Applied Soft Computing 38 (2016) 637–646

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

Combined VMD-SVM based feature selection method for classification of power quality events

Ali Akbar Abdoos*, Peyman Khorshidian Mianaei, Mostafa Rayatpanah Ghadikolaei

Department of Electrical and Computer Engineering, Babol Noshirvani University of Technology, Babol, Mazandaran, Iran

ARTICLE INFO

Article history:
Received 25 March 2015
Received in revised form
24 September 2015
Accepted 18 October 2015
Available online 28 October 2015

Keywords:
Feature selection
Pattern recognition
Support vector machines
Signal analysis
S-transform
Variational mode decomposition

ABSTRACT

Power quality (PQ) issues have become more important than before due to increased use of sensitive electrical loads. In this paper, a new hybrid algorithm is presented for PQ disturbances detection in electrical power systems. The proposed method is constructed based on four main steps: simulation of PQ events, extraction of features, selection of dominant features, and classification of selected features. By using two powerful signal processing tools, i.e. variational mode decomposition (VMD) and S-transform (ST), some potential features are extracted from different PQ events. VMD as a new tool decomposes signals into different modes and ST also analyzes signals in both time and frequency domains. In order to avoid large dimension of feature vector and obtain a detection scheme with optimum structure, sequential forward selection (SFS) and sequential backward selection (SBS) as wrapper based methods and Gram–Schmidt orthogonalization (GSO) based feature selection method as filter based method are used for elimination of redundant features. In the next step, PQ events are discriminated by support vector machines (SVMs) as classifier core. Obtained results of the extensive tests prove the satisfactory performance of the proposed method in terms of speed and accuracy even in noisy conditions. Moreover, the start and end points of PQ events can be detected with high precision.

1. Introduction

The power quality (PQ) problem can be interpreted as voltage quality. In nowadays-electrical networks, the voltage waveforms are more distorted due to augmented use of nonlinear loads like solid-state switching devices, power electronically switched loads, computers and data processing equipment, industrial plant rectifiers and inverters. Distorted voltage signals can affect the proper operation of these sensitive electronic devices. Therefore, PQ monitoring is an essential task to evaluate the performance of power systems. Before any corrective solution, the type of PQ disturbances should be determined by an accurate detection scheme [1].

Different solutions based on pattern recognition methods have been proposed for detection of PQ events [2]. Most of them are realized through two main steps: feature extraction and feature classification. Different signal analysis tools have been used for extraction of dominant features. Some of them analyze signals in time domain like empirical mode decomposition (EMD) [3,4] and some analyze signals in frequency domain such as Fourier transform (FT) [5,6] and some like wavelet transform (WT) [7–11], S-transform (ST) [12–16] and Hilbert transform [17,18] analyze wave shapes in both time and frequency domains, simultaneously.

The main drawback of EMD is that the number of intrinsic mode functions (IMFs) changes according to waveform of PQ events. Moreover, it provides no information about frequency contents of signals. WT decomposes signals into details and approximation levels in consecutive steps using special filter called mother wavelet. Extracted features from approximation and detail levels can be used for classification [19,20]. WT based methods suffer from two disadvantages: firstly the different types of mother wavelet and different decomposition levels should be scrutinized to obtain the best performance and secondly extracted features from details levels are sensitive to noise. ST as another signal processing tool provides a complete visualization of signal in both time and frequency domains, simultaneously. The output of ST is a matrix whose rows and columns pertain to frequency and time, respectively. Each element of the matrix is a complex number, thus the magnitude and phase of each frequency component is determined for each time instant.

In order to classify extracted features, different classifiers have been implemented such as: feed forward neural network (FFNN) [3,5,15,17], decision tree (DT) [14], support vector machines (SVM) [7,8,11] and probabilistic neural network (PNN) [3]. Neural networks (NNs) are well known for their learning ability but

* Corresponding author. Tel.: +98 9125314761; fax: +98 1132310977.
E-mail address: abdoos_a@yahoo.com (A.A. Abdoos).

http://dx.doi.org/10.1016/j.asoc.2015.10.038
1568-4946/

its learning process is a very time consuming task and there is no definite rule for optimum setting of NNs' structures. PNN needs a large number of exemplar patterns to yield acceptable classification accuracy. This obstacle leads to slow execution speed and large memory requirements.

In spite of large number of research works in the field of PQ events detection, there is still a lack of analysis of extracted features effects on classifiers detection accuracy. On the other hand large dimension of extracted features may mislead the classifier which results in the reduction of detection accuracy. So, redundant features should be eliminated from extracted features. Moreover, most of the proposed method are not able to detect the interval time in which the signal is distorted [3–17].

In this study, a relatively large dimension feature vectors are extracted using ST [21] and VMD [22], and more useful features are selected applying several feature selection methods based on filter and wrapper. Wrapper based method are very time consuming due to their huge computational burden while filters are faster since they rank features based on intrinsic attributes. Sequential forward selection (SFS) [23] and sequential backward selection (SBS) [24] as wrapper based methods and Gram–Schmidt orthogonalization (GSO) based feature selection [25] as a filter based method are used to eliminate redundant features. In the last step, dominant selected features are discriminated by support vector machines (SVMs) classifier [26,27]. The proposed detection scheme has advantages from the following aspects:

- VMD and ST have few tuning parameters as compared to WT which has many mother wavelet filters and decomposition levels.
- SVM classifier with simple structure and only few adjustable parameters has better performance as compared to FFNN.
- Elimination of redundant features by using feature selection methods, augments the generalization capability and detection accuracy of the proposed method.
- Results have been presented for both filter and wrapper based feature selection methods.
- The start and end points of events can be detected by using mode 2 of VMD analysis.
- The proposed algorithm is robust in noisy conditions.

2. Required tools

2.1. Variational mode decomposition

The VMD is a signal processing technique that decomposes a real valued signal into different levels named modes u_k, which have specific sparsity properties while producing main signal. It is assumed that each mode k to be concentrated around a center pulsation ω_k determined during the decomposition process. Thus, the sparsity of each mode is chosen to be its bandwidth in spectral domain. In order to obtain the mode bandwidth, the following steps should be implemented: (1) applying Hilbert transform to each mode u_k in order to obtain unilateral frequency spectrum. (2) Shifting the mode's frequency spectrum to "baseband", by using an exponential tuned to the respective estimated center frequency. (3) Estimation of the bandwidth through the H^1 Gaussian smoothness of the demodulated signal, i.e. the squared L^2 – norm of the gradient. Thus, the decomposition process is realized by solving the following optimization problem [22]:

$$\begin{aligned} &\min \quad \left\{ \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-i\omega_k t} \right\|_2^2 \right\} \\ &\text{s.t.} \quad \sum_k u_k = f(t) \end{aligned} \tag{1}$$

where $f(t)$ is the main signal to be decomposed, $\{u_k\} = \{u_1, \ldots, u_K\}$ and $\{\omega_k\} = \{\omega_1, \ldots, \omega_K\}$ implicates the set of all modes and their center frequencies, respectively. $\delta(t)$ is the Dirac distribution and * denotes convolution. In order to address the constraint, both penalty term and Lagrangian multipliers λ are considered. The combination of the two terms benefits both from the nice convergence properties of the quadratic penalty at finite weight, and the strict enforcement of the constraint by the Lagrangian multiplier. So, the above optimization problem is changed to unconstraint one as below [22]:

$$\begin{aligned} \mathcal{L}(\{\boldsymbol{u_k}\}, \{\boldsymbol{\omega_k}\}, \boldsymbol{\lambda}) = \alpha \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-i\omega_k t} \right\|_2^2 \\ + \left\| f(t) - \sum_k u(t) \right\|_2^2 + \lambda(t), f(t) - \sum_k u(t) \end{aligned} \tag{2}$$

Then the alternate direction method of multipliers (ADMM) is used for solving the original minimization problem (2) by finding the saddle point of the augmented Lagrangian $\mathcal{L}$ in a sequence of iterative sub-optimizations. Plugging the solutions of the sub-optimizations into the ADMM, and directly optimizing in Fourier domain, the complete algorithm for variational mode decomposition is summarized in the following algorithm [22].

Algorithm: Complete optimization of VMD
Initialize $\left\{\hat{u}_k^1\right\}, \left\{\omega_k^1\right\}, n \leftarrow 0$
repeat
$n \leftarrow n+1$
for $k = 1: K$ **do**
Update $\left\{u_k^1\right\}$ for all $\omega \geq 0$:

$$\left\{\hat{u}_k^{n+1}\right\} \leftarrow \frac{\hat{f}(\omega) - \sum_{i<k} \hat{u}_i^{n+1}(\omega) - \sum_{i>k} \hat{u}_i^n(\omega) + (\hat{\lambda}^n(\omega)/2)}{1 + 2\alpha(\omega - \omega_k^n)^2}$$

Update ω_k: $\omega_k^{n+1} \leftarrow \dfrac{\int_0^\infty \omega \left|\hat{u}_k^{n+1}(\omega)\right|^2 d\omega}{\int_0^\infty \left|\hat{u}_k^{n+1}(\omega)\right|^2 d\omega}$
end for
Dual ascent for all $\omega \geq 0$:

$$\hat{\lambda}^{n+1}(\omega) \leftarrow \hat{\lambda}^n(\omega) + \tau \left(\hat{f}(\omega) - \sum_k \hat{u}_k^{n+1}(\omega) \right)$$

until convergence: $\sum_k \left\| \hat{u}_k^{n+1} - \hat{u}_k^n \right\|_2^2 / \left\| \hat{u}_k^n \right\|_2^2 < \varepsilon$

According to above algorithm, u_k and ω_k should be updated to realize the VMD analysis process. To update the modes, the optimization problem of relation (2) is solved with respect to u_k. This sub optimal problem is represented as follows [22]:

$$\begin{aligned} u_k^{n+1} = \arg\min_{u_k \in X} \left\{ \alpha \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-i\omega_k t} \right\|_2^2 \right. \\ \left. + \left\| f(t) - \sum_i u_i(t) + \frac{\lambda(t)}{2} \right\|_2^2 \right\} \end{aligned} \tag{3}$$

The solution of this quadratic optimization problem is readily found by letting the first variation vanish for the positive frequencies [22]:

$$\hat{u}_k^{n+1}(\omega) = \frac{\hat{f}(\omega) - \sum_{i \neq k} \hat{u}_i(\omega) + (\hat{\lambda}(\omega)/2)}{1 + 2\alpha(\omega - \omega_k)^2} \tag{4}$$

which is clearly identified as a Wiener filtering of the current residual, with signal prior $1/(\omega - \omega_k)^2$. The full spectrum of the real mode is then simply obtained by Hermitian symmetric completion.

Conversely, the real part of the inverse Fourier transform of this filtered analytic signal yield the mode in time domain.

In the second subproblem, the optimization problem of relation (2) is solved with respect to ω_k. The center frequencies do not appear in the reconstruction fidelity term, but only in the bandwidth prior. The relevant problem thus reads [22]:

$$\omega_k^{n+1} = \underset{\omega_k}{\operatorname{argmin}} \left\{ \alpha \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-i\omega_k t} \right\|_2^2 \right\} \tag{5}$$

The solution of above suboptimization problem in frequency domain is obtained as follows [22]:

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega \left| \hat{u}_k(\omega) \right|^2 d\omega}{\int_0^\infty \left| \hat{u}_k(\omega) \right|^2 d\omega} \tag{6}$$

which puts the new ω_k at the gravity center of the corresponding mode's power spectrum. This means carrier frequency is the frequency of a least squares linear regression to the instantaneous phase observed in the mode [22].

2.2. Support vector machines

The purpose of SVM is to find an optimal separating hyperplane by maximizing the margin between the separating hyperplane and the data set [26]. Given a set of data $T = \{x_i, y_i\}_{i=1}^m$ where $x_i \in R^n$ denotes the input vectors, $y_i \in \{+1, -1\}$ stands for two classes, and m is the sample number. It is supposed that the hyperplane $f(x) = 0$ separates the given data which are linearly separable.

$$f(x) = w \cdot x + b = \sum_{k=1}^{m} w_k \cdot x_k + b = 0 \tag{7}$$

where w and b denotes the weight and bias terms. These variables adjust the position of the separating hyperplane. The separating hyperplane should satisfy the following constraints:

$$y_i f(x_i) = y_i (w \cdot x_i + b) \geq 1, \quad i = 1, 2, \ldots, m \tag{8}$$

Positive slack variables ξ_i are defined to measure the distance between the margin and the vectors x_i that lie on the wrong side of the margin. Then, the optimal hyperplane separating the data can be obtained by the following optimization problem:

$$\begin{aligned} &\text{Minimize} \quad \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{m} \xi_i, \quad i = 1, 2, \ldots, m \\ &\text{subject to} \quad \left\{ \begin{array}{c} y_i(w \cdot x + b) \geq 1 - \xi_i \\ \\ \xi_i \geq 0 \end{array} \right\} \end{aligned} \tag{9}$$

where C is the error penalty.

By introducing the Lagrangian multipliers α_i, the above-mentioned optimization problem is transformed into the dual quadratic optimization problem, that is:

$$\text{Maximize} \quad L(\alpha) = \sum_{i=1}^{m} \alpha_i - \frac{1}{2}\sum_{j=1}^{m} \alpha_i \alpha_j y_i y_j (x_i, x_j) \tag{10}$$

$$\text{subject to} \quad \sum_{i=1}^{m} \alpha_i y_i = 0, \quad \alpha_i \geq 0, \quad i = 1, 2, \ldots, m \tag{11}$$

Thus, the linear decision function is created by solving the dual optimization problem, which is defined as:

$$f(x) = sign\left(\sum_{ij=1}^{m} \alpha_i y_i (x_i, x_j) + b \right) \tag{12}$$

SVM can also be used in nonlinear classification using kernel function. The nonlinear mapping function φ is applied to map the original data x into a high-dimensional feature space, where the linear classification is possible. Then, the nonlinear decision function is:

$$f(x) = sign\left(\sum_{ij=1}^{m} \alpha_i y_i K(x_i, x_j) + b \right) \tag{13}$$

where $K(x_i, x_j)$ is called the kernel function, $K(x_i, x_j) = \phi(x_i)\phi(x_j)$. The most commonly used kernel functions are as follows [26,27]:

(1) Linear: $K(x_i, x_j) = x_i^T x_j$
(2) Polynomial: $K(x_i, x_j) = (\beta x_i^T x_j + r)^d, \quad \beta > 0$
(3) Gaussian radial basis function: $K(x_i, x_j) = \exp(-\rho x_i - x_j^2), \quad \rho > 0$
(4) Sigmoid: $K(x_i, x_j) = \tanh(\beta x_i^T x_j + r)$

Here β, r, ρ and d are kernel parameters. Currently in the literature, there is no method available for deciding the value of C, for choosing the best kernel function and for setting the kernel functions. As the proper setting of the SVM parameters has direct effect on the detection accuracy, the appropriate kernel function and other parameters are obtained using heuristic optimization techniques such as continuous ant colony optimization algorithms [28].

The SVMs find an optimal hyperplane on the feature space. Therefore, to classify more than two classes, two straightforward approaches could be used. The first approach is to compare class by class with several machines and combine the outputs using some decision rule. The number of machines N_m needed for the m classes, separation problem is given by [26,27]:

$$N_m = \frac{m!}{2(m-2)!} \tag{14}$$

In this approach, each class is associated with $m-1$ outputs. The advantage of such method is that it gives information about class-by-class separation, which could be used in another system to solve the problem of simultaneous events. In this work, a simple decision rule is used: the winner class is the one which all related outputs $m-1$ are greater than zero.

The second approach to solve the multiple class ($m > 2$) separation problem using SVMs is to compare one class against others. Therefore, the required number of machines N_m is the same number of classes m. The decision rule applied to the m outputs is the winner takes all. In this case, the decision rule is simplified because each output is related to one class. An advantage of this solution is the expected lower computational cost in comparison to the first approach as fewer machines are needed.

2.3. Sequential forward selection

Sequential forward selection was first proposed in [23]. It operates in the bottom-to-top manner. The selection procedure starts with an empty set initially. Then, at each step, the feature maximizing the criterion function is added to the current set. This operation continues until the desired number of features is selected. The nesting effect is present such that a feature added into the set in a step cannot be removed in the subsequent steps [11]. As a consequence, SFS method can offer only suboptimal result.

2.4. Sequential backward selection

Sequential backward selection [24] method proposed in works in a top-to-bottom manner. It is the reverse case of SFS method.

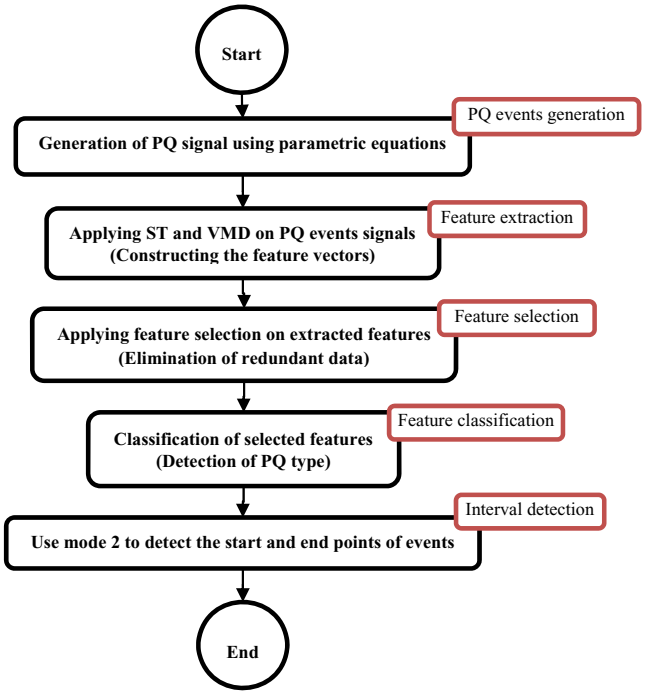

Fig. 1. Flowchart of the proposed algorithm.

Initially, complete feature set is considered. At each step, a single feature is deleted from the present set so that the criterion function is maximized for the remaining features within the set. Removal operation continues until the desired number of features is obtained. Once a feature is eliminated from the set, it cannot enter into the set in the subsequent steps [11].

2.5. Gram–Schmidt orthogonalization based feature selection method

GSO process is a simple forward selection method which can be effectively used for features ranking. Suppose the kth feature of M patterns is denoted by vector $X_k = [x_{k1}, x_{k2}, \ldots, x_{kM}]^T$ and $Y = [y_1, y_2, \ldots, y_M]^T$ represents the vector of output target. In order to select the best correlated feature with output, the cosine of angle between each input feature X_k and target Y is calculated as an evaluation criterion [25]:

$$\cos(\varphi_k) = \frac{\langle X_k \cdot Y \rangle}{\|X_k Y\|} \tag{15}$$

where φ_k is the angle between input kth feature vector X_k and output target Y, N is the number of all features and $\langle X_k \cdot Y \rangle$ denotes the inner product between X_k and Y. If the output is fully proportional to input, the φ_k is zero, and inversely if the output is fully uncorrelated to input, the φ_k is $\pi/2$ [25]. So, in an iterative procedure the feature that maximizes the above mentioned evaluation criterion, is selected as the most correlated feature to target. For selection of the next feature, the output vector and all other candidate features are mapped to null space of the selected feature and then input features and output vectors are updated with new data. The ranking procedure is repeated until all candidate features are ranked, or when a predetermined stopping condition is met [25].

3. Proposed method

Fig. 1 demonstrates the flowchart of the proposed method. The proposed detection scheme is realized through four main steps: generation of PQ signal using parametric equations, feature extraction, feature selection and feature classification. All steps of the proposed method are simulated in MATLAB software [29]. The algorithm presented in this paper is designed to recognize nine classes of different PQ disturbances including: C1: Pure signal, C2: Sag, C3: Swell, C4: Interruption, C5: Harmonics, C6: Transient, C7: Sag with Harmonics, C8: Swell with Harmonics, C9: Flicker. Feature extraction and feature selection play an important role in the proposed detection scheme. In the second step some useful features are extracted using two powerful signal analysis tools, i.e. ST and VMD. In the third stage, to enhance the proposed pattern recognition scheme, redundant features are removed using feature selection methods. Forth stage is the classification in which a cluster label is assigned to each PQ events using the well-known classifier SVM.

3.1. Generating of PQ disturbances

Table 1 shows a detailed summary of PQ disturbances types as well as their controlling parameters, definitions, and equations. Generating of PQ exemplar signals by adjustable parametric equations for recognition purpose has advantages from different viewpoints. The parameters allow the training and testing PQ disturbances to be changed in a wide range and in a controlled manner. The simulated signals in this way are also very similar to the PQ events occur in real power systems. On the other hand, the generalization capability of classifier core improves using different signals belonging to the same class. Two distinct data sets are generated for training and testing phases. A total of 100 cases of each class with various parameters are generated for training and 100 cases are generated for testing. Ten-cycle data window of PQ event signals is considered for extraction of features. In real power system the sampling frequency can be increased up to 10 KHz. If the high sampling frequency is considered in the simulation, more samples will appear in each cycle. More samples increase the signal resolution but the computational burden is consequently increased. On the other hand, few samples give poor resolution of the signal which contains less information about it. For power system applications, high-frequency components usually do not appear in voltage signals. Therefore, the sampling frequency is set to 3200 which is suitable for analysis of PQ disturbances in power systems. Thus, each cycle contains 64 samples for power systems with frequency of 50 Hz. According to Nyquist theorem, the harmonic contents can be monitored up to 1600 Hz which is suitable for analysis of PQ events. Our proposed method can be applied to systems with different sampling frequency but for higher sampling frequency, the required time for implementation of the proposed method is increased.

3.2. Feature extraction

Feature extraction is the most important part of the intelligent pattern recognition schemes. As mentioned in Section 1, many signal processing techniques have been used for extraction of features. Some DFT based algorithms analyze signal only in frequency domain and some DWT based algorithm are very sensitive. In this paper ST as well as VMD are used as powerful signal processing tools for extraction of potential features. ST can analyze signal in both time and frequency domains, simultaneously. On the other hand VMD decomposes signal into different modes so that each mode contains specific spectrum. Thus, VMD is not affected by noise as it can separate high frequency components in a high level mode. Moreover the VMD can trace the signal changes more accurately so that the start and end points of disturbances can be recognized more accurately.

After generation of the PQ events using parametric equations presented in Table 1, some features are extracted by analysis of

Table 1
PQ disturbance model.

PQ disturbance	Class symbol	Model	Parameters
Pure signal	C1	$f(t)=A\sin(\omega t)$	Frequency: 50 Hz, $A=1$
Sag	C2	$f(t)=A(1-\alpha(u(t-t_1))-(u(t-t_2)))\sin(\omega t)$	$0.1\le\alpha\le0.9$; $T\le t_2-t_1\le 9T$
Swell	C3	$f(t)=A(1+\alpha(u(t-t_1))-(u(t-t_2)))\sin(\omega t)$	$0.1\le\alpha\le0.9$; $T\le t_2-t_1\le 9T$
Interruption	C4	$f(t)=A(1-\alpha(u(t-t_1)-u(t-t_2)))\sin(\omega t)$	$0.9\le\alpha\le1$; $T\le t_2-t_1\le 9T$
Harmonic	C5	$f(t)=A(\alpha_1\sin(\omega t)+\alpha_3\sin(3\omega t)+\alpha_5\sin(5\omega t)+\alpha_7\sin(7\omega t))$	$0.05\le\alpha_3\le0.15$; $0.05\le\alpha_5\le0.15$; $0.05\le\alpha_7\le0.15$; $\sum\alpha_i^2=1$
Transient	C6	$f(t)=(\sin(\omega t)+\alpha_{osc}\exp(-(t-t_1)/\tau_{osc}))\sin(\omega_{nosc}(t-t_1))$	$\tau_{osc}=0.008-0.04$ s; $\omega_{nosc}=100-400$ Hz
Sag+ Harmonics	C7	$f(t)=A(1-\alpha(u(t-t_1)-u(t-t_2)))(\alpha_1\sin(\omega t)+\alpha_3\sin(3\omega t)+\alpha_5\sin(5\omega t))$	$T\le t_2-t_1\le 9T$; $0.1\le\alpha\le0.9$; $0.05\le\alpha_3\le0.15$; $0.05\le\alpha_5\le0.15$; $0.1\le\alpha_7\le0.15$; $\sum\alpha_i^2=1$
Swell+ Harmonics	C8	$f(t)=A(1+\alpha(u(t-t_1)-u(t-t_2)))(\alpha_1\sin(\omega t)+\alpha_3\sin(3\omega t)+\alpha_5\sin(5\omega t))$	$T\le t_2-t_1\le 9T$; $0.1\le\alpha\le0.9$; $0.05\le\alpha_3\le0.15$; $0.05\le\alpha_5\le0.15$; $0.1\le\alpha_7\le0.15$; $\sum\alpha_i^2=1$
Flicker	C9	$f(t)=(1+\alpha_f\sin(\beta_f\omega t))\sin(\omega t)$	$\alpha_f=0.1-0.2$; $\beta_f=5$—10

disturbances using ST and VMD. Typical sag and swell signals as well as their decomposed modes are depicted in Figs. 2 and 3, respectively. As it can be seen in Fig. 2, the mode 1 shows the approximation of the main signal which includes low frequency components. As opposed to mode 1, the higher level modes, i.e. modes 2 and 3 include higher frequencies components. This attribute of the VMD helps to detect the start and end points of PQ events. It is clearly seen that at start and end points of sag event, the magnitudes of modes 1 and 2 have an oscillatory behavior. Identically, this characteristic is also observed in the swell waveform depicted in Fig. 3. By setting a proper threshold value, the duration interval of PQ disturbances can be determined precisely. Extensive simulations have been performed and it is found that by using the absolute value of mode 2 and setting the threshold value to 0.05, the duration of events are recognized with high detection accuracy. In this study some features are extracted from the decomposed modes. The standard deviation, energy and maximum absolute value of each mode are calculated as extracted features. Since three decomposition levels are considered in the analysis, nine features are extracted using VMD analysis.

ST is also used for analysis of PQ events for extraction of some features. ST yields a complete visualization of the signal in both

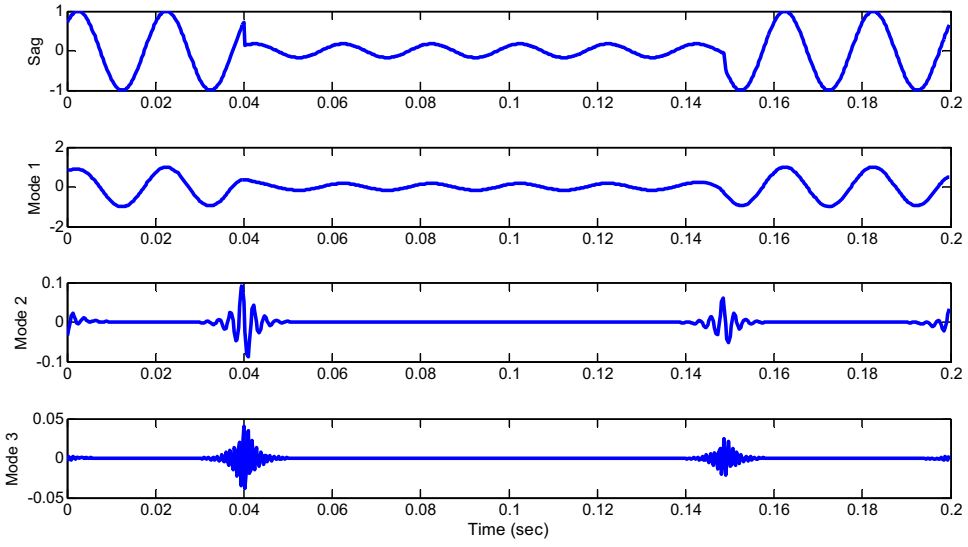

Fig. 2. A typical sag signal and its decomposed modes.

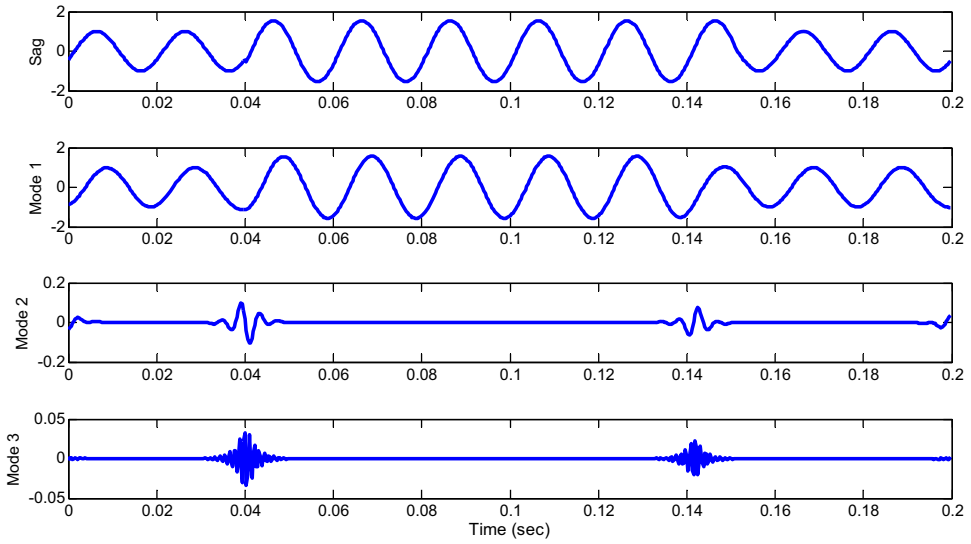

Fig. 3. A typical swell signal and its decomposed modes.

time and frequency domains. The output of ST is a matrix with complex elements whose rows pertain to frequencies and columns show the time. The maximum elements appear in columns and rows yield magnitude and frequency contours, respectively. The phase contour is obtained by calculating the correspondence phase of the element having maximum amplitude in each column. The S-contours of previously illustrated sag and swell have been shown in Figs. 4 and 5, respectively. Ten features are extracted using ST analysis as follows [13]:

- *Feature 10:* Standard deviation of the magnitude contour.
- *Feature 11:* Energy of the magnitude contour.
- *Feature 12:* Standard deviation of the frequency contour.
- *Feature 13:* Standard deviation of phase contour.
- *Feature 14:* Energy of the frequency contour.
- *Features 15–19:* Energy of contour levels 1–5.

3.3. Feature selection

In pattern recognition schemes, elimination of redundant features is a challenging problem. This main stage has been neglected in the most researches in the field of PQ events detection techniques while it has important impacts from following aspects:

- Required memory for data storage decreases and consequently less computational burden is needed for training time.
- The generalization capability of classifier increases and hence the probability of convergence difficulties decreases.

Therefore, feature selection as an essential stage should be considered in the intelligent pattern recognition techniques.

There are many feature selection methods which can be generally classified into two main groups: wrapper and filter. The wrapper method scores feature subset based on the performance of the forecaster model that needs a huge computational burden [11,13]. Filter methods use arithmetic analysis of features without evaluation of forecaster model performance.

In this step, the most useful subset features are selected among the extracted features because inseparable features cannot be recognized even by powerful classifiers. Feature selection is performed using different methods based on filter and wrapper. SFS, SBS as wrappers and GSO as a filter are examined to evaluate the performance of the proposed method.

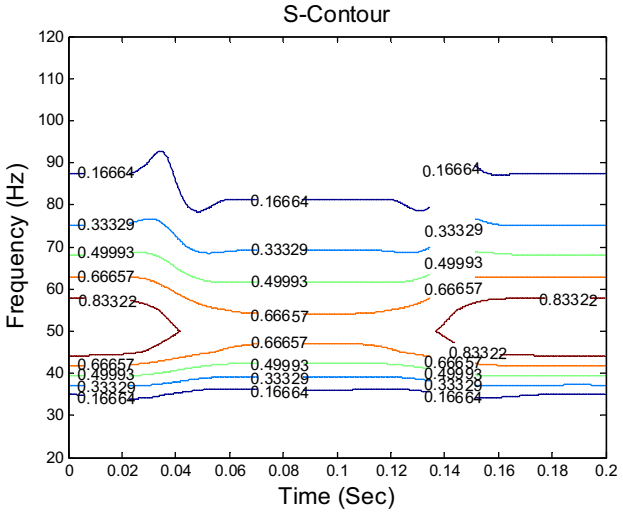

Fig. 4. S-contour of sag.

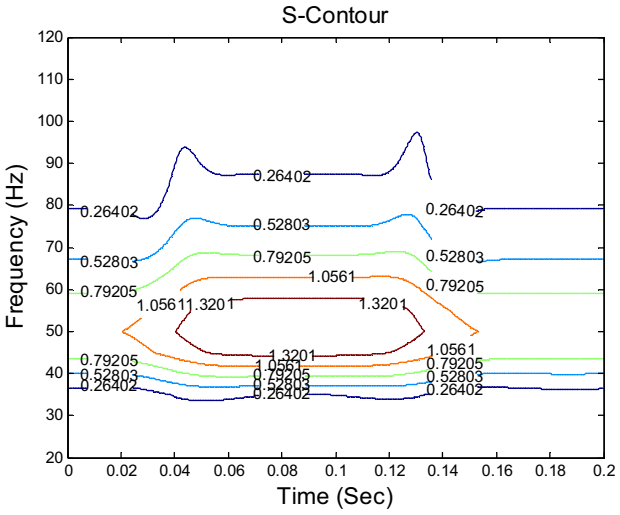

Fig. 5. S-contour of swell.

Table 2
Acquired results considering only extracted features from ST analysis.

Dimension	Feature selection method					
	SFS		SBS		GSO	
	Selected features	Classification accuracy (%)	Selected features	Classification accuracy (%)	Selected features	Classification accuracy (%)
1	5	51.00	5	51.00	2	43.66
2	3, 5	79.44	3, 5	79.44	2, 7	57.55
3	3, 5, 9	90.00	3, 5, 9	90.00	2, 7, 9	81.66
4	3, 4, 5, 9	92.22	3, 4, 5, 9	92.22	2, 3, 7, 9	87.66
5	3, 4, 5, 7, 9	92.89	3, 4, 5, 6, 9	92.89	2, 3, 5, 7, 9	91.55
6	3, 4, 5, 6, 7, 9	93.66	3, 4, 5, 6, 7, 9	93.66	2, 3, 5, 6, 7, 9	92.44
7	**3, 4, 5, 6, 7, 8, 9**	**94.33**	**3, 4, 5, 6, 7, 8, 9**	**94.33**	2, 3, 4, 5, 6, 7, 9	93.77
8	2, 3, 4, 5, 6, 7, 8, 9	94.33	2, 3, 4, 5, 6, 7, 8, 9	94.33	**2, 3, 4, 5, 6, 7, 8, 9**	**94.33**
9	1, 2, 3, 4, 5, 6, 7, 8, 9	93.66	1, 2, 3, 4, 5, 6, 7, 8, 9	93.66	1, 2, 3, 4, 5, 6, 7, 8, 9	93.66

Table 3
Acquired results considering only extracted features from VMD analysis.

Dimension	Feature selection method					
	SFS		SBS		GSO	
	Selected features	Classification accuracy (%)	Selected features	Classification accuracy (%)	Selected features	Classification accuracy (%)
1	12	61.89	12	61.89	11	48.66
2	10, 12	85.78	10, 12	85.78	11, 13	49.22
3	10, 12, 14	95.56	10, 12, 14	95.56	11, 13, 17	68.66
4	10, 12, 14, 17	96.56	10, 12, 14, 17	96.56	11, 13, 15, 17	77.66
5	**10, 12, 13, 14, 17**	**96.66**	**10, 12, 13, 14, 17**	**96.66**	11, 13, 14, 15, 17	92.66
6	10, 12, 13, 14, 16, 17	96.44	10, 12, 13, 14, 16, 17	96.44	**11, 12, 13, 14, 15, 17**	**93.77**
7	10, 12, 13, 14, 16, 17, 18	96.22	10, 12, 13, 14, 16, 17, 18	96.22	11, 12, 13, 14, 15, 17, 18	93.77
8	10, 11, 12, 13, 14, 16, 17, 18	95.88	10, 11, 12, 13, 14, 16, 17, 18	95.88	11, 12, 13, 14, 15, 16, 17, 18	93.77
9	10, 11, 12, 13, 14, 15, 16, 17, 18	94.55	10, 11, 12, 13, 14, 15, 16, 17, 18	94.55	10, 11, 12, 13, 14, 15, 16, 17, 18	94.55
10	10, 11, 12, 13, 14, 15, 16, 17, 18, 19	93.11	10, 11, 12, 13, 14, 15, 16, 17, 18, 19	93.11	10, 11, 12, 13, 14, 15, 16, 17, 18, 19	93.11

3.4. Feature classification

At last step, the SVM classifier is used as classifier core for detection of PQ events while subset selected features from preprocessing steps (feature extraction and feature selection) are utilized for training and testing process. The radial basis function is selected as a kernel function of the SVM and the penalty factor (C) and the adjustable parameter of radial basis function (ρ) are determined using improved ant colony optimization algorithm presented in [28]. After the detection of PQ disturbance type, the mode 2 is used for detection of start and end points of events.

4. Results

To investigate the effect of selected features on the detection accuracy of the proposed method, extracted features from ST and VMD analysis are considered separately. Tables 2 and 3 represent the detection accuracy of the proposed method using extracted features obtained from ST and VMD analysis. For each feature selection method, the subset features with the highest accuracy and the lowest feature dimensions are shown in bold font. The first nine features are those extracted using ST and the rest 10 features are extracted features from VMD analysis which are numbered

Table 4
SFS method with combination of all extracted features using VMD and ST.

Dimension	Selected Features	Classification accuracy (%)
1	12	61.88
2	10, 12	85.77
3	3, 10, 12	95.88
4	3, 10, 12, 14	97.44
5	3, 5, 10, 12, 14	98.77
6	3, 5, 10, 12, 13, 14	99.11
7	3, 5, 8, 10, 12, 13, 14	99.33
8	3, 5, 8, 10, 12, 13, 14, 16	99.33
9	3, 5, 6, 8, 10, 12, 13, 14, 16	99.33
10	3, 5, 6, 8, 10, 12, 13, 14, 16, 17	99.22
11	3, 5, 6, 8, 9, 10, 12, 13, 14, 16, 17	99.11
12	3, 5, 6, 8, 9, 10, 11, 12, 13, 14, 16, 17	99.33
13	3, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 16, 17	99.44
14	3, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 16, 17, 18	99.44
15	**2, 3, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 16, 17, 18**	**99.55**
16	2, 3, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 16, 17, 18, 19	99.22
17	2, 3, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19	98.66
18	2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19	96.33
19	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19	96.11

Boldface indicates the highest accuracy together with the lowest feature dimensions.

Table 5
SBS method with combination of all extracted features using VMD and ST.

Dimension	Selected features	Classification accuracy (%)
1	3	48.88
2	3, 11	86.44
3	3, 9, 11	96.00
4	3, 9, 10, 11	98.00
5	3, 9, 10, 11, 14	99.00
6	3, 9, 10, 11, 12, 14	99.22
7	3, 7, 9, 10, 11, 12, 14	99.33
8	3, 7, 9, 10, 11, 12, 13, 14	99.44
9	2, 3, 7, 9, 10, 11, 12, 13, 14	99.55
10	2, 3, 5, 7, 9, 10, 11, 12, 13, 14	99.55
11	2, 3, 5, 6, 7, 9, 10, 11, 12, 13, 14	99.55
12	2, 3, 5, 6, 7, 9, 10, 11, 12, 13, 14, 16	99.55
13	**2, 3, 5, 6, 7, 9, 10, 11, 12, 13, 14, 16, 17**	**99.66**
14	2, 3, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 16, 17	99.44
15	2, 3, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 16, 17, 19	99.33
16	2, 3, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 19	99.00
17	1, 2, 3, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 19	96.66
18	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 19	96.55
19	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19	96.11

Boldface indicates the highest accuracy together with the lowest feature dimensions.

Table 6
GSO method with combination of all extracted features using VMD and ST.

Dimension	Selected features	Classification accuracy (%)
1	2	43.66
2	2, 7	57.55
3	2, 7, 9	81.66
4	2, 3, 7, 9	87.66
5	2, 3, 7, 9, 18	92.44
6	2, 3, 7, 9, 16, 18	92.88
7	2, 3, 7, 9, 12, 16, 18	93.88
8	2, 3, 7, 9, 12, 14, 16, 18	95.44
9	2, 3, 7, 9, 11, 12, 14, 16, 18	98
10	2, 3, 7, 9, 11, 12, 14, 15, 16, 18	98
11	2, 3, 6, 7, 9, 11, 12, 14, 15, 16, 18	98.11
12	2, 3, 6, 7, 8, 9, 11, 12, 14, 15, 16, 18	98.33
13	**2, 3, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 18**	**98.77**
14	2, 3, 4, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 18	96.66
15	2, 3, 4, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 17, 18	96.77
16	2, 3, 4, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18	97
17	2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18	96.77
18	2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19	96.33
19	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19	96.11

Boldface indicates the highest accuracy together with the lowest feature dimensions.

from 10 to 19. Table 2 shows that for all SFS, SBS and GSO methods, the best detection accuracy is 94.33. Features 3, 4, 5, 6, 7, 8 and 9 appear in all the best solutions, so these feature can be considered as the "must select" features, regarding the classification accuracy. Similarly, the detection accuracy of the proposed method is calculated for extracted features while VMD by itself is used for analysis of PQ events. The best detection accuracies are 96.66%, 96.66% and 93.77% for SFS, SBS and GSO methods, respectively. Features 12, 13, 14 and 17 as dominant features have been repeated in all the best solutions of different feature selection methods.

Obtained results show that extracted features using a single analysis tool cannot yield acceptable detection accuracy. Thus, in the following investigation, the combination of all extracted features using both ST and VMD analysis tools are considered as candidate features. Then by applying different feature selection methods to 19-dimension vector of extracted features, the best

Table 7
Percentage of correct classification results under noiseless and different noisy conditions considering three classifiers.

Class	20 dB			30 dB			40 dB			50 dB			Noiseless condition		
	SVM	ANN	KNN	SVM	ANN	KNN	SVM	ANN	KNN	SVM	ANN	KNN	SVM	ANN	KNN
C1	**100**	99	99	**100**	100	100	**100**	100	100	**100**	100	100	**100**	100	100
C2	**96**	93	95	**97**	96	95	**98**	97	96	**98**	98	97	**99**	98	98
C3	**99**	94	95	**100**	95	96	**100**	95	97	**99**	96	97	**99**	97	97
C4	**100**	100	98	**100**	100	99	**100**	100	100	**100**	100	100	**100**	100	100
C5	**99**	96	97	**99**	97	98	**99**	98	99	**100**	99	100	**100**	99	100
C6	**98**	96	98	**99**	97	98	**99**	97	98	**100**	97	99	**100**	98	99
C7	**99**	96	97	**100**	97	98	**100**	98	98	**100**	98	98	**99**	98	99
C8	**97**	93	95	**97**	95	96	**97**	97	97	**98**	97	96	**100**	97	98
C9	**95**	94	95	**96**	95	95	**97**	96	96	**99**	97	97	**100**	98	97
Overall accuracy (%)	**98.11**	**96**	**96.55**	**98.66**	**96.88**	**97.22**	**99.00**	**97.55**	**97.88**	**99.33**	**98.00**	**98.22**	**99.66**	**98.33**	**98.66**

Boldface indicates the highest accuracy together with the lowest feature dimensions.

Table 8
Performance comparison in terms of percentage of correct classification results.

Reference	Method	Investigated PQ disturbance types	Overall accuracy (%)	
			Noiseless condition	Noisy condition (20 dB SNR)
[7]	WT + SVM	Pure signal, sag, swell, interruption, harmonics, transient, sag & harmonics, swell & harmonics, flicker	98.89	96.33
[8]	WT + RVM	Pure signal, sag, swell, interruption, harmonics, transient, sag & harmonics, swell & harmonics, flicker	99.03	98.47
[9]	WT + NN	Pure signal, sag, swell, interruption, harmonics, sag & harmonics, swell & harmonics	95.71	89.92
[10]	WT + NN	Frequency capacitor switching (high, low), normal, impulsive transient, sag, interruption	92.3	–
[13]	(WT + ST) + FS + PNN	Pure signal, sag, swell, interruption, harmonics, transient, sag & harmonics, swell & harmonics, flicker	99.22	97.44
[17]	HT + RBFN	Normal signal, sag, swell, harmonics, transients, voltage flicker	97.00	94
This paper	(ST + VMD) + FS + SVM	Pure signal, sag, swell, interruption, harmonics, transient, sag & harmonics, swell & harmonics, flicker	99.66	98.11

subset features are selected. Tables 4–6 represent the detection accuracy of different number of selected features. By applying the SFS, 15-dimension feature subset yields the best detection accuracy, i.e. 99.55%. As observed from Table 4, the detection accuracy of the proposed method improves by dimension increase of selected subset features until it reaches to its maximum value at which the subset feature vector size reaches to 15. Thereafter, the detection accuracy decreases until all extracted features are selected. Similarly, Tables 5 and 6 give the calculated detection accuracies for SFS and GSO methods for different dimension of subset features. The best selected subsets yield 99.66% and 98.77% detection accuracy for SBS and GSO feature selection methods, respectively. Features 2, 3, 6, 7 and 9 which have been extracted using ST appear in all the best subsets. On the other hand, features 10, 11, 12, 14 and 16 are those selected features of VMD analysis which appear in all the best solutions of applying different feature selection methods. Results demonstrate that the best performance belongs to SBS method with detection accuracy of 99.66%.

Moreover, the signals in real electric power systems are usually contaminated with noise. To consider the noisy conditions, additive white Gaussian noise (AWGN) is added to PQ events [8,13]. The sensitivity of the proposed algorithm under different noise levels with

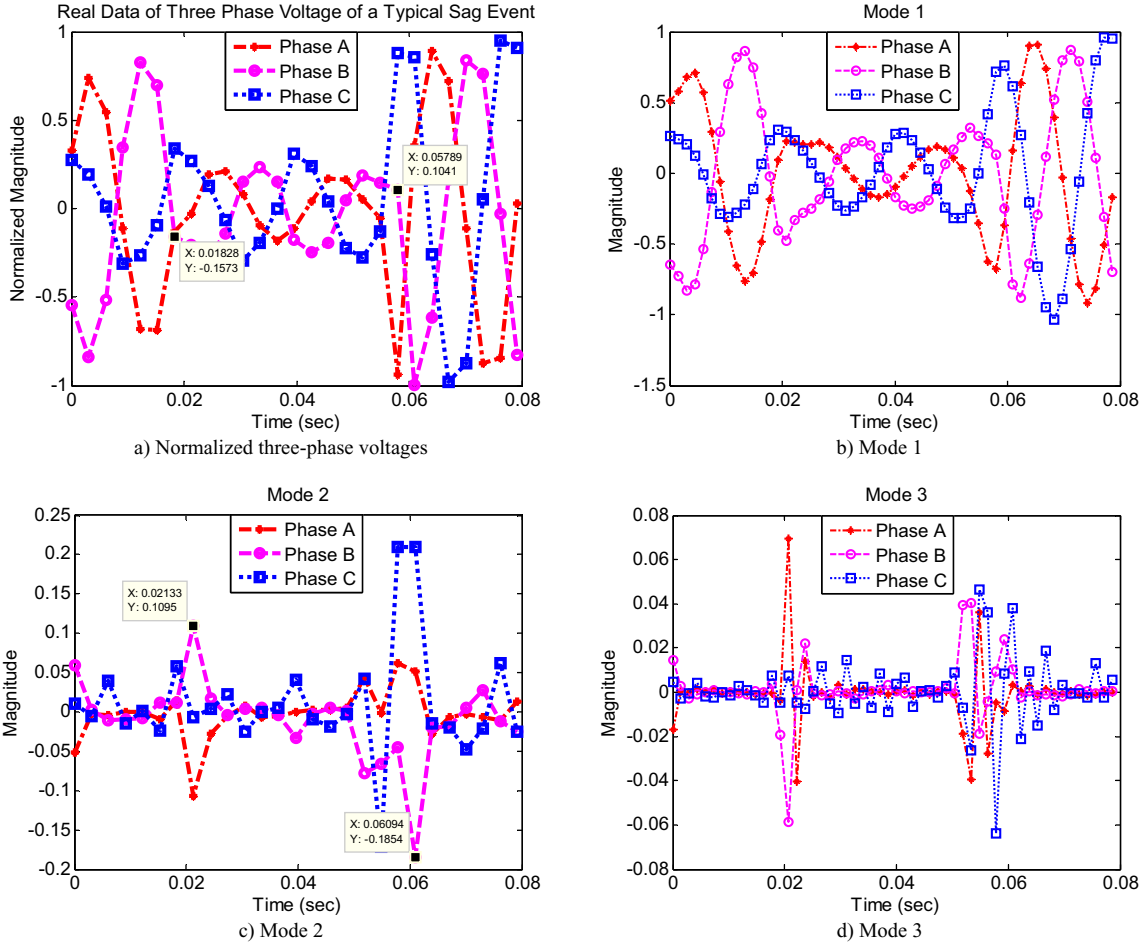

Fig. 6. Real three-phase sag voltages together with decomposed modes of VMD analysis. (a) Normalized three-phase voltages, (b) mode 1, (c) mode 2, (d) mode 3.

the signal-to-noise ratio (SNR) between 20 and 50 dB are tested to see how well the proposed scheme would work in the presence of noise. Besides, different classifiers have been considered as classification core of the detection scheme to investigate the effect of classifier type on the detection scheme accuracy. Overall, the detection accuracy of the proposed method decreases when SNR decreases but the SVM classifier is less sensitive to noise. Results depict the SVM classifier has the best performance in comparison with two other classifiers, i.e. artificial neural network (ANN) and K-nearest neighbor (KNN) in noiseless and all noisy conditions (Table 7).

To evaluate the effectiveness and feasibility of the proposed hybrid method, a comparison results in terms of PQ events classification accuracy has been made as given in Table 8. In our proposed method, nine types of PQ events have been considered and simulated in wide range. Ref. [9] contains only seven types of disturbances (exclude transient and flicker) but uses the same mathematical equations as used in this paper. Some disturbances such as sag and harmonics that may occur simultaneously are not considered in Ref. [17]. According to Table 8, the performance of the classification process with the proposed hybrid feature extraction method surpasses the performance of the detection accuracy of some proposed methods of other papers in the field of PQ classification.

To justify the applicability of the proposed method in real power system, a typical waveform of the sag event and its decomposed modes have been shown in Fig. 6. It can be seen from Fig. 6(a) that mode 1 containing low frequency contents is an approximation of the main signal. In contrast with mode 1, modes 2 and 3 (Fig. 6 (b) and (c)) contain high frequency components so that the start and ends of even can be determined easily. As depicted in Fig. 6(a), the sag signal of phase B starts at $t = 0.018$ s and it ends at $t = 0.058$ s, so the interval time of sag event is about 0.04 s. Fig. 6(c) shows that the magnitude of mode 2 exceeds the threshold value of 0.1 at $t = 0.02$ and $t = 0.06$ s.

The start and end points of the event are detected precisely with only 0.002 s estimation error.

5. Conclusions

Despite an extensive number of research works in the area of PQ events detection, there is still a need of precise analysis of combination of extracted features. At first large dimension of extracted feature vectors are constructed using well-known analysis tools VMD and ST, and then more meaningful and important feature subsets are selected using several feature selection methods on the basis of filter and wrapper methods. By elimination of redundant features, required memory and computational burden of detection process decrease while its generalization capability is improved, considerably. Results show that extracted features using only a single signal analysis tool cannot yield the best detection accuracy. On the other hand, the extracted features using both signal analyzers, i.e. VMD and ST together with SVM classifier has the best detection accuracy. Among the applied feature selection methods, SBS gives the best detection accuracy of 99.66%. Besides, mode 2 which is the output of VMD analysis can be used for detection of start and end of PQ events. Thus, this new presented algorithm can be effectively used for monitoring of PQ disturbances in real time application.

References

[1] A.B. Baggini, Handbook of Power Quality, John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England, 2008.
[2] O.P. Mahela, A.G. Shaik, N. Gupta, A critical review of detection and classification of power quality events, Renew. Sustain. Energy Rev. 41 (2015) 495–505.
[3] M. Manjula, S. Mishra, A.V.R.S. Sarma, Empirical mode decomposition with Hilbert transform for classification of voltage sag causes using probabilistic neural network, Int. J. Electr. Power Energy Syst. 44 (2013) 597–603.
[4] O. Ozgonenel, T. Yalcin, I. Guney, U. Kurt, A new classification for power quality events in distribution systems, Electr. Power Syst. Res. 95 (2013) 192–199.
[5] J.V. Wijayakulasooriya, G.A. Putrus, P.D. Minns, Electric power quality disturbance classification using self-adapting artificial neural networks, IEE Proc. Gener. Transm. Distrib. 149 (2002) 98–101.
[6] G.T. Heydt, P.S. Fjeld, C.C. Liu, D. Pierce, L. Tu, G. Hensley, Applications of the windowed FFT to electric power quality assessment, IEEE Trans. Power Deliv. 14 (1999) 1411–1416.
[7] Z. Moravej, A.A. Abdoos, M. Pazoki, Detection and classification of power quality disturbances using wavelet transform and support vector machines, Electr. Power Compon. Syst. 38 (2009) 182–196.
[8] Z. Moravej, M. Pazoki, A.A. Abdoos, Wavelet transform and multi-class relevance vector machines based recognition and classification of power quality disturbances, Eur. Trans. Electr. Power 21 (2010) 212–222.
[9] M. Uyar, S. Yildirim, M.T. Gencoglu, An effective wavelet-based feature extraction method for classification of power quality disturbance signals, Electr. Power Syst. Res. 78 (2008) 1747–1755.
[10] A.S. Yilmaz, A. Subasi, M. Bayrak, V.M. Karsli, E. Ercelebi, Application of lifting based wavelet transforms to characterize power quality events, Energy Convers. Manag. 48 (2007) 112–123.
[11] S. Gunal, O.N. Gerek, D.G. Ece, R. Edizkan, The search for optimal feature set in power quality event classification, Expert Syst. Appl. 36 (2009) 10266–10273.
[12] H.S. Behera, P.K. Dash, B. Biswal, Power quality time series data mining using S-transform and fuzzy expert systems, Appl. Soft Comput. 10 (2010) 245–255.
[13] A.A. Abdoos, Z. Moravej, M. Pazoki, A hybrid method based on time frequency analysis and artificial intelligence for classification of power quality events, J. Intell. Fuzzy Syst. 28 (2015) 1183–1193.
[14] M. Biswala, P.K. Dash, Detection and characterization of multiple power quality disturbances with a fast S-transform and decision tree based classifier, Digit. Signal Process. 23 (2013) 1071–1083.
[15] C.N. Bhende, S. Mishra, B.K. Panigrahi, Detection and classification of power quality disturbances using S-transform and modular neural network, Electr. Power Syst. Res. 78 (2008) 122–128.
[16] M.J.B. Reddy, K. Sagar, D.K. Mohanta, A multifunctional real-time power quality monitoring system using Stockwell transform, IET Sci. Meas. Technol. 8 (2014) 155–169.
[17] T. Jayasree, D. Devaraj, R. Sukanesh, Power quality disturbance classification using Hilbert transform and RBF networks, Neurocomputing 73 (2010) 1451–1456.
[18] N. Senroy, S. Suryanarayanan, P.F. Ribeiro, An improved Hilbert–Huang method for analysis of time-varying waveforms in power quality, IEEE Trans. Power Syst. 22 (2007) 1843–1850.
[19] K. Manimala, K. Selvi, R. Ahila, Hybrid soft computing techniques for feature selection and parameter optimization in power quality data mining, Appl. Soft Comput. 11 (2011) 5485–5497.
[20] B. Biswal, M. Biswal, S. Hasan, P.K. Dash, Nonstationary power signal time series data classification using LVQ classifier, Appl. Soft Comput. 18 (2014) 158–166.
[21] R.G. Stockwell, L. Mansinha, R.P. Lowe, Localization of the complex spectrum: the S-transform, IEEE Trans. Signal Process. 4 (1996) 998–1001.
[22] K. Dragomiretskiy, D. Zosso, Variational mode decomposition, IEEE Trans. Signal Process. 62 (2014) 531–544.
[23] A.W. Whitney, A direct method of nonparametric measurement selection, IEEE Trans. Comput. 20 (1971) 1100–1103.
[24] T. Marill, D.M. Green, On the effectiveness of receptors in recognition system, IEEE Trans. Inf. Theory 9 (1963) 11–17.
[25] H. Stoppiglia, G. Dreyfus, R. Dubois, Y. Oussar, Ranking a random feature for variable and feature selection, J. Mach. Learn. Res. 3 (2003) 1399–1414.
[26] C. Cortes, V. Vapnik, Support vector networks, Mach. Learn. 20 (1995) 273–297.
[27] V.N. Vapnik, Statistical Learning Theory, Wiley, New York, 1998.
[28] W.-M. Hung, W.-C. Hong, Application of SVR with improved ant colony optimization algorithms in exchange rate forecasting, Control Cybern. 38 (2009) 863–891.
[29] MATLAB 7.7.0 (2008b) Version, The Math Works Company, Natick, MA, 2008.