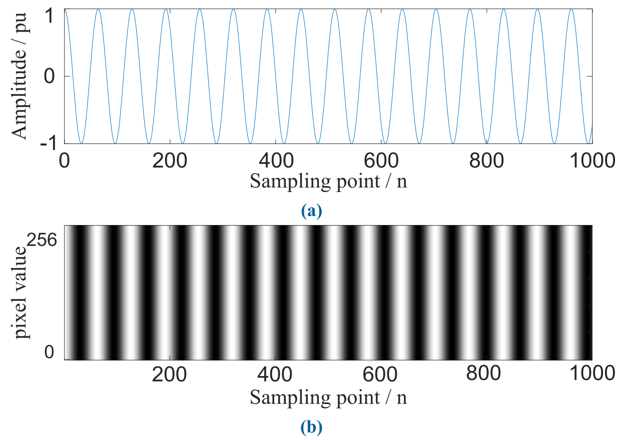


FIGURE 2. Standard power quality signal and its gray image, (a) Standard power quality signal, (b) Gray image of the standard power quality signal.

II. IMAGE ENHANCEMENT TECHNIQUES

Essence of the gray image is a set of data whose values are in a scope of a certain intensity, and the value of each pixel represents its intensity. The range of the pixel values is 0-256 in gray image, the intensity 0 represents black and the intensity 256 represents white. The original PQD signals are converted into gray image according to the following rules [12]:

$$f : \Re^1 \longrightarrow [0, 256]^1 \tag{1}$$

Waveforms of the original PQD signals are sampled, and each sampling point corresponds to one pixel of the gray image. The standard power quality signal and its gray image are shown in Figure 2.

According to Fig. 2, the gray image transformed from standard power quality signal has smooth gray transformation and no mutation. If the disturbance component appears in the standard signal, the obvious gray change can be found in the gray image. This characteristic of gray images is the basis of PQD signals recognition using image enhancement techniques. After the gray image is obtained, image enhancement techniques include gamma correction, edge detection and peaks and valley detection are used to highlight the disturbance components of the gray image.

A. GAMMA CORRECTION

Gamma correction [27] is also known as exponential transformation, it is a linear gray transformation. It can be expressed as:

$$s = (cy)^\gamma \tag{2}$$

where both c and γ are positive constants. γ is gamma coefficient, which is a very important parameter in gamma correction. According to the difference of γ value, gamma transform can enhance the contrast of low gray area or high gray area selectively. The effect of curve generated by $\gamma > 1$ is completely opposite to that generated by $\gamma < 1$. The contrast of the high gray area of the image is enhanced when $\gamma > 1$; The contrast of the low gray area of the image is

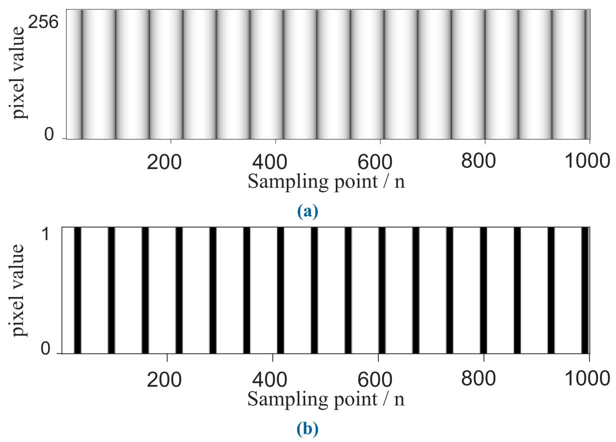


FIGURE 3. Gray image and binary image of the standard power quality signal when $\gamma = 0.125$, (a) Gray image obtained by gamma correction, (b) Binary image obtained by the otsu algorithm.

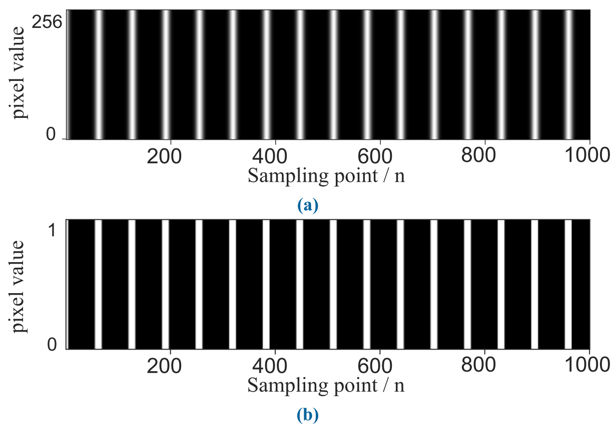


FIGURE 4. Gray image and binary image of the standard power quality signal when $\gamma = 8$, (a) Gray image obtained by gamma correction, (b) Binary image obtained by the otsu algorithm.

enhanced when $\gamma < 1$; The original image will not change when $\gamma = 1$.

The image obtained by gamma correction is still gray image. In order to further highlight the features of disturbance signal, the otsu algorithm is used to select the optimal threshold and the gray image is transformed into binary image. Binary image represents a logical array with values of 0 and 1. If the gray level of the gray image is L and the gray range is $[0, L-1]$, the process of calculating the optimal threshold using the otsu algorithm is expressed as:

$$t = Max[w0(t) * (u0(t) - u)^2 + w1(t) * (u1(t) - u)^2)] \tag{3}$$

where $w0$ is the background ratio, $u0$ is the background mean value, $w1$ is the foreground ratio, $u1$ is the foreground mean value and u is the mean value of the whole image. The optimal segmentation threshold is t which maximizes the value of the expression (3).

Figures 3 and 4 show the gray images of standard power quality signal after gamma correction when γ takes different values, and the binary images obtained by the otsu algorithm.

B. EDGE DETECTION

The region with the most significant change in local intensity in an image is called image edge. Edge detection aims to remove the redundant information from original image, and the image features that relatively important are retained for image recognition. The essence of edge detection is to extract the boundary between the object and the background from the image using special algorithms. The algorithm used in this paper is canny edge operator [28], the algorithm has the advantages of lower erroneous judgment and high positioning accuracy. The steps of canny edge operator are as follows:

1) Gauss filter is used to process the image to remove noise and smooth the image. The process can be expressed as:

$$J = I \otimes G \tag{4}$$

where G is the gradient intensity, I is the image to be smoothed and J represents the smoothed image.

2) Calculate the gradient direction and gradient magnitude of each pixel in the smoothed image. The process can be expressed as:

$$\nabla J = (J_x, J_y) = \left(\frac{\partial J}{\partial x}, \frac{\partial J}{\partial y}\right) \tag{5}$$

where J_x and J_y represent the gradients of the smoothed image J in the x, y directions respectively. The gradient direction and gradient magnitude are expressed as formula (6) and formula (7).

$$|\nabla J| = \sqrt{J_x^2 + J_y^2} \tag{6}$$

$$\theta = \arctan\left(\frac{|J_y|}{|J_x|}\right) \tag{7}$$

3) All the maximum values of gradient image are saved, and other non-maximum values are deleted to achieve non-maximum suppression, so that the blurred edges in gradient image is transformed into clear edges.

4) The remaining pixels can represent the actual edges of the image more accurately after non-maximum suppression, but there are still some edge pixels caused by noise. Therefore, the dual threshold method is used by canny edge operator to further determine the actual edge. The edge pixel will be marked as a strong edge pixel if its gradient value is higher than the high threshold value. The edge pixel will be marked as a weak edge pixel if its gradient value is less than the high threshold but greater than the low threshold; The edge pixel will be suppressed if the gradient value of edge pixels is less than the low threshold value.

5) It is determined whether the weak edge pixels obtained by the above steps are extracted from the real edges through suppressing the isolated low threshold points. The weak edge pixel can be retained as the real edge if there are strong edge pixels in the weak edge pixel and its eight adjacent pixels. Figure 5 shows the binary image of standard power quality signal obtained by canny edge detection.

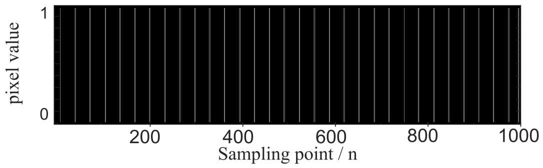


FIGURE 5. Binary image of standard power quality signal obtained by canny edge detection.

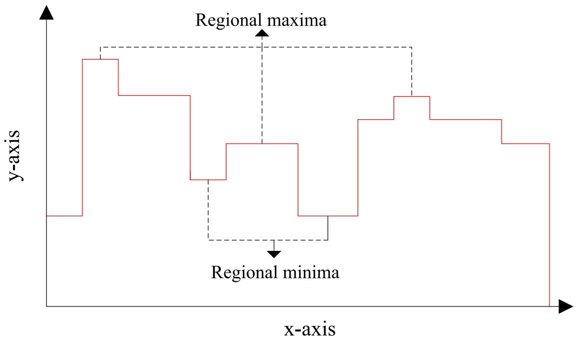


FIGURE 6. Maximum and minimum in gray image.

C. PEAKS AND VALLEY DETECTION

The gray image of the signal is regarded as a three-dimensional image. The x axis and y axis represent the position of the pixel respectively, and the z axis represents the pixel intensity. As shown in Figure 6, the positions with higher gray values and lower gray values correspond to the peaks and valleys in topographic maps [29]. The maxima and minima of the PQD signals are displayed by the intensity of different regions in the gray image, so as to reflect the change of the amplitude of the PQD signals. Therefore, the peaks and valley detection has a good detection effect on voltage sag, swell and interruption.

Binary image of the standard power quality signal obtained by peak and valley detection are shown in Figure 7.

III. POWER QUALITY FEATURE SELECTION AND DISTURBANCE RECOGNITION BASED ON RF

A. CLASSIFICATION THEORY OF RF

RF [21] is an excellent integrated classification algorithm. It combines DT with ensemble learning to form a new set of classifier:

$$\{h(x, \Theta_k), k = 1, \ldots\} \tag{8}$$

where $h(x, \Theta_k)$ represents the meta classifier. x is the input vector and k is the number of meta classifiers, Θ_k represents the random vector. RF generates a subset of random features at different nodes, and chooses the feature with the best classification effect as the classification feature. Finally, all the classification conclusions are summarized to achieve accurate classification.

The classifier set $H(x) = \{h_1(x), h_2(x), \ldots, h_k(x)\}$ is given, and the training set of each classifier is extracted from the original data set (X, Y). The margin function is

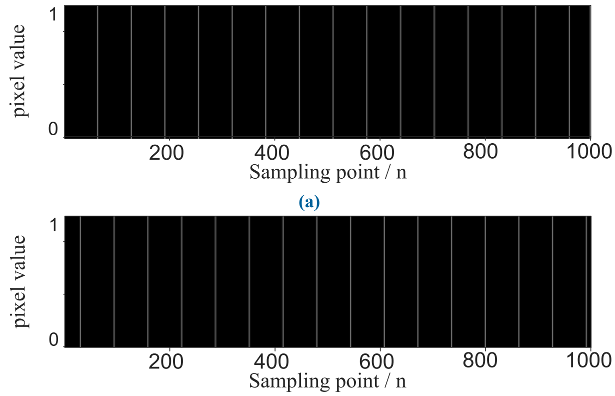


FIGURE 7. Binary image of standard power quality signal obtained by peak and valley detection, (a) binary image of local maxima and (b) binary image of local minimum.

expressed as:

$$mg(X, Y) = av_k I(h_k(X) = Y) - \max_{j \neq Y} av_k I(h_k(X) = j) \tag{9}$$

where $I(\cdot)$ represents the indicator function, $av_k(\cdot)$ represents average value. Y is the correct classification of the vector and j is the incorrect classification of the vector. The generalization error PE^* used to evaluate the classification ability of the classifier is expressed as:

$$PE^* = P_{X,Y}(mg(X, Y) < 0) \tag{10}$$

where X and Y is the definition of space. The process of the PQD classification based on RF is as follows:

1) n samples were randomly selected from the original feature set N to form a self-service sample set, the process is repeated L times. The value of L in feature selection stage is 200.

2) In the training stage, the features are randomly extracted from the feature space M to be the candidate feature of non-leaf nodes. Each candidate feature is used to segment the node and the best feature is selected as the segmentation feature of the node. Repeat the process until the non-leaf nodes of each tree are classified and the training process is completed.

3) In the classification stage, the majority voting method is used for each meta classifier to determine the optimal classification results.

B. PTHEORY OF GINI IMPORTANCE

In the training process of RF, the classification effect of different feature segmentation nodes can be measured by Gini index [22], [23]. The classification ability of each feature is sorted by Gini importance, and the optimal feature subset is determined by feature selection.

Assume S is a sample set contains s samples, and the classes of the samples are n. Let s_i represents the sample size of the class i ($i = 1, 2, \ldots, n$), and the Gini index of S is expressed as:

$$Gini(S) = 1 - \sum_{i=1}^{n} P_i^2 \tag{11}$$

where $P_i = s_i/s$ represents the probability that the sample comes from class i. When RF uses the feature to split the node, the S is split into m subsets ($S_j, j = 1, 2, \ldots, m$). The Gini index of S is expressed as:

$$Gini_{split}(S) = \sum_{j=1}^{m} \frac{s_j}{s} Gini(S_j) \tag{12}$$

where s_j is the number of samples in S_j. It can be seen from the formula (12) that the feature classification with the minimum A value is the best.

The calculation process of feature Gini importance ranking is as follows: Firstly, the RF classifier is trained by the original feature set, and the $Gini_{split}$ value of each feature is calculated. Use the Gini index of the feature before dividing minus the value of the $Gini_{split}$ to get the Gini importance of the feature. Then, the feature with the maximum Gini importance is used for segmentation. Finally, all Gini importance of the feature are superimposed and sorted, and the Gini importance ranking of all features is obtained.

After the Gini importance of all the features in the original feature combination are obtained, the SFS method based on RF classifier and Gini importance can be implemented. Firstly, arrange the Gini importance of features in descending order, and the features are added to the feature set Q in turn. Then, the classifier is retrained with the updated set Q after a feature is added, and the corresponding recognition accuracy is recorded. Repeat the process until all features in the original feature set are added to Q. Finally, the optimal feature subset is determined by considering the classification accuracy and the dimensions of the feature subset.

IV. EXPERIMENT RESULTS

A. PQD SIGNALS FEATURE EXTRACTION BASED ON IMAGE ENHANCEMENT TECHNIQUES

Referring to [30], [31], the standard power quality signal and 8 kinds of PQ signals include sag(C1), swell(C2), interruption (C3), flicker(C4), transient(C5), harmonic(C6), notch(C7), spike(C8) are generated by simulation. The fundamental frequency of signals is 50 Hz, and the sampling rate is 3200 Hz. In order to ensure the reliability and authenticity of simulation signals, the starting and stopping time, amplitude and oscillation frequency of each disturbance signal are generated by random function in the standard range. Meanwhile, the noise between 20 dB and 50 dB is added to the signal.

In order to extract the effective features that can reflect the characteristics of PQD signals, the disturbance signals are transformed into gray images. After that, Three kinds of image enhancement techniques include gamma correction, peaks and valley detection and edge detection are used to enhance the disturbance features of gray images, and binary images of the disturbance signal are obtained. The original waveforms and gray images of eight kinds of disturbance signals are shown in Fig. 8.

According to Fig. 8, gray images can reflect the signal characteristics of different disturbance classes. Gray intensity of the disturbance area in sag and interrupt signals are darker than that in the normal area. Gray intensity of the disturbance area in swell signals are lighter than that in the normal area. The brightness of the disturbance area in the flicker signal changes periodically. The fringes in the disturbance area of transient oscillation signal are dense and close to each other. The fringes in harmonic signals are more granular than those in normal signals. The fringes in notch signal are thin and dark when disturbance occurs. The fringes in spike signal are thin and light when disturbance occurs. Therefore, the characteristics reflected in different gray images can be used as the basis for disturbance classification.

In order to highlight the features of various disturbance signals, the image enhancement technologies are used to enhance the gray image and the binary images of disturbance signals are obtained. In this study, the gray images are enhanced respectively by gamma correction, edge detection and peaks and valley detection. Binary images of eight kinds of disturbance signals obtained by gamma correction, peaks and valley detection, edge detection are shown in Figures 9, 10 and 11 respectively.

The enhancement effect of gamma correction on gray images are different when different γ values are taken. The contrast of high grey level district in gray images are enhanced when $\gamma > 1$, and the contrast of low grey level district are enhanced when $\gamma < 1$. In order to fully characterize the perturbation characteristics, the gray image is enhanced when γ is taken as $\gamma = 0.125$ and $\gamma = 8$ respectively.

According to Fig. 9, 10 and 11, because the principles of gamma correction, peaks and valley detection and edge detection are different, their enhancement effects on various disturbance signals are also different. Gamma correction method is applicable to sag, swell and flicker signals. Peaks and valley detection method is applicable to interruption, notch and spike signals. Edge detection method is applicable to transient signal. Because the classes of PQDs are unknown in practical engineering applications, in order to comprehensively analyze the characteristics of signals, the three image enhancement techniques are used to process the disturbance signals at the same time. After that, the disturbance features are extracted from the binary images, and the original feature set with abundant and effective information is obtained.

Referring to [13], the area and Euler number of binary images are calculated to construct the feature vector. Furthermore, the following image features: angular two moment, contrast, correlation, mean value, variance value, inverse difference moment and entropy are calculated to construct the feature vector.

Assume L is the number of gray levels in binary images, $P(m, n)$ represents the gray values of row m and column n. The calculation methods of each feature are expressed as:

Angular two moment:

$$\beta = \sum_{m=0}^{L-1} \sum_{n=0}^{L-1} P(m, n)^2 \tag{13}$$

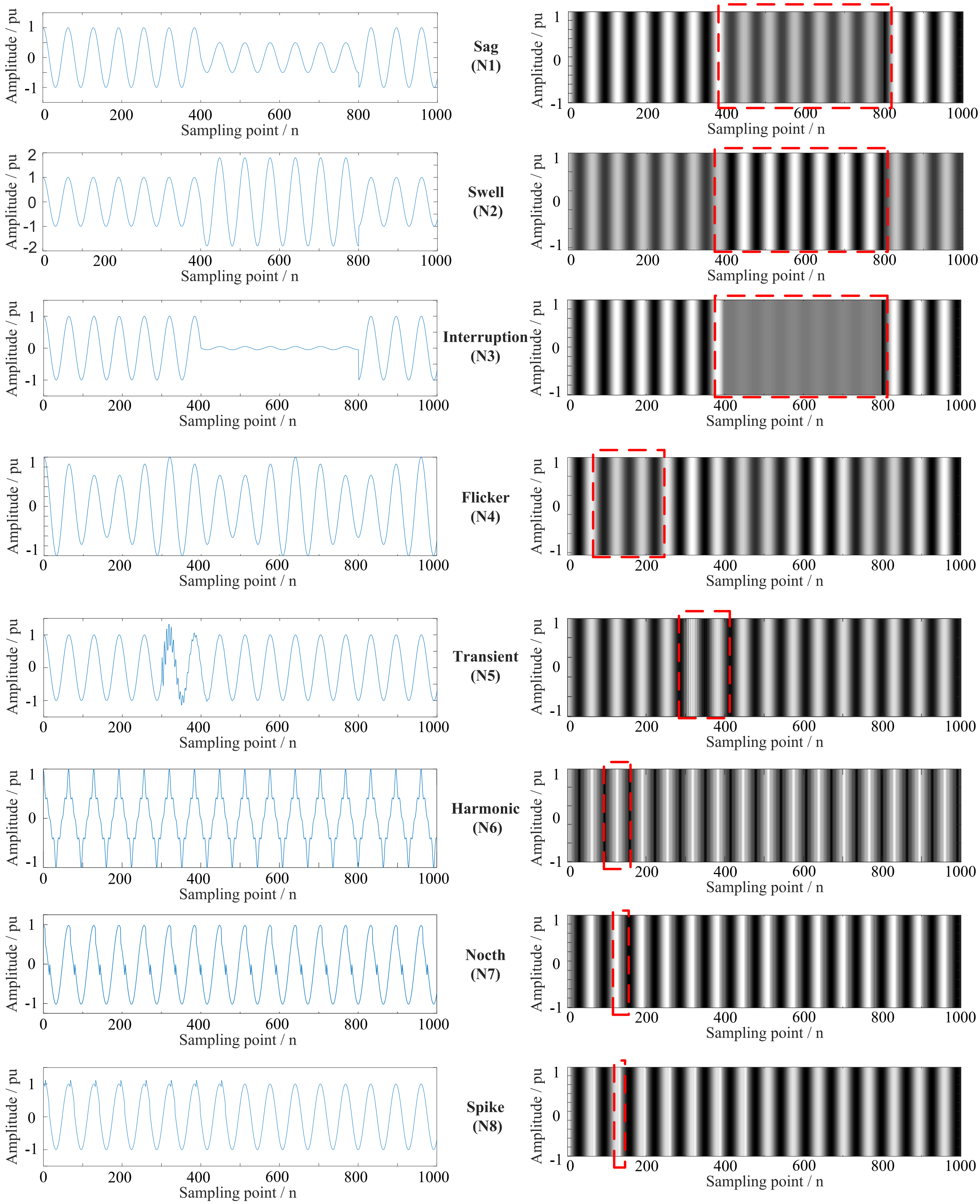


FIGURE 8. Original waveforms and gray images of eight kinds of disturbance signals.

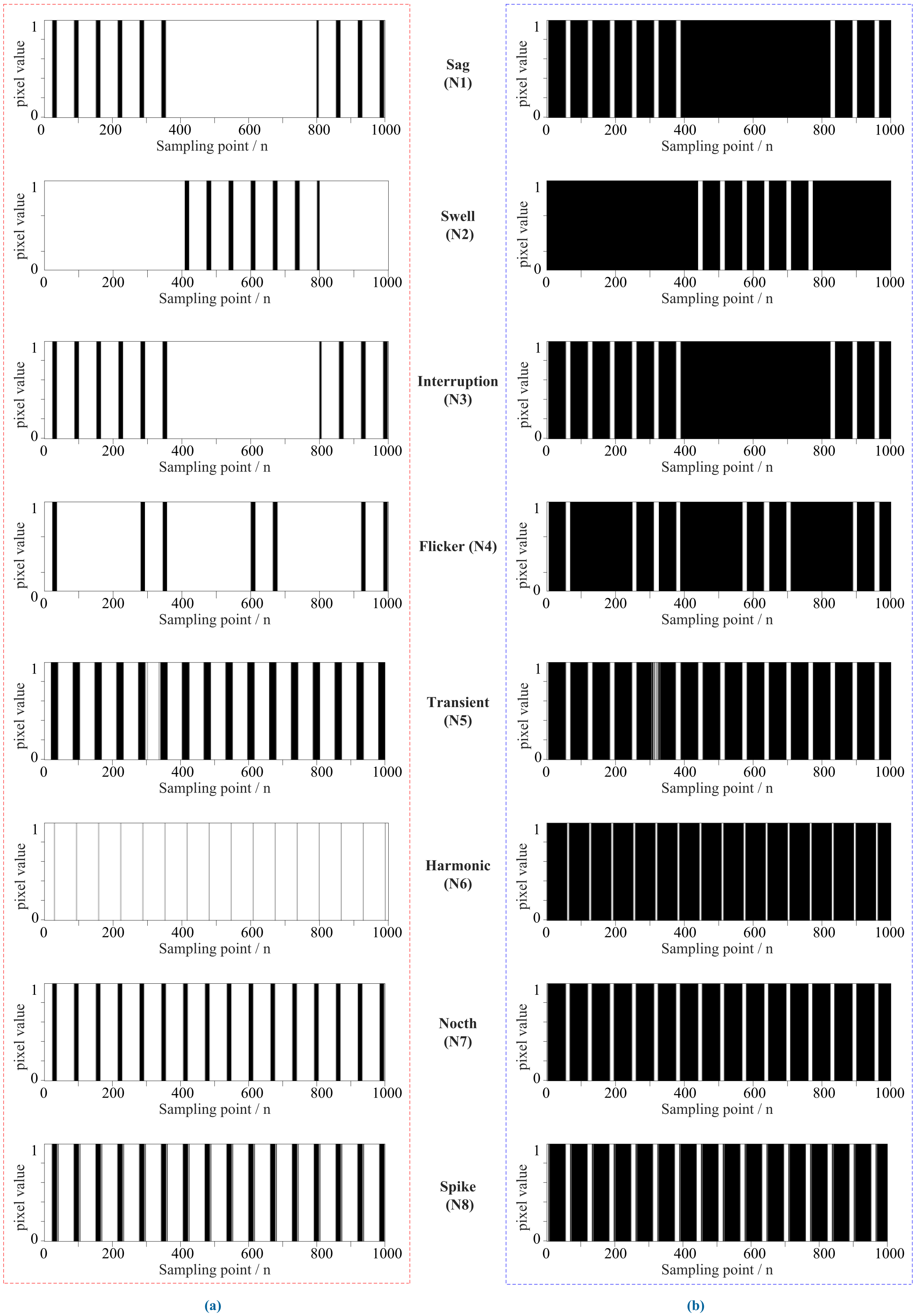


FIGURE 9. Binary images obtained by gamma correction and otsu algorithm, (a) binary images when $\gamma = 0.125$ and (b) binary images when $\gamma = 8$.

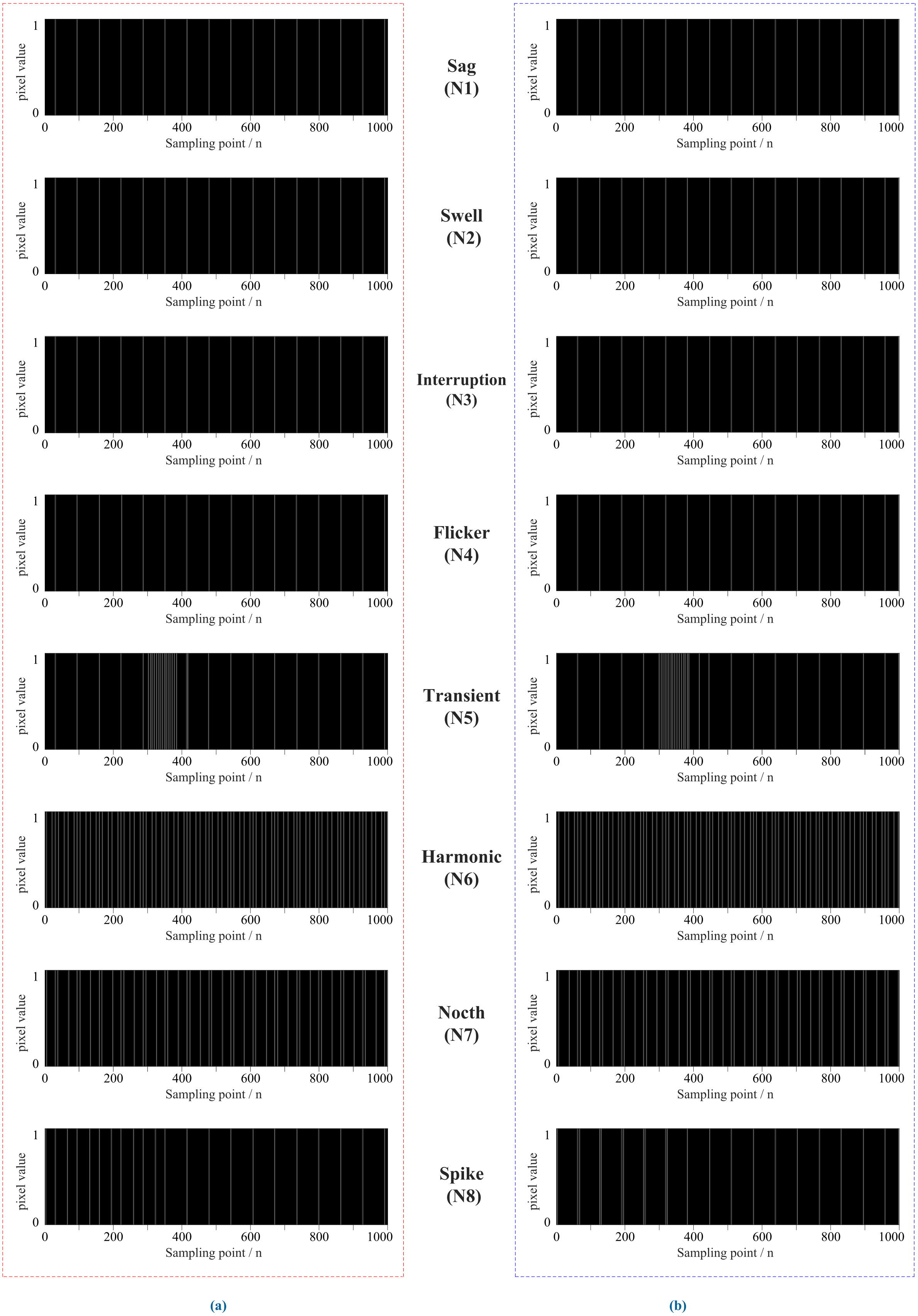


FIGURE 10. Binary images obtained by peaks and valley detection, (a) binary image of local maxima and (b) binary image of local minimum.

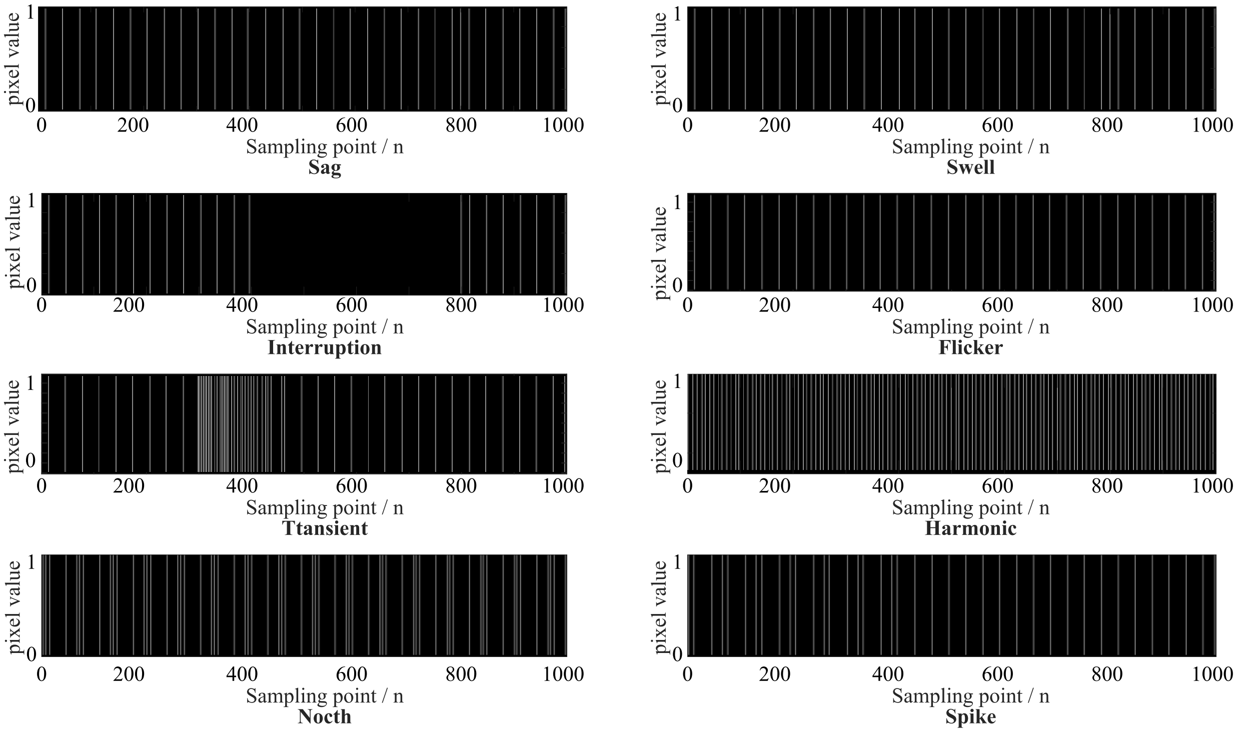


FIGURE 11. Binary images obtained by edge detection.

Contrast:

$$\alpha = \sum_{t=0}^{L-1} t^2 \left\{ \sum_{m=0}^{L-1} \sum_{n=0}^{L-1} P(m, n) \right\} \tag{14}$$

where t represents gray levels.

Correlation:

$$G = \sum_{m=0}^{L-1} \sum_{n=0}^{L-1} \frac{mnP(m, n) - \mu_1 \mu_2}{\sigma_1^2 \sigma_2^2} \tag{15}$$

where μ_1 and μ_1 represent mean values, σ_1 and $\sigma_{.2}$ represent variance values. The expressions of each value are expressed as:

$$\mu_1 = \sum_{m=0}^{L-1} m \sum_{n=0}^{L-1} P(m, n) \tag{16}$$

$$\mu_2 = \sum_{m=0}^{L-1} n \sum_{n=0}^{L-1} P(m, n) \tag{17}$$

$$\sigma_1 = \sum_{m=0}^{L-1} (m - \mu_1)^2 \sum_{n=0}^{L-1} P(m, n) \tag{18}$$

$$\sigma_2 = \sum_{m=0}^{L-1} (m - \mu_2)^2 \sum_{n=0}^{L-1} P(m, n) \tag{19}$$

Mean value:

$$E = \sum_{m=0}^{L-1} \sum_{n=0}^{L-1} mP(m, n) \tag{20}$$

Variance value:

$$\sigma = \sum_{m=0}^{L-1} \sum_{n=0}^{L-1} (m - \mu)^2 P(m, n) \tag{21}$$

where μ represents the mean of $P(m, n)$.

Inverse difference moment:

$$I = \sum_{m=0}^{L-1} \sum_{n=0}^{L-1} \frac{P(m, n)}{1 + (m - n)^2} \tag{22}$$

Entropy:

$$Entropy = \sum_{m=0}^{L-1} \sum_{n=0}^{L-1} P(m, n) \log(P(m, n)) \tag{23}$$

Referring to the above methods of feature calculation, nine features are extracted from the five kinds of binary images obtained by three image enhancement techniques. The five binary images are: Two kinds of binary images obtained by gamma correction method when $\gamma = 0.125$ and $\gamma = 8$ respectively; Binary image of local maxima and binary image of local minimum obtained by peaks and valley detection; Binary image obtained by edge detection. The calculation principles of 45-dimensional features are shown in Table 1.

B. PQD FEATURE SELECTION BASED ON GINI IMPORTANCE AND SFS METHOD

The RF classifier is trained with the original feature set, and Gini importance of all features are obtained. Based on the descending order of Gini importance, the SFS method is used

TABLE 1. Construction principle of original feature set.

Feature classes	Gamma correction		Edge detection	Peaks and valley detection	
	Binary image 1	Binary image 2	Binary image 3	Binary image 4	Binary image5
Area	F1	F2	F3	F4	F5
Euler number	F6	F7	F8	F9	F10
Angular two moment	F11	F12	F13	F14	F15
Contrast	F16	F17	F18	F19	F20
Correlation	F21	F22	F23	F24	F25
Mean value	F26	F27	F28	F29	F30
Variance value	F31	F32	F33	F34	F35
Inverse difference moment	F36	F37	F38	F39	F40
Entropy	F41	F42	F43	F44	F45

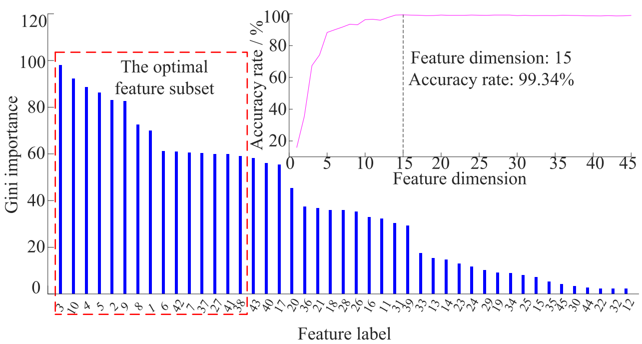


FIGURE 12. The process of feature selection based on Gini importance and SFS method.

to determine the optimal feature subset. Firstly, 9 kinds of PQ signals with random disturbances parameters and the signal-to-noise ratio (SNR) between 20 dB and 50 dB (random SNR) are generated by simulation. 600 samples of each type are applied to train RF classifier and get the Gini importance of features. Furthermore, 200 samples of each type are used to build the validation set and complete the feature selection. After that, according to the descending order of Gini importance, each feature is added to feature subset Q and the accuracy of RF classification under Q is obtained. Finally, the optimal feature subset is determined by considering the classification accuracy and feature dimension. The descending order of Gini importance and the classification accuracy under each feature subset are shown in Figure 12.

Classification stability is an important index to measure the performance of classifiers, which will affect the process of feature selection. Due to the influence of experimental

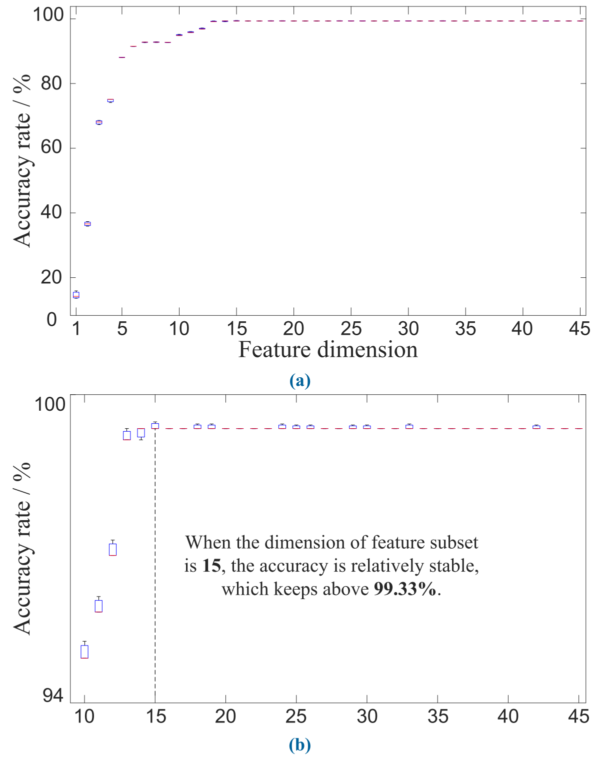


FIGURE 13. The fluctuation of RF accuracy under different feature subsets, (a) overall presentation and (b) enlarge the accuracy to 94%-100%.

environment, the accuracy of classifier will fluctuate in many times experiments under the same feature set. If the fluctuation of classifier is large, it is not conducive to the determination of the optimal feature subset. In order to analyze the stability of the RF classifier, several experiments were carried out on the same feature subset in the process of feature selection. By extracting the minimum value, the lower quartiles, the median, the upper quartile and the maximum value elements of the classification results, the overall classification results are finally displayed in the form of a box chart. The experimental results are shown in Fig. 13.

Fig. 12 and Fig. 13 show that RF has high stability under each feature subset, and the stability of RF is more higher with the enrichment of the feature set, When the feature subset is 15 dimensions, the accuracy of RF can be maintained above 99.33%. After that, if we continue to add features, the accuracy does not improve significantly. On the premise of ensuring good classification accuracy, reducing feature dimension can effectively improve the efficiency of feature extraction and simplify the structure of classifier. Therefore, considering the classification accuracy and feature dimension, the first 15 dimensional features in the descending ranking of feature Gini importance degree are determined to be the optimal feature subset. The optimal feature subset is: [F3 F10 F4 F5 F2 F9 F8 F1 F6 F42 F7 F37 F27 F41 F38]. RF can achieve a higher classification accuracy under the optimal feature subset, and the classification result is more stable. The experimental results show that RF has high

TABLE 2. Classification of new method (the number of features is 15, SNR is 20dB-50 dB).

class	C0	C1	C2	C3	C4	C5	C6	C7	C8	Accuracy
C0	199	0	0	0	0	1	0	0	0	99.5%
C1	0	197	0	3	0	0	0	0	0	98.5%
C2	0	0	200	0	0	0	0	0	0	100.0%
C3	1	3	0	196	0	0	0	0	0	98.0%
C4	0	0	0	0	200	0	0	0	0	100%
C5	0	0	0	0	0	200	0	0	0	100.0%
C6	0	0	0	0	0	0	200	0	0	100.0%
C7	0	0	0	0	0	0	0	200	0	100.0%
C8	3	0	0	1	0	0	0	0	196	98.0%
Average accuracy：99.33%										

classification stability, which is conducive to the determination of the optimal feature subset in the process of feature selection.

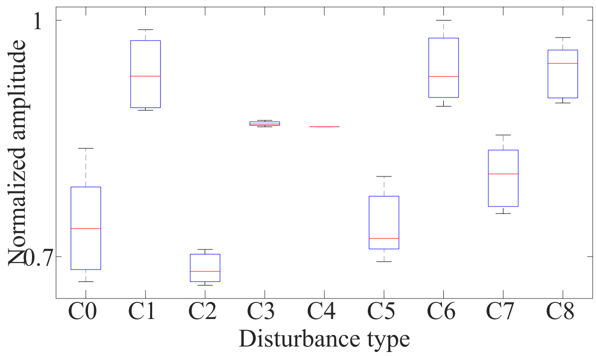
The classification accuracy of the validation set under the optimal feature subset are shown in table 2.

In order to verify the effectiveness of Gini importance as a measure of feature classification ability, the numerical distributions of PQD signals at the highest Gini importance (F3) and lowest Gini importance (F12) are analyzed respectively. The feature F3 and F12 are calculated from nine kinds of PQD signals, and the distribution of features are shown in Fig. 14.

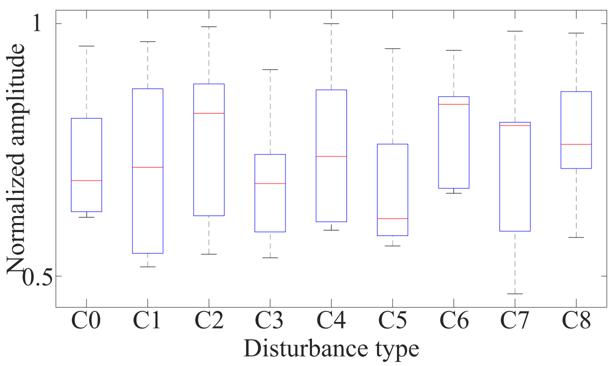
According to Fig.14, the numerical distribution of different signals under feature F3 is dispersed, and the degree of intersection between different types is small. Therefore, feature F3 has good classification ability. The numerical distribution of different signals under feature F12 is concentrated, so that the classification ability of feature F12 is poor.

The classification ability of single feature is limited, so that different feature combinations are needed to recognize PQD signals effectively. In order to analysis the classification ability of different feature combinations, 15 features contained in the optimal feature set are extracted from 9 kinds of PQ signals and shown in Fig. 15. The signal number of each class is 600, and the SNR of features is distributed randomly between 20 dB and 50 dB.

Fig. 15 shows that the different combinations of features in the optimal feature subset have excellent classification ability for different disturbance signals. The combination of feature F3, F5 and F10 has a good preliminary recognition effect on nine kinds of disturbance signals. For disturbance signals C0 and C5, C3 and C7 that cannot be separated by feature F3, F5 and F10, they can be further identified by feature F4 and F9, F1 and F6 respectively. Furthermore, the features that not involved in the Fig. 14 are not useless, they also play a role in the details of actual disturbance recognition. This proves that the optimal feature subset determined by the new methods is effective.



(a)



(b)

FIGURE 14. Classification ability analysis of the highest Gini importance (F3) and the lowest Gini importance (F12), (a) the numerical distribution of feature F3 and (b) the numerical distribution of feature F3.

At present, a large number of power quality data are collected by mass sampling equipment, and the sampling rate of signals is getting higher and higher. In the process of signal analysis, the amount of data is large and the task of analysis is heavy. Therefore, the main difficulty of online implementation of power quality disturbance identification method is the identification efficiency. The online implementation of disturbance identification method requires that the type of the signals can be accurately identified on the

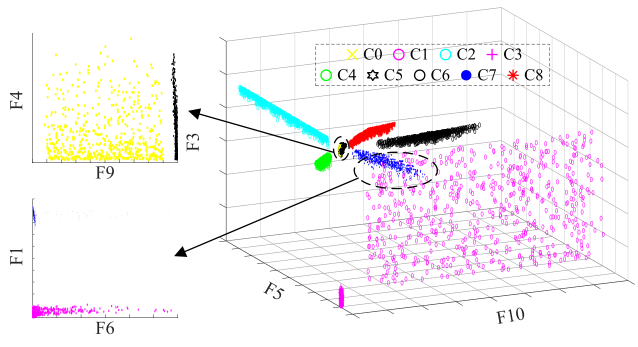


FIGURE 15. Classification ability analysis of different feature combinations.

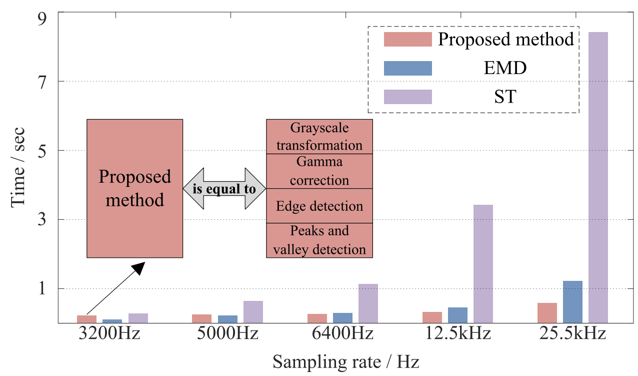


FIGURE 16. Signal processing time of each method at different sampling rates.

premise of high recognition efficiency. The proposed method has determined the optimal feature subset through effective feature selection, and the classifier model has been trained by the optimal feature subset. Therefore, the work of online implementation of identifying unknown disturbance signals includes three process: signal processing, feature extraction and pattern recognition (without feature selection and model training steps). The time of the above three process is the total time of online unknown signal recognition.

In order to show the advantages of the proposed method in signal processing efficiency, the proposed method, ST method and EMD method are used to process a group of disturbance signals at different sampling rates. Fig. 16 shows the signal processing time of each method at different sampling rates. The signal processing time of the proposed method is the sum of gray transformation time, Gamma corrections time, Edge detection time and Peaks and Valley detection time. The computer memory used in the experiment is 16GB, with Intel core i5 processor.

Fig. 16 shows that when the sampling rate is 3200 Hz, the total time of gray transformation and three image enhancement methods of the proposed method is slightly higher than that of EMD method. However, the signal processing time of EMD method and ST method increases obviously with the increase of signal sampling rate, and the advantages of the proposed method in signal processing efficiency become more obvious. Especially at high sampling rates of 12.5 kHz

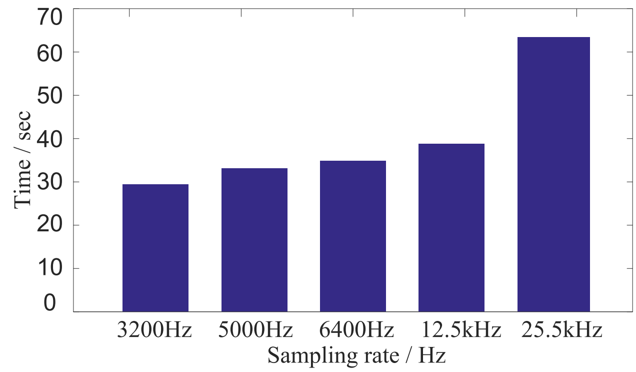


FIGURE 17. Total recognition time of new methods at different sampling rates.

and 25.5 kHz, the signal processing time of the proposed method is much lower than that of ST method. The experimental results show that the new method has high signal processing efficiency, so it is more suitable for the analysis of high sampling rate disturbance signals.

In order to analyze the whole recognition time of the proposed method, the total time of signal processing, feature extraction and pattern recognition at different sampling rates is calculated. The number of disturbance signals analyzed in the experiment is 100 groups. The experimental results are shown in Fig. 17.

Figure 17 shows that at the sampling rate is 12.5 kHz or less, the total time for the proposed method to identify 100 sets of disturbance signals is less than 40 seconds. At the high sampling rate of 25.5 kHz, the total recognition time of the proposed method can also be maintained at about 60 seconds. It is worth mentioning that due to the limited experimental environment, the computer memory used in the experiment is 16GB, with Intel core i5 processor. The proposed method can achieve satisfactory disturbance recognition efficiency in this experimental environment, while the computer used in practical engineering online applications usually has large memory and high processing efficiency. Therefore, we have reason to believe that the efficiency of disturbance identification can be further improved in practical engineering applications.

The above analysis proves that the proposed method has a high disturbance recognition efficiency, and the it has more obvious advantages in signal processing efficiency under high signal sampling rate. Therefore, the new method can meet the real-time analysis requirement of high sampling rate disturbance signal, so as to realize online implementation in actual engineering environment.

C. CLASSIFIER CONSTRUCTION AND DISTURBANCE RECOGNITION BASED ON RF

The optimal feature subset is used to train the RF classifier for PQD recognition. RF is an integrated classification algorithm, the classification effect of the classifier is related to the number of trees in the forest. The larger the forest scale, the smaller the classification error of RF, and the more

TABLE 3. Comparison of classification accuracy of different classifiers.

SNR	classifier	Classification Accuracy /%									Average accuracy
		C0	C1	C2	C3	C4	C5	C6	C7	C8	
50db	RF	100	100	100	100	100	99.0	100	100	100	**99.89**
	DT	98.0	100	100	100	96.5	97.5	100	100	100	99.11
	SVM	95	96	100	100	95	97.5	97.5	99	100	97.77
	ELM	95.5	95.5	100	99.5	96	100	97.5	99.5	100	98.17
40db	RF	100	100	100	100	100	98.5	100	100	100	**99.83**
	DT	98.5	100	100	100	97	97.5	96.5	100	100	98.83
	SVM	96	97.5	100	100	95	92	94	96	96.5	96.33
	ELM	96.5	97	100	99	95.5	98	95	100	100	97.89
30db	RF	98.5	99.5	100	100	100	97	100	100	100	**99.44**
	DT	95	98.5	100	100	95	96.5	97	100	100	98.00
	SVM	90.5	95.5	96.5	94	95	94	91	99.5	100	95.11
	ELM	98	96.5	99.5	100	96	93	93	96.5	99	96.83
20db	RF	95	96	99.5	99	95.5	94	93	95	99	**96.22**
	DT	93	94.5	94	95.5	91.5	94.5	92	94.5	97.5	94.11
	SVM	90	89.5	94	90	88	94	89.5	95	96.5	91.83
	ELM	89.5	91.5	92.5	92	89.5	92	89.5	96	97.5	92.22

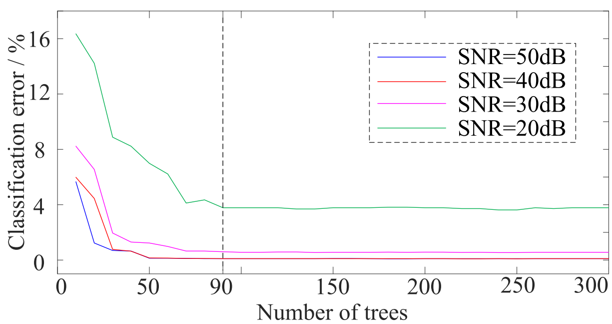


FIGURE 18. The influence of RF scale on classification error.

accurate the Gini importance analysis of features. Therefore, in the feature selection stage, the number of trees in RF is set at 200. However, the classification efficiency of RF will decrease with the increase of the number of trees. Therefore, on the premise of ensuring the optimal classification accuracy, the number of trees should be minimized to improve the classification efficiency of RF. The classification error of RF varies with the number of trees under different noise environments are shown in Fig.18.

Fig. 18 shows that when the number of trees exceeds 90, the recognition accuracy of RF under different noise environments achieve the stable level. The contribution of increasing the number of trees to the classification effect is not obvious, and the classification efficiency will decrease significantly with the increase of the number of trees. Therefore, considering the classification efficiency and classification accuracy, the number of trees in RF is set at 90. While ensuring the good classification effect, the RF still has high classification efficiency.

In order to verify the effectiveness of the new method under different noise environments, 200 samples of each type are generated by simulation to structure the test sets. The SNR of the signals are 50 dB, 40 dB, 30 dB and 20 dB respectively (specific SNR). The classification accuracy of the new method is compared to SVM, [19] ELM and DT, and the parameters setting of the comparison methods are from [19]. The result of the experiment is shown in Table 3.

Table 3 shows the accuracy of identifying nine kinds of disturbance signals by RF method, DT method, SVM method and ELM method under different noise environments. The experimental results show that the recognition accuracy of RF is 99.89% under the SNR of 50 dB, which is 0.78%, 2.12% and 1.72% higher than DT, SVM and ELM respectively. The recognition accuracy of RF is 99.83% under the SNR of 40 dB, which is 3.5% and 1.94% higher than SVM and ELM respectively. The recognition accuracy of RF is 99.44% under the SNR of 30 dB, which is 1.44%, 4.33% and 2.61% higher than DT, SVM and ELM respectively. Under the SNR of 20dB, the recognition accuracy of each classifier decreases because of the great influence of noise. However, the recognition accuracy of RF reaches 96.22%, which is 2.11%, 4.39% and 4% higher than DT, SVM and ELM respectively. The experimental results show that RF achieves the highest recognition accuracy under different noise environments. With the increase of noise, the advantages of RF method are more obvious than other three methods. This proves that RF has excellent disturbance recognition performance and anti-noise ability.

In conclusion, the RF classifier used in the proposed method has good disturbance recognition performance and anti-noise ability, so it can realize high-precision recognition of disturbances under different noise environments. In addition, the good classification stability of RF is conducive to the determination of the optimal feature subset in feature selection. Therefore, RF can be used as an effective tool for power quality disturbance identification.

V. CONCLUSIONS

In this paper, in order to improve the efficiency of signal processing in PQD recognition and remove redundant features from the original feature set, a new PQD recognition method based on image enhancement techniques and feature selection is proposed. The main contributions of the new method include:

1) The original power quality signals are converted into gray image, and three image enhancement techniques include gamma correction edge detection and peaks and valley detection are used to enhanced disturbance features in gray images. On this basis, the disturbance features are extracted and the original feature set is constructed. The new method can extract disturbance features effectively, and has higher signal processing efficiency compared with ST and EMD methods.

2) The classification ability of all features in original feature set are measured by the Gini importance, and the SFS method is used to determine the optimal feature subset. The classification ability of different features is analyzed effectively by the new method, and the redundant features in the original feature set are removed. Therefore, the computational efficiency of features is improved and the structure of the classifier is simplified.

3) RF classifier is used to recognize PQD signals under different noise environments, and the effectiveness of the new method is proved by simulation and comparison experiment.

REFERENCES

[1] S. De and S. Debnath, "Real-time cross-correlation-based technique for detection and classification of power quality disturbances," *IET Gener. Transmiss. Distrib.*, vol. 12, no. 3, pp. 688–695, 2017.

[2] U. Singh and S. N. Singh, "Detection and classification of power quality disturbances based on time–frequency-scale transform," *IET Sci. Meas. Technol.*, vol. 11, pp. 802–810, Sep. 2017.

[3] U. Singh and S. N. Singh, "Application of fractional Fourier transform for classification of power quality disturbances," *IET Sci. Meas. Technol.*, vol. 11, no. 1, pp. 67–76, 2017.

[4] F. A. S. Borges, R. A. S. Fernandes, I. N. Silva, and C. B. S. Silva, "Feature extraction and power quality disturbances classification using smart meters signals," *IEEE Trans. Ind. Informat.*, vol. 12, no. 2, pp. 824–833, Apr. 2017.

[5] R. Kumar, B. Singh, and D. T. Shahani, "Symmetrical components-based modified technique for power-quality disturbances detection and classification," *IEEE Trans. Ind. Appl.*, vol. 52, no. 4, pp. 3443–3450, Jul./Aug. 2016.

[6] Y. Baiqiang, H. E. Yigang, and Z. Yanqing, "Detection and classification of power quality multi-disturbances based on generalized S-transform and fuzzy SOM neural network," *Proc. CSEE*, vol. 35, no. 4, pp. 866–872, 2015.

[7] C. A. Naik and P. Kundu, "Power quality disturbance classification employing S-transform and three-module artificial neural network," *Int. Trans. Electr. Energy Syst.*, vol. 24, no. 9, pp. 1301–1322, 2015.

[8] O. P. Mahela and A. G. Shaik, "Recognition of power quality disturbances using S-transform based ruled decision tree and fuzzy C-means clustering classifiers," *Appl. Soft Comput.*, vol. 59, pp. 243–257, Oct. 2017.

[9] W. Qinghao, W. Tianshi, S. Jidong, Z. Xinwei, and S. Guodong, "The power quality disturbance detection and classification method based on EMD," in *Proc. IEEE PES Innov. Smart Grid Technol.*, May 2012, pp. 1–3.

[10] Z. Liu, Y. Cui, and W. Li, "A classification method for complex power quality disturbances using EEMD and rank wavelet SVM," *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 1678–1685, Jul. 2017.

[11] D. De Yong, S. Bhowmik, and F. Magnago, "An effective power quality classifier using wavelet transform and support vector machines," *Expert Syst. Appl.*, vol. 42, pp. 6075–6081, Sep. 2015.

[12] H. Shareef, A. Mohamed, and A. A. Ibrahim, "An image processing based method for power quality event identification," *Int. J. Elect. Power Energy Syst.*, vol. 46, pp. 184–197, Mar. 2013.

[13] J. Hui *et al.*, "An image processing based method for transient power quality classification," *Power Syst. Protection Control*, vol. 43, no. 13, pp. 72–78, 2015.

[14] U. Singh and S. N. Singh, "Optimal feature selection via NSGA-II for power quality disturbances classification," *IEEE Trans. Ind. Informat.*, vol. 14, no. 7, pp. 2994–3002, Jul. 2017.

[15] N. Huang, D. Xu, X. Liu, and L. Liu, "Power quality disturbances classification based on S-transform and probabilistic neural network," *Neurocomputing*, vol. 98, no. S1, pp. 12–23, 2012.

[16] P. K. Ray, S. R. Mohanty, N. Kishor, and J. P. S. Catalão, "Optimal feature and decision tree-based classification of power quality disturbances in distributed generation systems," *IEEE Trans. Sustain. Energy*, vol. 5, no. 1, pp. 200–208, Jan. 2014.

[17] 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 Trans. Ind. Appl.*, vol. 51, no. 2, pp. 1249–1258, Mar./Apr. 2015.

[18] 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 Trans. Instrum. Meas.*, vol. 65, no. 10, pp. 2302–2312, Oct. 2016.

[19] S. Zhang, P. Li, L. Zhang, H. Li, W. Jiang, and Y. Hu, "Modified S transform and ELM algorithms and their applications in power quality analysis," *Neurocomputing*, vol. 185, pp. 231–241, Apr. 2016.

[20] 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 Trans. Smart Grid*, vol. 9, no. 4, pp. 3122–3132, Jul. 2016.

[21] A. Cutler, D. R. Cutler, and J. R. Stevens, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 157–176, 2011.

[22] W. Shang, H. Huang, Y. Liu, Y. Lin, Y. Qu, and H. Dong, "Research on the algorithm of feature selection based on Gini index for text categorization," *J. Comput. Res. Develop.*, vol. 43, no. 10, pp. 1688–1694, 2006.

[23] R. I. Lerman and S. Yitzhaki, "A note on the calculation and interpretation of the Gini index," *Econ. Lett.*, vol. 15, nos. 3–4, pp. 363–368, 1984.

[24] J. Wu, J. Xin, and N. Zheng, "SVM learning from imbalanced microaneurysm candidate datasets used feature selection by Gini index," in *Proc. IEEE Int. Conf. Inf. Automat. (ICIA)*, Aug. 2015, pp. 1637–1641.

[25] H. Liu, M. Zhou, X. S. Lu, and C. Yao, "Weighted Gini index feature selection method for imbalanced data," in *Proc. IEEE 15th Int. Conf. Netw., Sens. Control (ICNSC)*, Mar. 2018, pp. 1–6.

[26] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, vol. 28, no. 4. Upper Saddle River, NJ, USA: Prentice-Hall, 2008, pp. 484–486.

[27] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-8, no. 6, pp. 679–698, Nov. 1986.

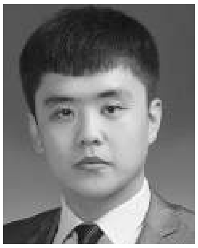
[28] P. Salembier and J. Serra, "Flat zones filtering, connected operators, and filters by reconstruction," *IEEE Trans. Image Process.*, vol. 4, no. 8, pp. 1153–1160, Aug. 1995.

[29] H. Nantian *et al.*, "Feature selection and optimal decision tree construction of complex power quality disturbances," *Proc. CSEE*, vol. 37, no. 3, pp. 776–785, 2017.

[30] 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, Jan. 2017.

[31] C.-Y. Lee and Y.-X. Shen, "Optimal feature selection for power-quality disturbances classification," *IEEE Trans. Power Del.*, vol. 26, no. 4, pp. 2342–2351, Oct. 2011.

LIN LIN received the B.Sc. degree from Lanzhou Jiaotong University, in 2003, and the M.Sc. degree from the Dalian University of Technology, in 2008. She is currently an Associate Professor with the Jilin Institute of Chemical Technology, China. Her research interests include power quality analysis and control, the fault diagnosis of power apparatus, and data mining in smart grids.



DA WANG received the B.Sc. degree from Shandong Agricultural University, in 2016, and the M.Sc. degree from Northeast Electric Power University, in 2019. He is currently with Dezhou Power Supply Company, State Grid Shandong Electric Power Co., Ltd. His research interest includes power quality disturbance analysis.

SHUYE ZHAO received the B.Sc. degree from the Harbin University of Science and Technology, in 2012, and the M.Sc. degree from the Harbin Institute of Technology, in 2014. He is currently with the Economic and Technical Research Institute, State Grid Inner Mongolia Eastern Electric Power Co., Ltd. His research interest includes power system planning and design.

LINGLING CHEN received the B.Sc. degree from Jilin University, in 2004, the M.Sc. degree from Yanbian University, in 2011, and the Ph.D. degree from Jilin University, in 2015. She is currently an Associate Professor with the Jilin Institute of Chemical Technology, China. Her research interests include power quality analysis and control, the fault diagnosis of power apparatus, and data mining in smart grids.

NANTIAN HUANG received the B.Sc. degree from the Harbin Institute of Technology, in 2003, the M.Sc. degree from the Dalian University of Technology, in 2006, and the Ph.D. degree from the Harbin Institute of Technology, in 2012. He is currently an Associate Professor and a M.Sc. supervisor with Northeast Electric Power University. His research interests include power quality analysis and control, condition monitoring, and the fault diagnosis of power apparatus.