Journal of Information Hiding and Multimedia Signal Processing ©2017 ISSN 2073-4212
Ubiquitous International Volume **8**, Number **4**, July **2017**

The Classification of Multiple Power Quality Disturbances based on Preliminary Classification and Improved Support Vector Machine

Liquan Zhao[1], Yan Long[1], Lin Wang[2]

[1]College of Information Engineering
Northeast Electric Power University
NO.169 Changchun Road,Chuanying District, Jilin City, China
zhao_liquan@163.com
[2]Telecommunication Branch
Jilin Power Supply Company
NO.124 Jilin Street, Changyi District, Jilin City, China

Received August, 2016; revised March, 2017

Abstract. *The classification of multiple power quality disturbances composes two parts, which are feature extraction and classification. To improve the accuracy rate of classification, a new classification method for multiple power quality disturbances is proposed based on S transform, wavelet transform and improved support vector machine. First, S transform is used to extract the features of disturbance signals, and a preliminary classification is done by setting a suitable threshold to class the extracted features to classify the disturbance into three groups. Then, the wavelet transformation is used to extract the energy difference of the wavelet coefficients of disturbance signals in each classified group. In every classified group, the support vector machine is used to class the extracted energy difference to further more class the multiple power quality disturbances. In order to improve the accuracy of the overall classification, the Gaussian kernel function of the support vector machine is improved. The aggregation degree of the feature vector at the center of the kernel function affects the number of support vectors, so this paper introduces a radial width factor and an amplitude modulation factor to the kernel function. It solves the problems that exist in the traditional Gaussian kernel function and reduces the number of support vectors as well as the computational complexity involved. When the improved algorithm is applied to the multiple power quality disturbance classification, it is proved to be feasible and to have higher classification accuracy.*
Keywords: Multiple power quality disturbance, Support vector machine, Gaussian kernel function, Classification accuracy.

---

1. **Introduction.** With the wide application of electronic technology, all forms of the nonlinear impact load cause power quality problems in the distribution networks, which will cause major economic losses and social impact. Problems caused by power quality have become a matter of concern [1]. The identification of power quality is the premise underlying the analysis and improvement of power quality. Classifying power quality disturbances is of great significance to solve the problem of power quality [2].

Typical power quality disturbances include voltage swell, sag, interruption, harmonics, impulse transient, oscillation transient and flicker. In practical application, power quality disturbance is a disturbance that occurrs in many forms simultaneously. Including swell+harmonic, sag+harmonic, swell+flicker and sag+flicker and so on. The identification of a multiple power quality disturbance requires two steps: feature extraction and classification. At present, the main methods of feature extraction of power quality disturbances include short time Fourier transform (STFT) [3], wavelet transform [4,5], S transform [6,7] and Hilbert-Huang transform(HHT) [8] and so on. The main methods of classifying the disturbances are: neural network [9], decision tree[10] and support vector machines (SVM) [11,12]. These classification methods are reliable, but the processes of feature extraction and construction of the classifier is complex, training time for a neural network classifier is long, and it is easy to fall into local minimum by using a neural network classifier. Although the decision tree classifier can be accomplished quickly, the rules are complex, and the model is difficult to deal with. Support vector machine is based on statistical learning theory and is a machine learning method used to solve pattern recognition problems. SVM overcomes the disadvantages of the artificial neural network, which easily relapses into local optimal solution and has a long training time, so we use SVM as classifier in this paper.

In this paper, we first use S transform to extract the disturbance feature and classify the disturbance signals into three groups by setting threshold to the extracted feature values. Then, the energy differences that are obtained by wavelet transform are used as input of SVM in every group to realize the final classification. This process reduces the complexity of the classification. To further more improve the performance of classification method, the improved Gaussian kernel function is applied to the SVM to classify the eleven kinds of power quality disturbances. The experimental results show that our proposed method is feasible and effective.

2. **Preliminary classification.**

2.1. **S Transform.** S transform was proposed by Stockwell in 1996 in a solution that may be regarded as both a phase correction of the wavelet transform, and also as the development of the short time Fourier transform. The height and width of the Gaussian window of the S transform vary with frequency, which overcomes the shortcomings of the window height and fixed width of the STFT[13]. Setting the input signal $h(t)$, when the S transform is $S(\tau, f)$,we obtain the following:

$$S(\tau, f) = \int_{-\infty}^{\infty} h(t)g(\tau - t, f) \cdot e^{-2\pi ift}dt \tag{1}$$

$$g(t, f) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{t^2}{2\sigma^2}} \tag{2}$$

where $g(t, f)$ is the Gaussian window, and the parameter τ controls the position of the Gaussian window on the timeline,$\sigma = 1/|f|$. The signal $h(t)$ can be reconstructed by the S transform $S(\tau, f)$,and S inverter transform is:

$$h(t) = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} S(\tau, f)d\tau \right) \cdot e^{2\pi ift}df \tag{3}$$

The $h[kt]$, $k = 0, 1, 2, \cdots, N-1$ is obtained by sampling the continuous time signal $h(t)$, and the total sample points value is N. From the discrete Fourier transform, we obtain the discrete representation of the S transform:

$$S[m, n] = \sum_{k=0}^{N-1} H(n + k) \cdot e^{\frac{-2\pi^2k^2}{n^2}} \cdot e^{\frac{i2\pi km}{N}}, n \neq 0 \tag{4}$$

$$S\left[m,n\right]=\frac{1}{N}\sum_{k=0}^{N-1}h(k),n=0 \tag{5}$$

$$H(n)=\frac{1}{N}\sum_{k=0}^{N-1}h(k)\cdot e^{\frac{-i2\pi kn}{N}} \tag{6}$$

where N is the total number of points.

2.2. **Preliminary classification based on S Transform.** The result of the discrete S transform is a complex matrix, which is called the S matrix. The row of the S matrix corresponds to the time, the column corresponds to the frequency, and the matrix element is the complex value. The row vector demonstrates that the signal amplitude changes with time at a certain frequency, and the column vector demonstrates that the signal amplitude changes with frequency at a certain time.

This paper mainly classifies seven kinds of single disturbances and four kinds of complex disturbances. These disturbances include voltage swell, sag, interruption, harmonics, impulse transient, oscillation transient, flicker, swell+harmonic, sag+harmonic, swell+flicker and sag+flicker. The eleven kinds of power quality disturbance are transformed by the S transform, and the occurrence of the disturbance is mainly reflected in the frequency and amplitude, so the disturbances can be distinguished according to the different frequencies and amplitudes. Voltage swell, sag and interruption are mainly amplitude changes, and voltage harmonics is the main sources of frequency changes.

Due to the different voltage fluctuation causes the different change of voltage amplitude, for example, the amplitude of the voltage swell is higher than normal voltage, while the amplitude of the voltage sag is lower than normal voltage. Therefore, this paper uses the amplitude difference of each disturbance signal as features. We randomly generate 200 disturbance signals for each kind of power quality disturbance according to the disturbance signal model. Each disturbance signal has 1218 sample points. First, we use the maximum amplitude value corresponding to each column of the S matrix to construct a vector. Then we use the maximum element of the constructed vector as one sample value. So we can get 200 samples for each kind of disturbance signal in our experiment. Each sample point corresponds to one of the 200 disturbance signals that are of the same kind of disturbance.

Figure 1 shows the relationship between the disturbance signals and their sample values that are obtained by the above method for eleven kinds of disturbance. From the figure 1, we can see that some kinds of disturbance have the same range of sample values, and some kinds of disturbance have the different range of sample values. Therefore, we use matlab to compute the maximum and minimum sample values of each kind of disturbance signals, to obtain the sample value range of different disturbance signals. Based on the above computation, the sample value range of voltage swell is 1.077 to 1.828, which is assumed to be the first class of disturbance. The sample value range of voltage sag, voltage interruption, harmonics, impulse transient, oscillation transient and flicker is 0.883 to 1.046 , which is assumed to be second class. The sample value range of voltage swell+harmonics, sag+harmonics, voltage swell+flicker, sag+flicker is 1.848 to 2.846 , which is assumed to be third class. Therefore, we can use the three sample value range to preliminarily classify the disturbance signals into three major classes.

After preliminary classification, we use wavelet transform to extract the energy difference of the wavelet coefficients as the feature vector in each classes, and then use SVM classifier to further more class the three classes to realize the final classification.

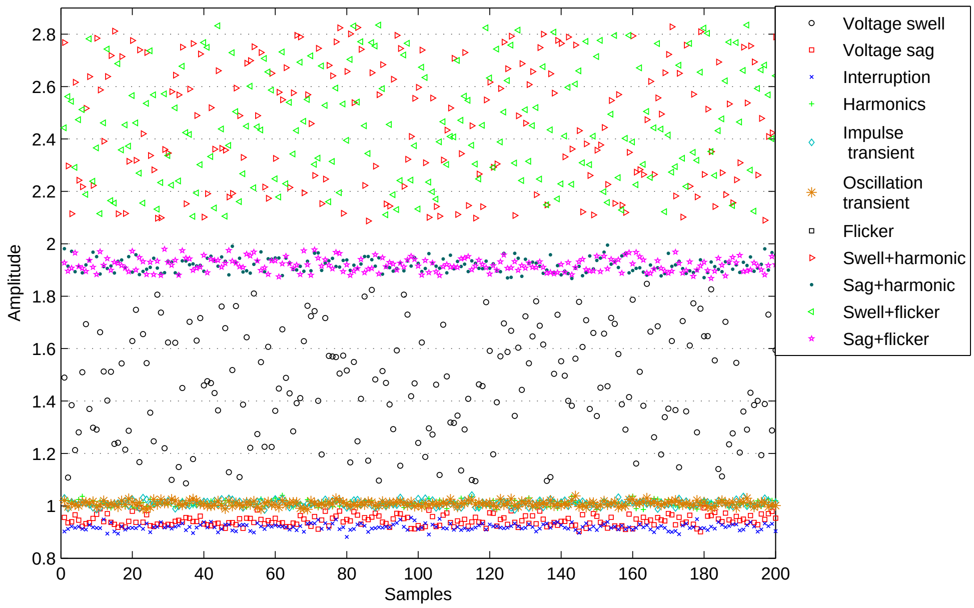

FIGURE 1. Relationship between the disturbance signals and their amplitudes for eleven kinds of disturbance

3. Improved Support Vector Machine.

3.1. Support Vector Machine. Support vector machine is a statistical learning tool based on the Vapnik-Chervonenkis theory and structural risk minimization principles. SVM provides great advantages for solving the problems of nonlinear and multi-dimensional pattern recognition. The problems posed by the nonlinear character of the patterns can be mapped to multi-dimensional space by the kernel function, which can be divided into a linear, separable problem. The separation of the two classes is performed by the optimal hyper plane [14]. Given a set of training data $T = \{x_1, x_2, \cdots, x_l\}, x_i \in R^n$.The fuzzy one-class SVM searches for a crisp hyperplane that separates the image of the target class from the origin:

$$f(x) = \omega_1^{\mathbf{T}} x + b \tag{7}$$

where ω_1 are weights, and b are biases [15]. The hyperplane is considered optimal when the separating margin between the two classes is maximal, and this is achieved by computing:

$$\varphi(\omega_1) = \min \frac{1}{2} ||\omega_1||^2 \tag{8}$$

which is subject to:

$$y_i(\omega_1^{\mathbf{T}} x + b) \geq 1, i = 1, 2, \cdots, N \tag{9}$$

This problem leads to the optimization of the Lagrange function:

$$L(\omega_1, b, \alpha) = \frac{1}{2} ||\omega_1||^2 - \sum_{i=1}^{N} \alpha_i [y_i(\omega_1^{\mathbf{T}} x_i + b) - 1] \tag{10}$$

where α_i is a Lagrange multiplier and x_i corresponding to α_i that is not zero is known as the support vector [16]. If the data is not linearly separable, a slack variable is introduced.

Thus, (8) is converted to (11).

$$\varphi(\omega,\xi) = \min\frac{1}{2}||\omega||^2 + C\left(\sum_{i=1}^{N}\xi_i\right) \tag{11}$$

If the data is not linearly separable, a conversion method is needed to transform the original data space into a multi-dimensional space. These transformations are performed using Kernel functions. To find α_i, it is necessary to solve the following problem:

$$Q(\alpha) = \max\sum_{i=1}^{N}\alpha_i - \frac{1}{2}\sum_{i,j=1}^{N}\alpha_i\alpha_j y_i y_j K(x_i\cdot x_j) \tag{12}$$

which is subject to:

$$\sum_{i=1}^{N} y_i\alpha_i = 0 \tag{13}$$

$$0 \le \alpha_i \le C, i = 1, 2, \cdots, N \tag{14}$$

where $K(x_i \cdot x_j)$ is the kernel function.

Support vector machine is originally designed for binary classification. For solving the problem of multi-class classification of power quality disturbances, we need to transform the multi-classification problem into two classification problems. There are several methods of combining binary classifiers to construct multi-class classifications. The most commonly used multi-class classification methods include the one-against-one method and the one-against-all method. This paper uses the one-against-one method

3.2. **Improvement of the Gaussian kernel function.** Kernel function is the key to changing the feature vector from nonlinear and non separable in a low-dimensional space to linearly separable in a multi-dimensional space. Different kernel functions can cause the SVM to have different performance. Kernel function K is defined as

$$K(x, x_i) =< \varphi(x), \varphi(x_i) > \tag{15}$$

where $x, x_i \in X$, X is the sample set,$\varphi(\cdot)$ is a mapping from the input space to the multi-dimensional feature space, and $<,>$ is the inner product. Commonly used kernel functions include the Gaussian kernel function, the linear kernel function and the polynomial kernel function. The most commonly used kernel function in SVM is the Gaussian kernel function. Its general form is:

$$K(x, x_i) = \exp\left(-\frac{||x - x_i||^2}{2\sigma^2}\right), \sigma > 0 \tag{16}$$

The Gaussian kernel function based on Euclidean distance equations $K(x, x_i) = ||x-x_i||$ has the problem that the feature vectors that distribute densely in the original space will become very sparse, when are mapped to the high-dimensional space. This will increase the number of support vectors which lead to increase the computation complexity and consume much more time to realize classification. From (16), it can see that the bandwidth σ is only one of the parameters of the Gaussian kernel function. By adjusting the values of σ , the classification performance of the SVM and the clustering ability of samples in the high-dimensional space can be changed. However, adjusting the parameters only will have little effect on the classification performance of the SVM and will not solve the problem that exists in the Gaussian kernel function.

To solve this problem, the kernel function must have a characteristic that allows the rate of decay near the support vector to be very fast. In order to solve this problem, this paper proposes a way to improve the Gaussian kernel function. An amplitude modulation parameter is introduced into the Gaussian kernel function. It is expressed as follows

$$K(x, x_i) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{||x - x_i||^2}{2\sigma^2}\right) \tag{17}$$

The performance of the function curve can make the sample data in the vicinity of the support vector have a faster decay rate.

In order to make the feature vectors more gather in the high-dimension space, a radial width adjustment parameter is introduced. This make the kernel function have a faster decay rate. The improved kernel function can be expressed as

$$K(x, x_i) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{c||x - x_i||^2}{2\sigma^2}\right) \tag{18}$$

Figure 2 shows the function curves are obtained by adjusting parameters c with a fixed σ. As can be seen from the figure 2, the faster the corresponding curve decay, and the more aggregated the data. Thus, the decay rate of the kernel function and the degree of aggregation of the data can be changed by controlling the parameter c.

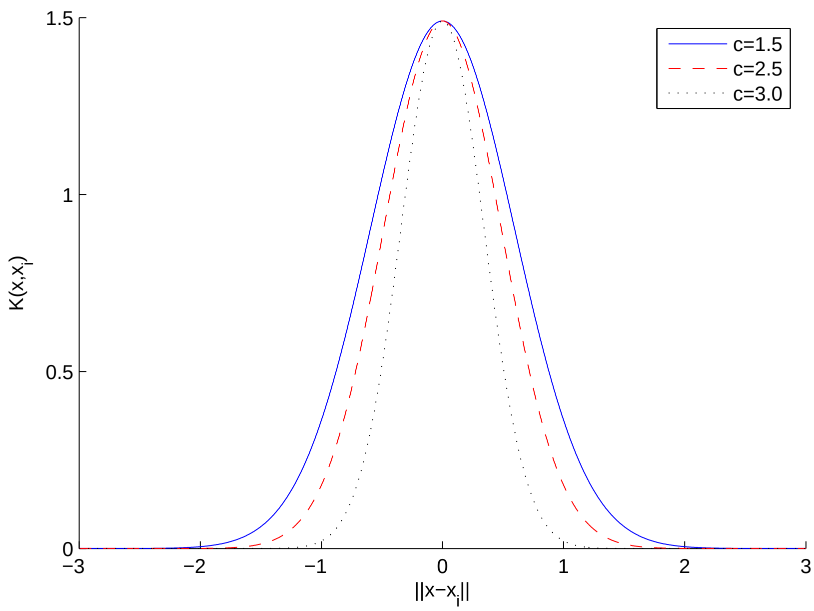

FIGURE 2. Kernel function curves with different c values

The figure 3 shows the original and improved kernel functions. The value of σ is 0.3535, and the value c of the improved Gaussian kernel function is 2.7. From figure 3, we can see that the improved Gaussian kernel function is more clustered near the support vector, and the decay rate near the support vector is faster than the Gaussian kernel function. By changing the value of c , the rate of decay of the kernel function can be changed. The greater the value of c , the faster of decay rate.

4. **Simulations and Analysis.** The classification of power quality disturbances is divided into two steps: feature extraction and classification. First of all, the feature is extracted by the ST of the disturbance, and the threshold is set to make a preliminary classification, which reduces the computation of the classification. Then, the power quality disturbance signals are transformed by wavelet transform. The Daubechies db4 wavelet

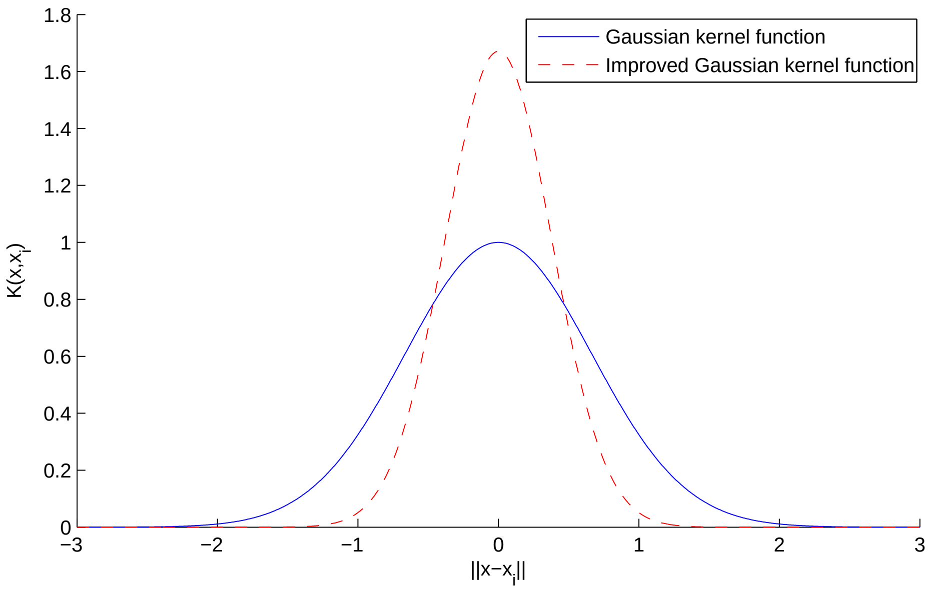

FIGURE 3. The comparison of the improved Gaussian kernel function and the original Gaussian kernel function

function is adopted to decompose the 8 scale wavelet, and the wavelet coefficients of each layer after decomposition are extracted. According to the Parseval theorem, the energy of each dimension is obtained. Then, the difference between the energy of the signal and the normal signal is obtained. The energy difference is normalized to be used as the feature vector. Finally, the improved SVM is used to classify the disturbance signal of the power quality.

We use power quality signals that contain eleven kinds of power quality disturbance signals that have been mentioned in the section 2.2 as test signals, to verify the effectiveness of the proposed classification. The$\sigma = 0.595$ and the penalty parameter $C = 1$ is obtained by cross validation and grid search. The SNR is 15dB. The classification method proposed in this paper is compared with the Gaussian kernel function classification method. The c value is determined after several tests. Table 1 shows the average classification accuracy and the total number of support vectors with the different c values. As can be seen from Table 1, when the value of c is greater than 2.7 or less than 2.7, the classification results of the program are more than the original results, but the results are lower than the results obtained when c=2.7. This is because as the value of c becomes larger, the improved Gaussian kernel function near the support vector will be gathered, and the decay rate will be faster. But when the value of c is greater than 2.7, the modified Gaussian kernel function becomes too clustered near the support vector, and the overall classification accuracy of the model is reduced, so the optimal value of c in this paper is 2.7.

Table 2 shows the classification accuracy of power quality disturbances using different methods.Among them, the training samples of each disturbance are 80 and the test samples are 120. From Table 2, we can see that the classification accuracy is improved for voltage sag,impulse transient, flicker, swell+harmonic, swell+flicker, sag+flicker using the preliminary classification +improved SVM method compared with the original SVM method. The classification accuracy is reduced for voltage swell and voltage interruption using the preliminary classification +improved SVM method compared with the original SVM method. The classification accuracy obtained from the preliminary classification

Table 1. Comparison of the average classification accuracy and the total number of support vectors with the different c values

	c=2.5	c=2.6	c=2.7	c=2.8	c=3.2
Average classification accuracy (%)	94.2424	94.5455	94.5455	92.44697	94.3939
The total number of support vectors	330	330	328	328	329

+improved SVM method for most of disturbance kinds is equal or lager than the other two methods. The average classification accuracy of preliminary classification +improved SVM method is the largest, followed by the improved SVM method and the original SVM method.

The average classification of primary support vector machine accuracy rate is 94.1667%, the accuracy rate has been very high, so this paper uses the improved support vector machine to improve the overall classification accuracy as far as possible.The average classification accuracy is improved by 0.4% using the improved support vector machine, Among them, the voltage sag, impulse transient, swell+ harmonic, Swell+flicker, Sag+flicker are increased by 0.84%, 0.8%, 2.5%, 2%, 0.84%, respectively. In order to improve the classification accuracy, this paper proposes a preliminary classification and improved SVM method to classify the disturbance signal. The classification results show that, although the classification accuracy rate of individual disturbance is reduced, but the overall classification accuracy increased by 1.02%. Among them, the voltage sag, impulse transient, flicker, swell+ harmonic, Swell+flicker, Sag+flicker are increased by 10.9%, 0.8%, 5%, 1.67%, 7.5%, 0.84%, respectively.Therefore, these results demonstrate that the method proposed in this paper has significant accuracy in the classification of power quality disturbances.

Table 2. Comparison of the classification accuracy of various disturbance signals

Disturbance type	Classification accuracy(%)		
	SVM	Improved SVM	Preliminary classification +improved SVM
Voltage swell	100	100	90
Voltage sag	89.1667	90	100
Voltage interruption	89.1667	86.6667	86.8333
Harmonic	100	100	100
Impulse transient	97.5	98.3333	98.3333
Oscillation transient	100	100	100
Flicker	95	94.1667	100
Swell+harmonic	90	92.5	92.6667
Sag+harmonic	91.6667	91.6667	91.6667
Swell+flicker	89.1667	91.6667	96.6667
Sag+flicker	94.1667	95	95
Average classification accuracy(%)	94.1667	94.5455	95.5606

Table 3 shows a comparison of the total number of support vectors, running time and the average classification accuracy for original SVM method, improved SVM method and preliminary classification +improved SVM method. Table 3 shows that by using the preliminary classification +improved SVM method, the total number of support vector is reduced, which makes the calculation simple and improves the average classification accuracy. Meanwhile, the classification time is reduced. The comparison results show that the method that uses the preliminary classification +improved SVM method can improve the average classification accuracy of power quality disturbances and have faster classification speed.

Table 3. Comparison of different methods for the number of support vector, running time and average classification accuracy

	The total number of support vector	Running time(s)	Average classification accuracy (%)
SVM	344	0.0918	94.1667
Improved SVM	328	0.0651	94.5455
Preliminary classification +improved SVM	217	0.0311	95.5606

5. **Conclusions.** In this paper, a new classification method based on preliminary classification and improved SVM method is proposed. The new method firstly uses S transform to obtain the threshold, and uses the threshold to classify disturbance signals into three groups. Then proposes an improved kernel function as SVM kernel function to improve its performance. Comparing with other methods, the new method has higher average accuracy and smaller running time.

REFERENCES

[1] R. Ahila , V. Sadasivam , and K. Manimala, An integrated PSO for parameter determination and feature selection of ELM and its application in classification of power system disturbances, *Applied Soft Computing*, vol.32, pp.23–37, 2015.

[2] A. Abdoos, P. K. Mianaei, and M. R. Ghadikolaei, Combined VMD-SVM based feature selection method for classification of power quality events, *Applied Soft Computing*, vol.38, pp.637–646, 2016.

[3] H. Xie, J. Lin, Y. Lei, and Y. Liao, Fast-varying AMCFM components extraction based on an adaptive STFT, *Digital Signal Process*, vol.22, no.4, pp.664–670, 2012.

[4] M. Masoum, S. Jamali, and N. Ghaffarzadeh, Detection and classification of power quality disturbances using discrete wavelet transform and wavelet networks, *Iet Science Measurement & Technology*, vol.4, no.4, pp.193–205, 2010.

[5] W. G. Morsi, C. P. Diduch, L. Chang, and M. E. El-Hawary, Wavelet-based reactive power and energy measurement in the presence of power quality disturbances, *IEEE Transactions on Power Systems*, vol.26, no.3, pp.1263–1271, 2011.

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

[7] P. K. Dash, B. K. Panigrahi,and G. Panda, Power quality analysis using S-transform, *IEEE Transactions on Power Delivery*, vol.18, no.2, pp.406–411, 2003.

[8] W. Tao , S. Yin , and M. Ding, Classification of power quality disturbance signals based on S-transform and HHT *Control Conference*, Xi'an, pp.3639–3644, 2013.

[9] C. Naik , and P. Kundu, Power quality disturbance classification employing S-transform and three-module artificial neural network, *International Transactions on Electrical Energy Systems*, vol.24, no.9, pp.1301–1322, 2014.

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

[11] H. Nantian, G. Dianguoand L. Xiaoshen, Identification of Power Quality Complex Disturbances Based on S-Transform and SVM, *Transactions of China Electrotechnical Society* , vol.26, no.10, pp.23–30, 2011.

[12] D. D. Yong, S. Bhowmik, and F. Magnago, An effective Power Quality classifier using Wavelet Transform and Support Vector Machines, *Expert Systems with Applications*, vol.42, no.15-16, pp.6075–6081, 2015.

[13] A. Rodrguez, J. A. Aguado , and F. Martn, Rule-based classification of power quality disturbances using S-transform, *Electric Power Systems Research*, vol.86, no.4, pp.113–121, 2012.

[14] G. Liu, F. Li, and G. Wen, Classification of power quality disturbances based on independent component analysis and support vector machine, *2013 International Conference on Wavelet Analysis and Pattern Recognition*, Tianjin, pp.115–123, 2013.

[15] J. Zhang, K. Wang, and W. Zhu, Least Squares Fuzzy One-class Support Vector Machine for Imbalanced Data, *International Journal of Signal Processing, Image Processing and Pattern Recognition*, vol.8, no.8, pp.299–308, 2015.

[16] A. A. Abdoos, P. K. Mianaei, and M. R. Ghadikolaei, Combined VMD-SVM based feature selection method for classification of power quality events, *Applied Soft Computing*, vol.38, pp.637–646, 2016.