



A Power Quality Disturbance Classification System Based on Wavelet Transformation and Back Propagation Network

¹Md Rabiul Islam, ²Shan Gao

¹Postgraduate Student, ²Professore
School of Electrical Engineering,
Southeast University, Nanjing, China

Abstract : This paper is proposed a method for Power Quality (PQ) disturbances classification based on Wavelet transformation and Back-propagation (BP) algorithm. Several types of PQ disturbance are taken in consideration in this paper. Wavelet transformation is used to construct the feature vector form these PQ disturbances based on Multi-resolution analysis (MRA). This feature vectors then are applied to BP network for training and testing. Comparison research between the proposed method and the existing literature shows that the proposed can provide accurate classification result. By the hypothesis test of the average classification accuracy values of different wavelets shows that one can choose any wavelet with short filter length to achieve a good classification result with a low computational cost. Gaussian white noise and Monte Carlo method are taken in consideration to simulate the performance of the proposed method in different noise environments.

Index-terms - Power quality (PQ), noise, disturbance classification, wavelet transformation, neural network, back-propagation.

I. INTRODUCTION

POWER QUALITY (PQ) disturbance is becoming a very important issue in the sector of power industry recently. It can cause distorted waves and degrade voltage quality to impair the operation of electrical apparatus. These disturbances cause problems in power systems, such as overheating, motor failures, inaccurate metering, and miss-operation of protective equipment. Power harmonic pollution is caused by the converters, reactors, and nonlinear loads. According to the data provide by Electrical Power Research Institute (EPRI), the US economy is losing between \$104 billion and \$164 billion a year to outages and another \$15 billion to PQ phenomena [1]. So in the recent decade, the research of power quality issue has captured exponentially increasing attention in the power engineering community.

This paper is focused on the PQ disturbances classification problem. Artificial intelligent (AI) and the machine learning is one of the most power tool which deal with this kind of issue. Many researcher have used these tools for PQ problems such as Expert System [2], Fuzzy logic [3]-[6] etc.

Recently for this PQ problem the Wavelet transformation has become a vary power tool. The wavelet transform involves time-frequency window which yield nice for PQ classification. The combination of wavelet transformation and Generalized Regression Neural Network (GRNN) is proposed in this paper [7]. U. D. Dwivedi et al. proposed an effective De-noising techniques with change-point approach for wavelet-based power-quality monitoring [8]. Automatic voltage disturbance detection and classification using wavelets and multi-class logistic regression is discussed in this paper [9]. Several typical power quality disturbances are correctly classified in these papers. Cong-Hui Huang et al. proposed chaos synchronization-based detector for power-quality disturbances classification in a Power System [10]. Although lots of researches achievements have been reported, the objective of classifying different kinds of power quality disturbances is still both difficult and challenging.

Motivated by the recent research in the area of wavelet analysis and the intelligent system, this paper propose an effective classification method based on wavelet transformation and back-propagation (BP) network. In this method, after multi-resolution analysis on the original sampled power waveform, the energy of the detail and approximation at each decomposition level is calculated to constructed feature vector for future training and testing. Then this data is then applied to the neural network which is constructed based on back-propagation algorithm (BP).

The rest of this paper is as follows. Section II, a wavelet based feature extraction scheme is proposed which shows the way to construct (l+1) dimension feature vector from l level decomposition. In this way feature vector keeps the necessary PQ

characteristics for future classification. Section III discusses the back-propagation (BP) algorithm for the classification. Section IV shows the result based on the method proposed in this paper. Comparison the research result between the proposed method and the existing literature report are discussed in detail in this section. Next the relationship between the performance of classification accuracy and the decomposition level is showed. By using the hypothesis test of the average classification accuracy values of different wavelets we show that there is no restriction in choosing any particular wavelet for the classification. This means one can choose any wavelet with short filter length to perform this operation according to the proposed method with good classification accuracy as well as small computational cost. In Section V, noise is taken in consideration with different level and Monte Carlo method is used to generate the training and testing data. The result shows the effectiveness of the proposed method in noisy environments. At last, a conclusion is given in section VI.

II. WAVELET BASED FEATURE EXTRACTION

The mathematics about the Wavelet transform were studied and this will be found in [11] and [12]. Mallat introduced the multi-resolution analysis and a detailed study about MRA introduces the multi-resolution analysis and this can be found in [13].

2.1 Continuous Wavelet Transform

If there is a finite signal $f(t) \in L^2(\mathbb{R})$, and it satisfies,

$$\int_{-\infty}^{+\infty} |f(t)|^2 dt < \infty \quad (2.1)$$

The continuous wavelet transform (CWT) of a signal $f(t)$ is then [14],

$$CWT_{\psi} f(a, b) = \int_{-\infty}^{+\infty} f(t) \frac{1}{\sqrt{|a|}} \left(\psi \frac{t-b}{a} \right) dt \quad (2.2)$$

Where $\Psi(t)$ is the mother wavelet, the constant a and b denotes ($a, b \in \mathbb{R}$) are scaling (dilation) and transform parameters, respectively. Here the scale parameter a will decide the oscillatory frequency and the length of wavelet, the parameter b will decide its shifting position. Mother wavelet $\Psi(t)$ is defined as,

$$\int_{-\infty}^{+\infty} \Psi(t) dt = 0 \quad (2.3)$$

2.2 Discrete Wavelet Transform

In a practical application generally the discrete wavelet transform (DWT) is used instead of CWT. This is implemented by using discrete values of the scaling parameter a and translation parameter b . For doing this, let set $a = 2^l$ and $b = 2^m$; $m \in \mathbb{Z}$, The Discrete wavelet transform (DWT) of a signal $f(t)$ is then [14],

$$DWT_{\psi} f(l, m) = \int_{-\infty}^{+\infty} f(t) \overline{\Psi_{l,m}(t)} dt \quad (2.4)$$

$$\text{Where, } \Psi_{l,m}(t) = \frac{1}{\sqrt{2}} \Psi(2^{-l}t - m) \quad (2.5)$$

We have used DWT for our classification analysis.

2.3 Multi-resolution Analysis

In MRA, any time series $f(t)$ can be decomposed in terms of approximations, which is provided by scaling functions $\Theta_m(t)$ and the details, which is provided by the wavelets $\Psi_m(t)$, where the $\Theta_{m,n}(t)$ and $\Psi_{m,n}(t)$ are defined as the follow,

$$\Theta_{m,n}(t) = 2^{-m/2} \theta(2^{-m}t - n) \quad \& \quad \Psi_{m,n}(t) = 2^{-m/2} \psi(2^{-m}t - n) \quad (2.6)$$

Now the scaling function is associated with the low-pass filters coefficients $\{k(n), n \in \mathbb{Z}\}$, and the wavelet function is associated with the high-pass filters with the filter coefficients $\{l(n), n \in \mathbb{Z}\}$. Then so called Two scale Equations (TSE) give rise to these filters.

$$\theta(t) = \sum_n k(n) \sqrt{2} \theta(2t - n) \quad \& \quad \psi(t) = \sum_n l(n) \sqrt{2} \theta(2t - n) \quad (2.7)$$

For these filters there are more important properties as follow,

$$\sum_n k(n)^2 = 1 \quad \& \quad \sum_n l(n)^2 = 1 \quad (2.8)$$

$$\sum_n k(n) = \sqrt{2} \quad \& \quad \sum_n l(n) = 0 \quad (2.9)$$

The characteristics about these filters and the way to control them, Daubechies has given a detailed discussion which will found in [11]. The decomposition of a signal is performed as follow,

1. The signal is passed through the filter,
2. The approximations are the low-frequency components of the time series and are found at the output of low-pass filter and the details are the high-frequency components are found at the output of high-pass filter.

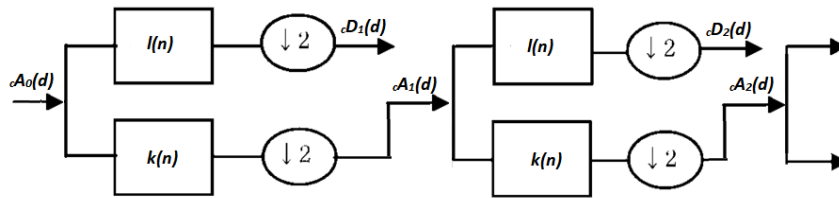


Fig. 1. Two-stage wavelet analysis filter bank

MRA leads to a hierarchical and fast scheme which can be implemented by a set of successive filter bank as shown in Fig.1, where $k(n)$ and $l(n)$ indicate low-pass and high-pass filters as defined (7)-(9). In Fig. 1, $\downarrow 2$ means the down sampling with a factor of $\downarrow 2$, d is the coefficient index at each decomposition level.

Now according to the filter bank implementation in Fig. 1. the relationship of approximation coefficients and the detailed coefficients are given as follow,

$$cA_j(d) = \sum_n K(2d - n)cA_{j-1}(n) \quad \& \quad cD_j(d) = \sum_n l(2d - n)cA_{j-1}(n) \quad (2.10)$$

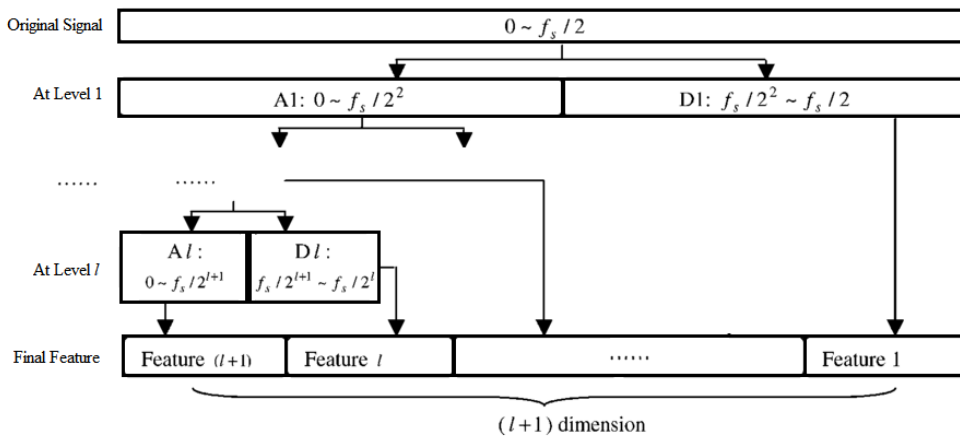


Fig. 2 MRA analysis and feature construction

Here cA_j and cD_j is the approximation coefficients and detail coefficients of the signal at level j respectively. So the decomposition coefficients of MRA analysis can be represent as follows,

$$\begin{aligned}
 [A_0] &\leftrightarrow [cA_1, cD_1] && \text{(level 1)} \\
 &\leftrightarrow [cA_2, cD_2, cD_1] && \text{(level 2)} \\
 &\leftrightarrow [cA_3, cD_3, cD_2, cD_1] && \text{(level 3)} \\
 &\leftrightarrow \dots\dots\dots &&
 \end{aligned}
 \quad (2.11)$$

In this way the decomposition of signal $f(t)$ is,

$$\begin{aligned}
 f(t) &= A_1(t) + D_1(t) && \text{(level 1)} \\
 &= A_2(t) + D_2(t) + D_1(t) && \text{(level 2)} \\
 &= A_3(t) + D_3(t) + D_2(t) + D_1(t) && \text{(level 3)} \\
 &= \dots\dots\dots &&
 \end{aligned}
 \quad (2.12)$$

Here $A_i(t)$ and $D_i(t)$ are called the approximation and the detail at level i respectively. In Fig. 2, the MRA decomposition in frequency domain for a signal which is sampled with a frequency f_s is demonstrated as both the high pass filter and low pass filter are half band.

2.4 Feature Extraction Multi-resolution Analysis

Now instead of using the approximation and detailed value for using training and testing we use the energy at each decomposition level as the input of Back-propagation network. The energy at each decomposition level is calculated by using the following equations,

Here, $i=1 \dots l$ is the wavelet decomposition level from 1 to l and N is the no. of the coefficients of approximation and detail at each decomposition level. FD_i is the energy of the detail at decomposition level i and FA_i is the energy of the approximation at decomposition level l . So, for a l level wavelet decomposition, a $(l+1)$ dimension feature vector is found for our future analysis. In Fig. 3, the flow of data is shown in this wavelet feature extraction.

$$FD_i = \sum_{j=1}^N |D_{ij}|^2 \quad \& \quad FA_i = \sum_{j=1}^N |A_{ij}|^2, \quad i = 1, 2, 3, \dots, l \quad (2.13)$$

Here, $i=1 \dots l$ is the wavelet decomposition level from 1 to l and N is the no. of the coefficients of approximation and detail at each decomposition level. FD_l is the energy of the detail at decomposition level l and FA_l is the energy of the approximation at decomposition level l .

So, for a l level wavelet decomposition, a $(l+1)$ dimension feature vector is found for our future analysis. In Fig. 3, the flow of data is shown in this wavelet feature extraction.

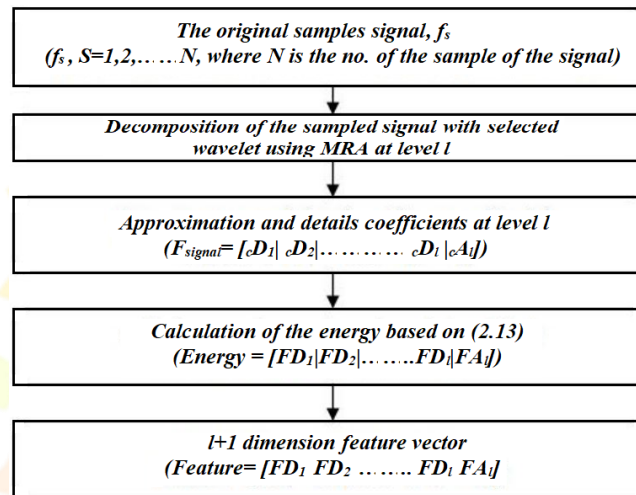


Fig. 3. Wavelet based feature extraction.

III. BACK PROPAGATION ALGORITHM

Back-propagation (BP) is most popular supervised learning method for multi-layer neural networks [15]. BP have been applied to a wide variety of problems, including pattern recognition, signal processing, image compression, speech recognition etc due to its most appealing features and and adaptive nature [16].

Back-propagation an abbreviation for "backward propagation of errors" is a common method of training artificial neural network. For a desired output, the network learns from many inputs, similar to the way a child learns to identify a dog from examples of dogs. The goal of supervised learning algorithm is to find a function that best maps a set of inputs to its correct output.

The learning process of back-propagation algorithm can be divided into two phases,

Phase 1: Propagation

Each propagation involves the following steps:

- 1) Forward propagation of a training pattern's input through the neural network in order to generate the propagation's output activation.
- 2) Backward propagation of the propagation's output activation through the neural network using the training pattern target in order to generate the deltas of all output and hidden neurons.

Phase 2: Weight update

For each weight-synapse follow the following steps:

- 1) Multiply its output delta and input activation to get the gradient of the weight.
- 2) Bring the weight in the opposite direction of the gradient by subtracting a ratio of it from the weight.

This ratio influences the speed and quality of learning; it is called the *learning rate*. The sign of the gradient of a weight indicates where the error is increasing; this is why the weight must be updated in the opposite direction. Repeat phase 1 and 2 until the performance of the network is satisfactory. The no. of repeats is called *iteration*.

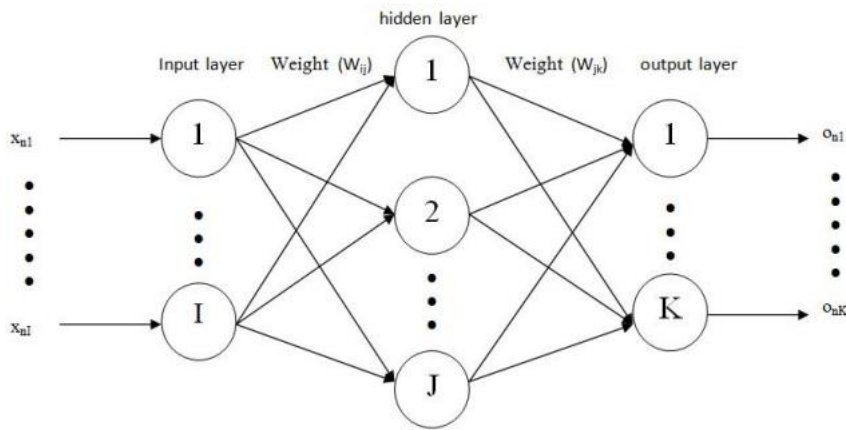


Fig. 4. Back propagation network

Now a feed-forward (FF) Back-propagation network with one input layer, one hidden layer and one output layer is shown in Fig. 4. Let I , J and K be the no. of input units, hidden units and output units respectively. The network weight that connect input unit and the hidden unit is W_{ij} and the weight that connect the hidden unit and the output unit is W_{jk} . The number of training examples is N and any arbitrary n -th training example is $(x_{n1}, x_{n2}, \dots, x_{nI}; y_{n1}, y_{n2}, \dots, y_{nK})$, where x_n is the input vector and y_n is the target output. h_{nj} and o_{nk} are the outputs of hidden unit j and output unit k for the n -th training example. Δ is the difference between the current and new value of the network weights. In the back-propagation algorithm the mathematical expression to update these weights to achieve the desired output according to Phase 1 and 2 is as follow,

$$h_{nj} = f\left(\sum_{i=1}^I w_{ij}x_{ni}\right) \quad \& \quad o_{nk} = f\left(\sum_{j=1}^J w_{jk}x_{nj}\right) \quad (3.1)$$

Here, the sigmoid function $f(x) = 1/(1+e^{-x})$ is used as activation function. The calculation of the changes of weight, Δw_{ij} and Δw_{jk} is as follow,

$$\Delta w_{jk} = -LR(\partial E_n / \partial w_{jk}) = -LR \times h_{nj} \times \delta_{nk} \quad \& \quad \Delta w_{ij} = -LR \times x_{ni} \times \delta_{nj} \quad (3.2)$$

$$\text{Where, } E_n = \frac{1}{2} \sum_{k=1}^K [o_{nk} - y_{nk}] \quad (3.3)$$

$$\text{weight update, } w_{ij}(\text{new}) = w_{ij}(\text{old}) + \Delta w_{ij} \quad \& \quad w_{jk}(\text{new}) = w_{jk}(\text{old}) + \Delta w_{jk} \quad (3.3)$$

A detail discussion can be found in this paper [15].

IV. SIMULATION AND RESULT

4.1 Generation of Data for Future analysis

The simulation data was generated in MATLAB based on model in [17]. Here, Seven classes (C1-C7) of different PQ disturbances which named as undisturbed sinusoidal (normal), swell, sag, harmonics, outage, sag with harmonic and swell with harmonic, were considered. These are shown in the Table 1 with their control parameters. For both training and testing 200 cases of each class are generated with different parameters. While training and testing data are generated, the sample frequency is 256 point/cycle and the normal frequency is 50 Hz. So, ten power frequency cycles which contain the disturbance are used for a total of 2560 points.

Table 1 Power Quality Disturbance Model

PQ disturbances	Class name	Model	Control parameters
Normal	C1	$f(t)=\sin(\omega t)$	
Swell	C2	$f(t)=A(1+ \alpha(u(t-t_1)- (u(t-t_2))))\sin(\omega t)$ $t_1 < t_2, u(t)=\begin{cases} 1, & t \geq 0 \\ 0, & t < 0 \end{cases}$	$0.1 \leq \alpha \leq 0.8$ $T \leq t_2 - t_1 \leq 9T$
Sag	C3	$f(t)=A(1- \alpha(u(t-t_1)- (u(t-t_2))))\sin(\omega t)$	$0.1 \leq \alpha \leq 0.9; T \leq t_2 - t_1 \leq 9T$
Harmonic	C4	$f(t)=A(\alpha_1 \sin(\omega t) + \alpha_3 \sin(3\omega t) + \alpha_5 \sin(5\omega t) + \alpha_7 \sin(7\omega t))$	$0.05 \leq \alpha_3 \leq 0.15, 0.05 \leq \alpha_5 \leq 0.15,$ $0.05 \leq \alpha_7 \leq 0.15,$ $\sum \alpha_i^2 = 1$
Outage	C5	$f(t)=A(1- \alpha(u(t-t_1)- (u(t-t_2))))\sin(\omega t)$	$0.9 \leq \alpha \leq 1; T \leq t_2 - t_1 \leq 9T$
Sag with harmonic	C6	$f(t)=A(1- \alpha(u(t-t_1)- (u(t-t_2))))(\alpha_1 \sin(\omega t) + \alpha_3 \sin(3\omega t) + \alpha_5 \sin(5\omega t))$	$0.1 \leq \alpha \leq 0.9; T \leq t_2 - t_1 \leq 9T$ $0.05 \leq \alpha_3 \leq 0.15, 0.05 \leq \alpha_5 \leq 0.15,$ $\sum \alpha_i^2 = 1$
Sag with swell	C7	$f(t)=A(1+ \alpha(u(t-t_1)- (u(t-t_2))))(\alpha_1 \sin(\omega t) + \alpha_3 \sin(3\omega t) + \alpha_5 \sin(5\omega t))$	$0.1 \leq \alpha \leq 0.9; T \leq t_2 - t_1 \leq 9T$ $0.05 \leq \alpha_3 \leq 0.15, 0.05 \leq \alpha_5 \leq 0.15,$ $\sum \alpha_i^2 = 1$

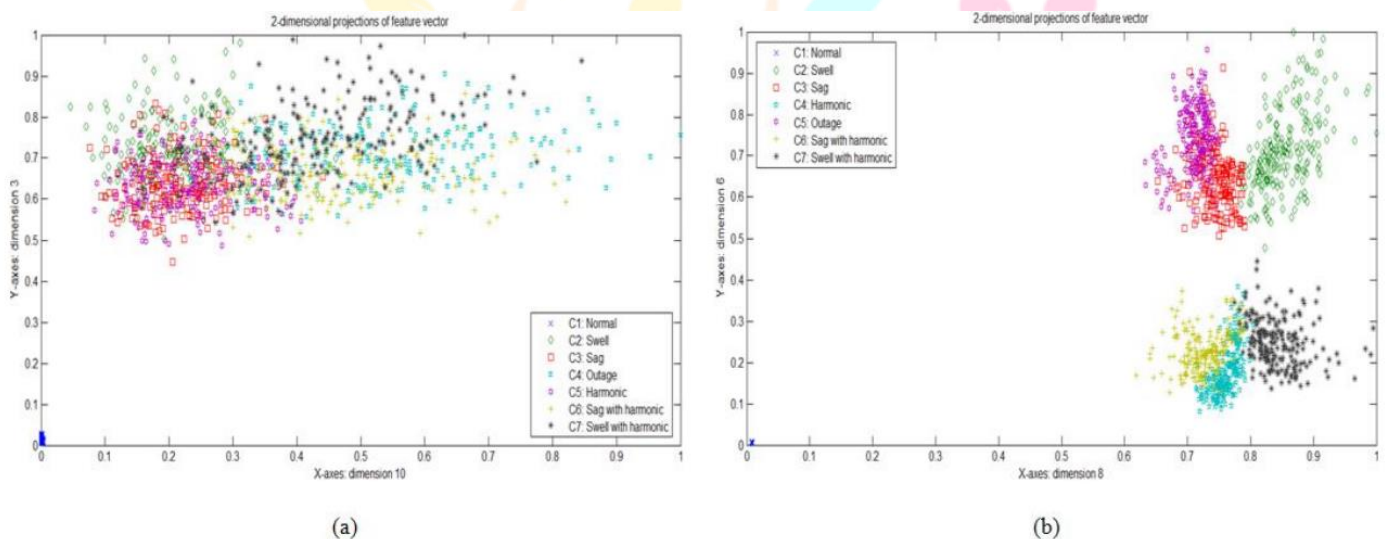


Fig. 5. Two-dimensional projection of the feature vector. (a) dimension 10 and dimension 3. (b) dimension 8 and dimension 6

4.2 Simulation Result and Analysis

Daubechie’s (Db4) Wavelet with 10 level decomposition were analysis to make 11 dimensions feature sets for training and testing based on Fig. 3. Taking 200 cases from each class we get a 1400×11 size of testing and training data. Each column of the data set was scaling 0 to 1 before being applied to the neural network. Now, Fig. 5(a) and Fig. 5(b) shows the two dimensional projections of the training set. Here, in Fig. 5(b) data are more separated.

Table 2 shows the simulation result for this seven-class PQ disturbance problem based on the proposed method and Table 3 shows the result reported in paper [17]. Here a 7×7 confusion matrix C is constructed to show the performance of the proposed method. In this 7×7 confusion matrix the diagonal elements indicates the correctly classified PQ types and the off-diagonal elements indicates the miss-classification. So we can see the method we proposed here is very effective to classify the different PQ disturbances.

Now here is something that we should analysis in order to get a detail performance of the proposed method,

- 1) The relationship between the classification performance and the decomposition level by which we can find the suitable decomposition level for effective classification.
- 2) The relationship between the classification performance and the wavelet by which we can find the suitable wavelet for effective classification

The following parts (4.3) and (4.4) will discuss these two issues.

4.3 Analysis about the classification performance and the decomposition levels

As we have discussed in section II that 1 level decomposition will provide the (1+1) dimension feature vector for future analysis and more levels of decomposition will increase the composition cost. So we should find a way to know a reasonable composition levels which will provide a good classification accuracy with satisfactory computational cost.

In this section, we choose Db4 wavelet to achieve our goal. Here for each PQ disturbance as mentioned in Table: I, 200 cases were generated for training and testing and evaluate the performance for levels 1 to 10. Fig. 6(a) shows the result for the proposed method and Fig. 6(b) shows the result for the method reported in paper [17].

	C1	C2	C3	C4	C5	C6	C7
C1	200	0	0	0	0	0	0
C2	0	200	0	0	0	0	0
C3	0	1	190	0	7	2	0
C4	0	0	0	200	0	0	0
C5	1	0	2	0	194	3	0
C6	0	0	0	2	1	196	1
C7	0	0	0	0	0	0	200
Overall Accuracy	98.57%						

	C1	C2	C3	C4	C5	C6	C7
C1	200	0	0	0	0	0	0
C2	0	200	0	0	0	0	0
C3	1	0	174	0	24	1	0
C4	0	0	0	200	0	0	0
C5	15	0	16	0	161	8	0
C6	0	0	2	1	2	194	1
C7	0	0	0	0	0	0	200
Overall Accuracy	94.93%						

From Fig. 6(a) we can see that when the decomposition level is small such as $l \leq 4$, the performance accuracy is poor. But when the decomposition level is higher $l \geq 4$ the classification accuracy is extremely good. Where from Fig.6(b) we can see that if we want to have a good classification accuracy (about 90%) we need decomposition level at least 6. So this comparison shows that the proposed method can have the ability to reduce the computational cost.

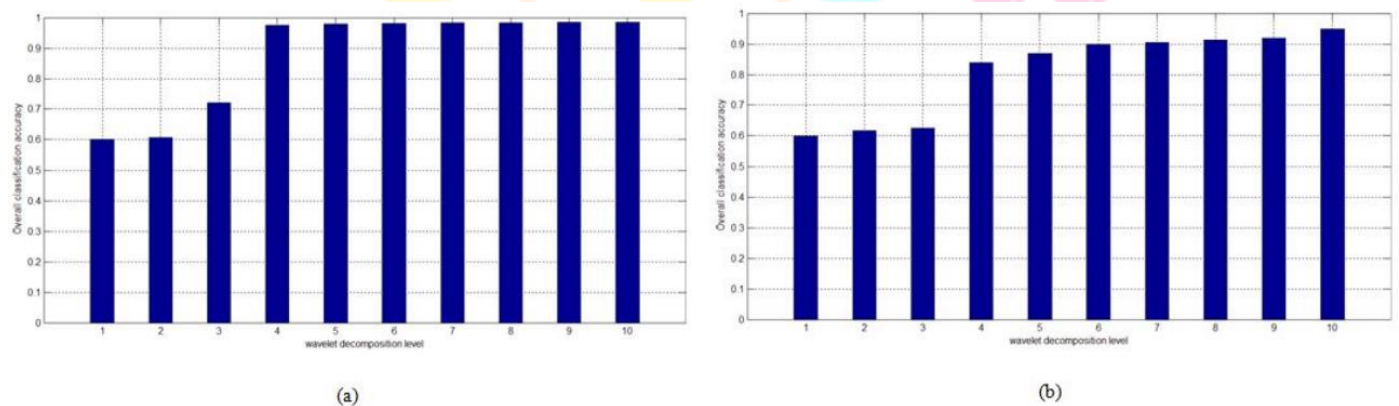


Fig. 6. Relationship between the wavelet decomposition levels and overall classification accuracy. (a) for the proposed method. (b) as reported in Paper [17].

4.4 Analysis about the classification performance and the wavelets

Now there is very important issue related with the proposed method and that is the choice of suitable wavelet. Of course the longer the wavelet filter length, the larger computational cost. In this section we will analysis the performance of the proposed method for different wavelet.

Assumption 1: No single wavelet transform has a significant advantage over other wavelets in performance of the proposed method.

Now to investigate the *Assumption 1*, four commonly used wavelet named Haar wavelet, Daubechies's wavelet, Symlets wavelet and coiflets wavelet were used. Table IV shows the characteristics of these wavelets. Here, Orthogonal means that the inner products of wavelet basis function are zero. The computational cost is measured by Filter length. It is the number of the filter coefficients. The longer the wavelet filter length means the larger computational cost. So it is very important factor to take account while a filter is designed. The frequency characteristics of the wavelet transform is determined by Compact support and the support width. These are used to measure the domain of the wavelet function with nonzero values on it. Here the compact support indicate that these nonzero values are only for a finite duration and the support width indicate the how long these nonzero values duration is. A wavelet with small compact support is fast to compute, but narrowness in time domain implies a large width frequency domain. Again a wavelet with large support width is smoother and has frequency resolution. Symmetry is much desired property for a wavelet. Symmetry prevent orthogonality. So from Table IV we can see that only Haar wavelet contains the property of orthogonality and symmetry with 2 coefficients. A detail discussion about the wavelet characteristics is found in paper [11] and [12].

Table 4 Wavelet Characteristics

Wavelet name	Orthogonal	Filter length	Compact support	Support width	Symmetry
Haar	Yes	2	Yes	1	Yes
Daubechies	Yes	2N	Yes	2N-1	Far from
Coiflets	Yes	6N	Yes	6N-1	Near from

Symlets	Yes	2N	Yes	2N-1	Near from
---------	-----	----	-----	------	-----------

Now to evaluate the performance of different wavelets, we use decomposition level 4 as we can see from Fig. 6(a). at level 4 with db4 wavelet the proposed method can classify the PQ disturbances with a good accuracy as well as small computational cost compared to higher levels of decomposition. Now as before here also 200 cases for each class are constructed for training and testing data for each wavelet form Table 1. For each wavelet we took 20 simulation runs to evaluate the performance. Table 5 shows the result.

After that we use hypothesis test of the average accuracy values to determine whether there is any significant difference in the performance among different wavelet families or not [18].

Wavelet	Avg. Accuracy, p	Wavelet	Avg. Accuracy, p	Wavelet	Avg. Accuracy, p
Haar (Db1)	0.9504	Coif1	0.9676	Sym1	0.9540
Db2	0.9800	Coif2	0.9800	Sym2	0.9794
Db3	0.9796	Coif3	0.9735	Sym3	0.9795
Db4	0.9690	Coif4	0.9821	Sym4	0.9667
Db5	0.9791	Coif5	0.9720	Sym5	0.9661
Db6	0.9769			Sym6	0.9697
Db7	0.9739			Sym7	0.9811
Db8	0.9751			Sym8	0.9793
Db9	0.9735				
Db10	0.9781				
Mean (μ)	0.9736		0.9750		0.9720
Standard deviation (σ)	0.0088		0.0059		0.0096

Wavelet name	Wavelet name	Z	Accept or Reject H_0
Daubechies	Coiflets	0.3841	Accept
Coiflets	Symlets	0.7129	Accept
Symlets	Daubechies	0.3618	Accept

From Table 5 we get the mean and the standard deviation values of each wavelet family. The mean μ and the standard deviation σ were calculated from the average accuracy values by the following equation:

$$\mu = \frac{1}{n} \sum_{i=1}^n p_i \quad \& \quad \sigma = \sqrt{\frac{\frac{1}{n} \sum_{i=1}^n p_i^2 - (\sum_{i=1}^n p_i)^2}{n(n-1)}} \quad (4.1)$$

Here p_i is the classification accuracy of different wavelet as presented in Table V and n is no. of wavelet in each family. For example in Table VI the no. of db wavelet is 10.

Now we formulate the hypothesis test:

$$\text{Null Hypothesis: } H_0 : \mu_1 = \mu_2 \quad \& \quad \text{Alternative Hypothesis: } H_1 : \mu_1 \neq \mu_2 \quad (4.2)$$

Here the test statistic for the z-test was calculated as follow:

$$Z = \frac{\mu_1 - \mu_2}{\sqrt{\frac{\sigma_1^2 + \sigma_2^2}{n_1^2 + n_2^2}}} \quad (4.3)$$

For a two-tailed test when the value of $|Z| > 1.96$, the results will be at a significant level of 0.05. So we will reject H_0 if $|Z| > 1.96$. Now from Table 6 we can see, we accept H_0 which directly means that there will be no significant difference in the classification performance if we choose different the wavelet families. Now the same test is performed within the one wavelet family and Fig. 7., shows the result.

So form the hypothesis test it has been shown that one can choose any wavelet as there is no significant change in classification for PQ disturbances. It means, the proposed method has a good ability in classification performance and has no restriction to choose a particular wavelet. We know from Table. 4 that different wavelets has different filter length. It has an important effect in computational cost as the computational cost increase with increase of filter length of wavelet. From the analysis result of the proposed method we can choose the shorter wavelet such as Haar wavelet, db2 wavelet etc to have a good classification accuracy with low computational cost.

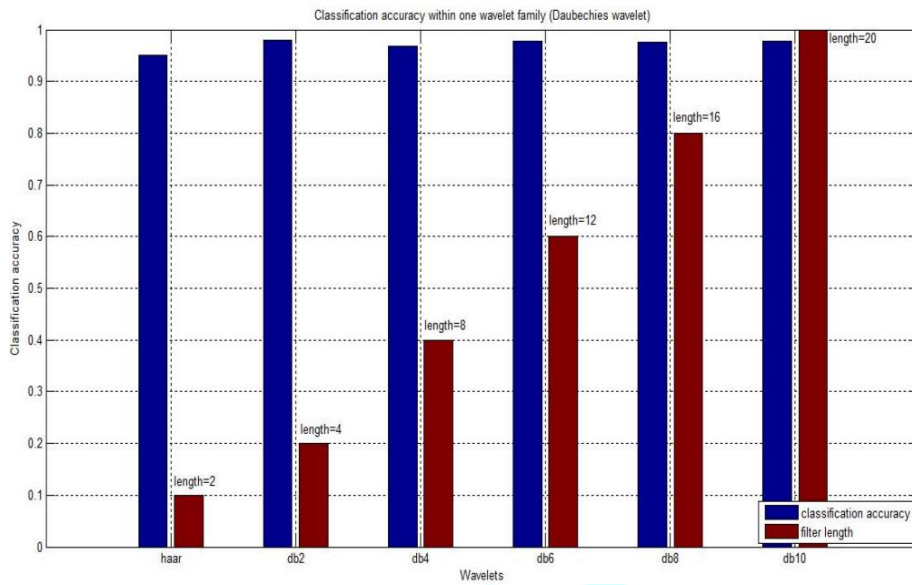


Fig. 7. Classification result and their corresponding wavelet filter length for different wavelets within one wavelet family (Daubechies Family)

V. PERFORMANCE IN THE NOISY CONDITION FOR THE PROPOSED METHOD

As we know noise is a very important factor in any electrical power distribution network, so now we will analysis whether the proposed method is good in a noisy environment or not.

In research of power quality issue the Gaussian white noise is widely considered [19], [20], [21]. Here we consider different signal to noise ratio (SNR) values ranging from 10dB to 50dB to test the performance of the proposed method in different noise condition. Now the SNR is defined as follow,

$$SNR = 10 \log \left(\frac{P_s}{P_n} \right) dB \quad (5.1)$$

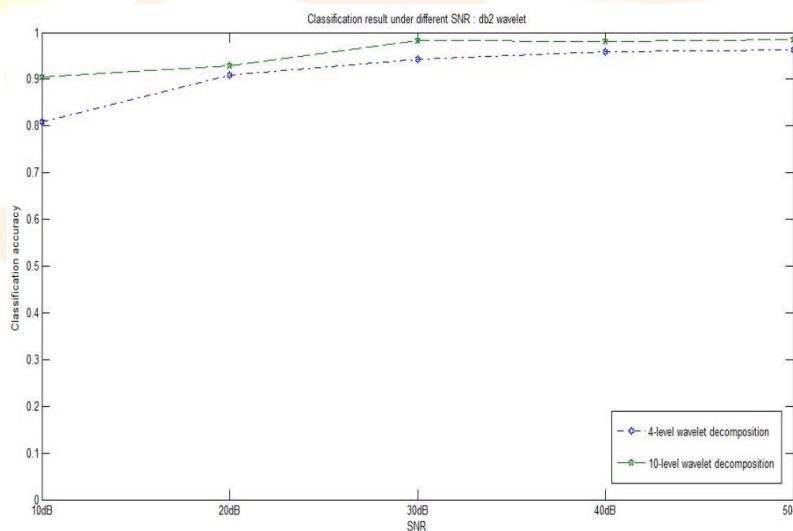


Fig. 8. Classification result under different SNR condition

Where P_s is the power (variance) of the signal and P_n is power (variance) of that noise. We used Monte Carlo method to generate training and testing data with different control parameters as shown in Table 1. 200 cases were considered for each class of PQ disturbances (C1 to C7). Fig. 8. shows the analysis result for 4-level decomposition and 10-level decomposition where db2 wavelet was chosen. From Fig. 8, we can see that even at a low SNR condition the proposed method can still achieve a good classification accuracy. Fig. 8. also show that the 10-level decomposition classification result is slightly better than 4-level decomposition result. So this result is consistent as the previous result in section 4.3.

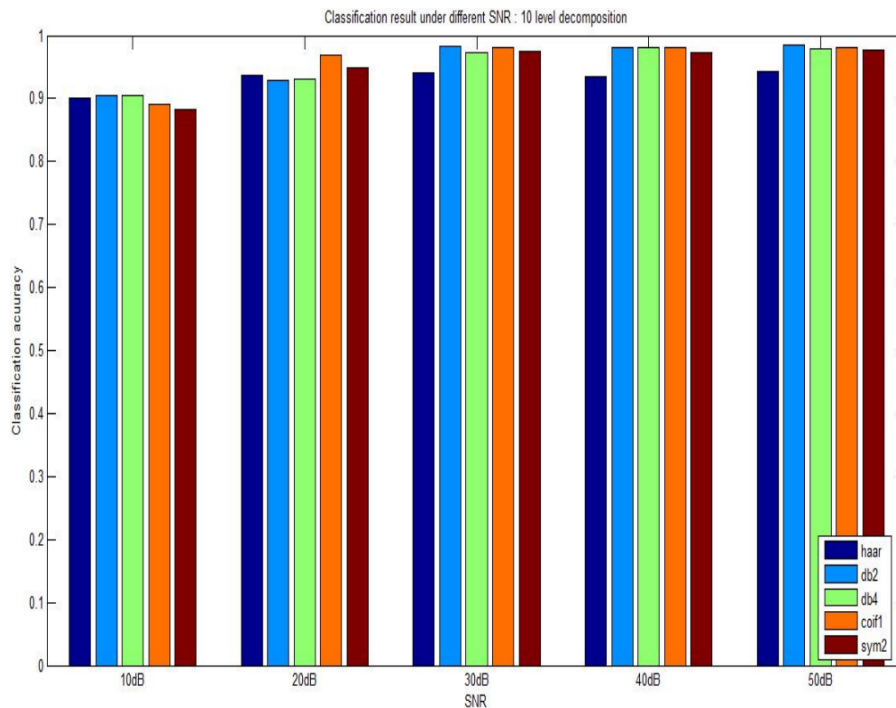


Fig. 9. Classification result under different SNR condition for different wavelets

We also test the performance for different wavelets. Fig. 9. shows the result for different wavelets. Here we choose 10-level decomposition in a noisy environment. Fig. 9. shows that there is no significant difference in the performance of the proposed method for different wavelets. So one can use different wavelet to achieve a good classification accuracy with low computational cost in a noisy environment too. This result is also consistent as our previous result in section 4.4.

VI. CONCLUSION

A PQ disturbances classification system based on Wavelet transformation and Back-propagation network is proposed in this paper. Here the training and testing data are contracted through Wavelet transformation based on multi-resolution analysis method. Then this training and testing data are applied to a Back-propagation network. Several types of PQ disturbance are taken in consideration in this paper. The classification result of the proposed method is compared to an existing literature report. This comparison shows that the proposed method can classify different types of PQ disturbance with an effective accuracy.

The classification result also shows that when different wavelets are chosen, there is no significant difference in the performance of the proposed method. This means the simplest wavelet to implement the proposed method can do as good a job as any other wavelet with good accuracy with low computational cost, at least for PQ disturbance classification problem. However, the analysis under different noise level shows that the proposed method can also do a good job under noisy environment too. This idea of combining the wavelet transformation and back propagation algorithm could potentially be applied in other sector such as this kind of problem.

ACKNOWLEDGMENT

The authors would like to thanks the reviewers for their helpful comments and suggestions.

REFERENCES

- [1] M. Samotyj, "The cost of power disturbance to industrial and digital economy companies," consortium for electrical infrastructure to support a digital society , an initiative by EPRI and the Electrical Innovation Institute, Jun. 2001. Primen.
- [2] W. R. A.Ibrahim and M. M. Morcos, "Artificial intelligent and advanced mathematical tools for power quality applications: A survey," IEEE Trans Power Del., vol. 17, no. 2, pp. 668-673, Apr.2002.
- [3] Ying-Tung Hsiao, Ying-Ming Wu, Yen-Hsing Lee and Fun Ye, " Fuzzy power quality indicator," Int. Confer. on Mach. Lear. And Cyber., Hong Kong, pp. 1418-1423 19-22 August 2007.
- [4] Bizjak and P. Planinsic. "Classification of power disturbances using fuzzy logic." In Proc. of IEEE Int. Conf. on Power Electronics and Motion Control, 2006.
- [5] Birendra Biswal, P. K. Dash and B. K. Panigrahi, "Power Quality Disturbance Classification Using Fuzzy C-Means Algorithm and Adaptive Particle Swarm Optimization," IEEE Trans. on Indust. Electr., vol. 56, no. 1, pp. 212-220 Jan 2009.

- [6] Liao and J.-B. Lee. "A fuzzy-expert system for classifying power quality disturbances. *Electrical Power and Energy Systems*," Elsevier Science, 26:199–205, 2004.
- [7] R. N. M. Machado, U. H. Bezerra, E. G. Pelaes, R. C. L. D. Oliveira and M. E. D. L. Tostes, "Use of Wavelet Transform and Generalized Regression Neural Network (GRNN) to the Characterization of Short-Duration Voltage Variation in Electric Power System," *IEEE Latin America Trans.*, vol. 7, no. 2, pp. 217-222, June 2009.
- [8] U. D. Dwivedi and S. N. Singh, "Denoising Techniques With Change-Point Approach for Wavelet-Based Power- Quality Monitoring." *IEEE Trans. on power delivery*, vol. 24, no. 3, pp.1719-1727, July 2009.
- [9] Dimce Kostadinov and Dimitar Taskovski, "Automatic voltage disturbance detection and classification using wavelets and multiclass logistic regression," EC project SEE ERA NET PLUS under grant ERA 078/01.
- [10] Cong-Hui Huang, Chia-Hung Lin, and Chao-Lin Kuo, "Chaos Synchronization-Based Detector for Power-Quality Disturbances Classification in a Power System," *IEEE Trans. on power delivery*, vol. 26, no. 2, pp. 944-953, APRIL 2011.
- [11] I. Daubechies, "Ten lecture on wavelets," Philadelphia, PA: Society for Industrial and Applied Mathematics, 1992.
- [12] C. S. Burrus, R. A. Gopinath and H. Gauo, "Introduction towavelets and wavelet transforms: A Primer ," Englewood Cliffs, NJ: Prentice-Hall, 1998.
- [13] S. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation ," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 7, pp. 674-693, jul.1989.
- [14] Yuhua Peng, "Wavelet Transform and Engineering Application." Beijing: Science Press, 1999.
- [15] M. K. S. Alsmadi, K. B. Omar, and S. A. Noah, "Back Propagation Algorithm: The Best Algorithm Among the Multi-layer Perceptron Algorithm", *Int. Journal of Computer Science and Network Security*, vol.9, no. 4, April 2009.
- [16] X. Zhou, J. Zheng, and K. Mao, "Application of Modified Backpropagation Algorithm to the Prediction of the Chloride Ion Concentration in Cracked Concrete", *In the Proc. of 3rd International Conference on Natural Computation*, pp. 257-261, 2007.
- [17] Haibo He and Janus A. Starzyk, "A self-organizing learning array system for power quality classification system based on wavelet transformation," *IEEE trans. On power delivery*, vol. 21, no. 1, pp. 286-295, Jan. 2006.
- [18] I. Miler and J. E. Fruend, "Probability and Statistics for Engineers," *Englewood Cliffs, NJ: Prentice-Hall*, 1965.
- [19] H. T. Yang and C-C Liao, "A de-noising scheme for enhancing wavelet based power quality monitoring system ," *IEEE Trans. Power del.*, vol. 17, no. 1, pp. 353- 360, Jul. 2001.
- [20] H. Zhang, P. liu and O. P. Malik, "Detection and classification of power quality disturbances in noisy conditions," *Proc. Inst.Electr. Eng. Gener. Transm. Distrib.*, vol. 150, pp. 567-672, Sep. 2003.
- [21] P. O'Shea, "A high-resolution spectral analysis algorithm for power system disturbance monitoring," *IEEE Trans. Power Syst.*, vol. 17, no. 8, pp. 676-680, Aug. 2002

