Wavelet transform based power quality events classification using artificial neural network and SVM

D. Saxena[1]*, K.S. Verma[2]

*[1]*Department of Electrical and Electronics Engineering, Invertis Inst. of Engg.& Tech., Bareilly (UP), INDIA*
[2] Department of Electrical Engineering, , K.N.I.T Sultanpur (UP), INDIA.
E-mails: diptisx@gmail.com (D. Saxena[1], [1]Corresponding author), ksv211@rediffmail.com (K.S. Verma)

Abstract

This paper demonstrates classification of PQ events utilizing wavelet transform (WT) energy features by artificial neural network (ANN) and SVM classifiers. The proposed scheme utilizes wavelet based feature extraction to be used for the artificial neural networks in the classification. Six different PQ events are considered in this study. Three types of neural network classifiers such as feed forward multilayer back propagation (FFML), learning vector quantization (LVQ), probabilistic neural network (PNN) are analyzed for effective classification of PQ events. The results show the superiority of PNN over FFML and LVQ. The test simulations show that SVM has higher performance than ANN with feed forward multilayer back propagation (FFML), learning vector quantization (LVQ), probabilistic neural network (PNN).

Keywords— Power Quality, Daubechies, Multi Resolution Analysis, Feature Extraction, Neural networks, SVM

1. Introduction

In the emerging power systems, power quality (PQ) issues have attained considerable attention in the last decades due to increased penetration of power electronics and/or microprocessor based non-linear controlled loads. On one hand, these devices introduce power quality problem and on other hand, these devices mal-operate due to the induced power quality problems. Large penetrations of power electronics based controllers and devices along with restructuring of the electric power industry and small-scale distributed generation require more stringent demand on the quality of electric power supplied to the customers (Arrillaga *et al*, 2000a; Arrillaga *et al*, 2000b; Dugan *et al,* 2003). The utilities or other electric power providers have to ensure a high quality of their service to remain competitive and to retain/attract the customers in new electricity market scenarios. Electricity is now treated as commercial product that is evaluated not only by its reliability but also by its quality. The customer will choose the supplier providing electrical energy having better power quality, at lower cost and acceptable reliability that meet his load needs.

PQ disturbances/events cover a broad frequency range with significantly different magnitude variations and can be of stationary or non-stationary signals. The on-line detection and identification of disturbances are essential to know the sources and cause of such disturbances for appropriate mitigation/corrective actions. A feasible approach to achieve this goal is to incorporate detection capabilities into monitoring equipment so that the events of interest can be recognized, captured and classified automatically. Hence, good performance monitoring equipment must be capable of detecting, localizing and classifying PQ events for minimizing their impact by devising suitable corrective and preventive measures.

Artificial neural network (ANN) systems with their unique ability to learn from examples and to generalize, i.e., to produce reasonable outputs for the new inputs not encountered during a learning process, can provide an effective solution to cope with such problems (Santoso *et al,* 1996), (Gauda *et al,* 1999). The distinct features of ANN as learning from examples, generalization ability, non-linearity of processing units, adaptability, massive parallel interconnection among processing units and fault tolerance make it a good choice as robust method for automatic classification. ANNs had attracted a great deal of attention because of their

inherent pattern recognition capabilities and their ability to handle noisy data. However, the choices of extracted feature parameters (magnitude, duration, frequency component, or waveform shape) as well building the ANN system are the vital factors that decide the complexity of the classifier structure. Although the ANN has been found very effective for the PQ disturbance classification, this technique has few limitations of local optimum and over fitting. SVM, which is quite a new method and also used in this paper, can overcome these deficiencies and provide efficient and powerful classification algorithms that are capable of dealing with high-dimensional input features and with theoretical bounds on the generalization error and sparseness of the solution provided by statistical learning theory (Vapnik, 1998; Cristianini *et al*, 2000). The features extracted by signal processing techniques are used as input to the PQ classification system. Feature extraction can be defined as transformation of the raw signal from its original form to a new form, from which suitable information can be extracted. An appropriately chosen feature set reduces the burden over the classifiers. Therefore, analyzing PQ related problems demands an effective signal processing technique.

Wavelet transform (WT) has emerged as a powerful signal processing tool used in the power quality analysis (Pillay *et al,* 1996; Robertson *et al, 1996;* Tunaboylu *et al,* 1996; Angrisani *et al,* 1996). The WT has the capability to analyze different power quality problems simultaneously in both time and frequency domain, hence giving the time frequency representation of signal like the short time Fourier transform (STFT). Unlike the STFT which uses a fixed window function, the wavelet transform makes use of a varied time frequency window, whose length depends on the frequency analyzed using long windows at low frequencies and short windows at high frequencies (Rioul *et al,* 1991). Therefore, the WT at low frequency provides accurate frequency resolution and poor time location, and at high frequency, WT gives accurate time location and bad frequency resolution. This feature of WT makes it very effective for real signals such as voltage sags and transient over voltages.

Wavelet based online disturbance detection for power quality applications are efficient as compared to conventional approaches in terms of speed and precision discrimination in the type of transient event (Karimi e*t al, 2000).* In addition to detection of PQ events, wavelet transform is also effective in power system protection (Charri *et al,* 1996), detection of high impedance faults (Michalik *et al, 2006*), and PQ data compression (Santoso *et al,* 1997).

The selection of suitable features is extremely important for classification of any problem. An appropriately chosen feature set reduces the burden over the classifiers. It is essential to extract a useful feature vector that can reduce data size as well incorporate the main characteristics of signal to build an efficient and robust classifier. Hence, in this paper, a wavelet feature extraction technique based on the energy distribution of the disturbances at different levels using multi-resolution analysis which is unique for a disturbance is proposed for automatic PQ disturbances classification. The disturbance classification schema is performed with wavelet-neural network (WNN) which utilizes feature from WT energy function and a neural network classifier. The featured extracted from wavelet transform are given to the neural network for training and subsequently, it is tested for an effective classification. WNN is applied on a set of different PQ disturbances, such as pure sine (normal), sag, swell, outage, harmonics, and sag with harmonic and swell with harmonic. Three types of well known neural network classifiers, namely, feed-forward multilayer back propagation (FFML), learning vector quantization (LVQ), probabilistic neural network (PNN) are used for comparison with proposed WNN. The same data is again classified using support vector machines for comparison of classification performance between WNN and SVM.

The paper is organized as follows. Section 2 gives a general introduction of the wavelet transform and multi-resolution analysis. Section 3 describes wavelet-neural network for classification of PQ events. Section 4 gives a brief overview of SVM (support vector machines), the other classifier used for classification of PQ events. Section 5 gives the methodology of the proposed scheme. Section 6 gives the results and finally conclusion and references are given in section 7 and section 8.

2. Wavelet transform and multi-resolution analysis

Fourier analysis consists of breaking up a signal into sine waves of various frequencies. Similarly, wavelet analysis is the breaking up of a signal into shifted and scaled versions of the original (or mother) wavelet. The resulting wavelets, called daughter wavelets, are localized both in time and frequency. Thus, wavelet transform provides a local representation of signal in both time and frequency unlike Fourier transform which gives a global representation of signal in terms of frequency. Continuous wavelet transform (CWT), wavelet series (WS) and discrete wavelet transform (DWT) are three ways by which wavelet transform can be implemented.

Continuous wavelet transform of the signal *x(t)* using the analysis wavelet ψ(.)

$$CWT_x^{\psi}(a,b) = \Psi_x^{\psi}(a,b) = \frac{1}{\sqrt{|a|}} \int_t f(t) \psi^* \left(\frac{t-b}{a} \right) dt \qquad (1)$$

where *a* is scale parameter, *b* is translation parameter and $1/\sqrt{a}$ is a normalization constant, *ψ* is the mother wavelet. All kernels are obtained by translating (shifting) and/or scaling the mother wavelet .The oscillatory frequency and the length of wavelet is decided by the scale parameter, *a,* and the shifting position of wavelet is determined by the translation parameter, *b*.

Discrete wavelet transform is implemented using discrete values of the scaling parameter and the translation parameter. This is achieved by replacing a by a_0^m and *b* by $nb_0 a_0^m$ where *m* is the translation step indicating frequency localization and *n* is scaling step indicating time localization.

$$DWT\,(m,n) = \frac{1}{\sqrt[m]{a_0}}\sum_k f(k)\Psi(\frac{k - nb_0 a_0^m}{a_o^m}) \tag{2}$$

Wavelet transform is called a dyadic-orthonormal wavelet transform when $a_o = 2$ and $b_o = 1$ and this results in a very efficient algorithm, known as multi-resolution signal decomposition technique. The muli-resolution analysis (MRA) introduced by Mallat (Mallat 1989) decomposes a signal into scale with different time and frequency resolution. In MRA a signal $f(t)$ can be completely decomposed into its detailed version (high frequency components) and smoothed/ approximated versions. The wavelet function serving as high pass filter with filter coefficients $h(n)$, generate the detailed version of the distorted signal, while the scaling function associated with low pass filter with filter coefficient $g(n)$, generate the approximated version of the distorted signal. Thus, by using MRA high frequency transients can be easily analyzed in presence of low frequency components such as non-stationary and non-periodic wide-band signals. MRA can be implemented by a set of successive filter banks as shown in Fig 1.
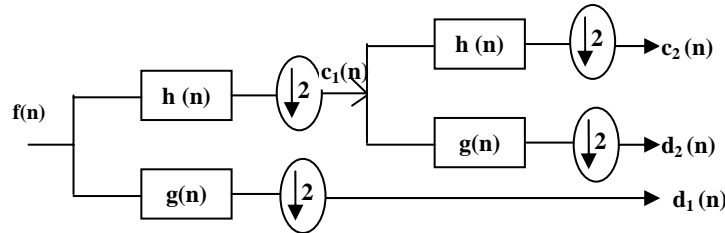

Fig. 1. Decomposition of *f(n)* into two scales

The $h(n)$ and $g(n)$ are the low pass and high pass filters. Let $f(n)$ is the discrete time signal, from MRA the decomposed signals at scale 1 are $c_1(n)$ and $d_1(n)$, where $c_1(n)$ is the smoothed version of the original signal, and $d_1(n)$ is the detailed version of the original signal down-sampled by a factor 2. Since both the high pass filter and low pass filters are half band, these decomposition halves the time resolution, since now only half the number of sample characterize the entire signal. However, this operation doubles the frequency resolution since the frequency band of signal now spans only half the previous frequency band, effectively reducing the uncertainty in the frequency by half. The next higher scale decomposition is now based on the signal $c_1(n)$, which decomposes it further into $c_2(n)$ and $d_2(n)$. At each scale, the filtering and sub sampling result in half the number of samples and thus half the time resolution and double the frequency resolution.

The choice of mother wavelet plays a significant role in detecting and localizing various types of disturbances. Daubechies' wavelets with 4, 6, 8, and 10 filter coefficients work well in most of the disturbance detection cases. At the lowest scale (scale 1), the mother wavelet is the most localized in time and oscillates most rapidly within a very short period of time. As the wavelet goes to higher scales, the analyzing wavelets become less localized in time and it oscillates less due to the dilation nature of the wavelet transform analysis. As a result of higher scale signal decomposition, the fast and short power quality disturbances will be detected at lower scales, whereas slow and long power quality disturbances will be detected at higher scales. Hence, both fast and slow power quality disturbances can be detected. Since Daub4 has the least number of filter coefficients and it gives the shortest support. Daub4 has been used in the work.

3. Artificial neural network

An artificial neural network (ANN) is an information-processing paradigm that is inspired by the biological nervous systems, such as the brain, process information (Bishop, 1995). The key element of this paradigm is the novel structure of the information processing system composed of a large number of highly interconnected processing elements (neurons) working in unison to solve the specific problems. Neural network is a non-linear, data driven self adaptive method and is a promising tool for classification. These can adjust themselves to the data without any explicit specification of functional or distributional form for the underlying model (Bishop, 1995). The neural network recognizes a given pattern by experience which is acquired during the learning or training phase when a set of finite examples is presented to the network. ANNs are not programmed in the conventional sense; rather they learn to solve the problem through interconnections with environment.

Fig. 2 shows a typical architecture of three-layer WNN with only single output, where $\mathbf{x}=(x_1, x_2,\ldots, x_n)\in \mathbb{R}^n$ (n-dimensional space of real numbers) is an input vector, $\mathbf{y}$ is the output vector, $\mathbf{w}=(w_1, w_2, \ldots, w_n)$ is weight matrix of input layer to output layer connections, $\mathbf{v}=(v_1, v_2, \ldots, v_m)$ is weight matrix of hidden layer to output layer connections and $\mathbf{b}=(b_1, b_2, \ldots, b_k)$ is bias vector of the output layer neurons.

Wavelet input feature based neural network classifier recognizes the waveform in the wavelet domain using neural networks. It does so by efficiently integrating the aspects of the wavelet transformation for the feature extraction with the characteristic decision capabilities of neural network approaches for classification (Santoso *et al*, 2000; Perunicic *et al*, 1998; Angrisani *et al*, 1998; Borras *et al*, 2001). In the preprocessing stage, the detailed energy distribution at different levels is calculated using multi-resolution analysis, and the post processing stage is implemented by using neural networks. The proposed NN based PQ classifier

is shown in Fig. 3. The input to the classifier is the energy features and the output is the type of the disturbance along with its corresponding degree of belief.
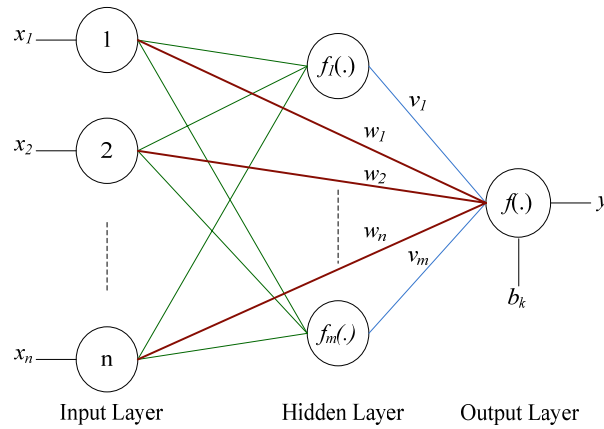

Fig. 2. A three layer artificial neural network (Jain *et al.* 2011)
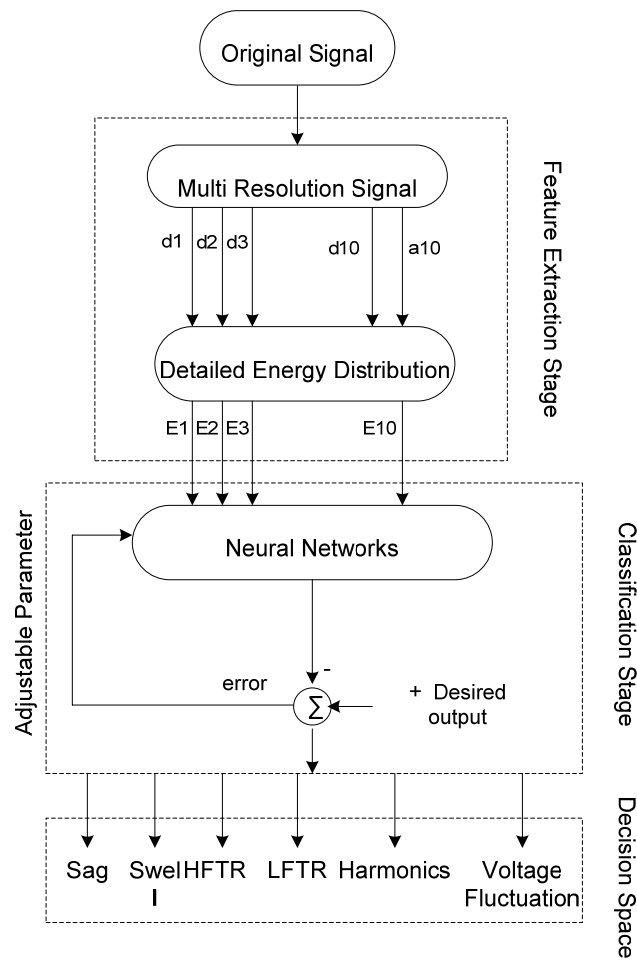

Fig. 3. Proposed NN structure model

4. Support vector machines

SVM is one of the relatively new and promising methods for learning separating functions in pattern recognition (classification) tasks as well as performing function estimation in regression problems (Vapnik, 1998; Cristianini *et al*, 2000; Bishop, 2008). It is originated from learning systems derived from the *statistical learning theory* introduced by Vapnik and co-workers for "distribution free learning from data". Learning is formulated as an optimization problem where classification of data is done by determining a set of support vectors, which are members of the set of training inputs that outlining hyper plane. SVM are able to find non linear boundaries if classes are linearly non separable. The main issue of interest in using SVMs for classification is the generalization performance and complexity of the classifiers, which is a practical implementation concern. Classifier also concerns with the problem of over fitting or under fitting. Another issue of concern is computational complexity. SVM classifier minimizes the generalization error on test set under Structural Risk Minimization (SRM) principle.

Main objective of the SRM principle is to choose the model complexity optimally for a given training sample. The input space in a SVM is nonlinearly mapped onto a high dimensional feature space. The result is that the classes are more likely to be separable than in a low-dimensional feature space. The idea is to map the feature space into a much bigger space so that the boundary is linear in the new space. Generally linear boundaries in the enlarged space achieve better training-class separation, and it translates to non-linear boundaries in the original space. Another important attraction of SVMs is the use of kernels. Kernel representations offer an alternative solution by nonlinearly projecting the input space onto a high dimensional feature space. A SVM classifier employs a kernel function $k(x_i,x)$ relating the support vectors x_i (which is small subset of training samples) and the input vectors drawn from the test data. The main purpose of the SVM algorithm used for classification is to construct an optimal decision function that accurately predicts unseen data into two classes and minimizes the classification error. This is achieved by structural risk minimization (SRM) which states that the classification error expectation R of unseen data is bounded by the sum of a training error rate and a term that depends on the Vapnik–Chevornenkis (VC) dimension (Vapnik, 1998).

$$R < \frac{t}{N} + \sqrt{\frac{h(\ln(\frac{2N}{h})+1) - \ln(\frac{\eta}{4})}{N}} \qquad (3)$$

where t is the number of training errors, N is the number of training samples, and R is a confidence measure.

The function in (4) is the decision boundary and is derived from a set of training samples

$$X = \{x_1, x_2, \dots x_n\}, x \in \Re^M \qquad (4)$$

where each training sample x_i has M features describing a particular signature and belongs to one of two classes

$$Y = \{y_1, y_2, \dots, y_n\}, y \in \{-1,1\}. \qquad (5)$$

The decision boundary is a hyper-plane

$$f(x) = <w, x> + b \qquad (6)$$

where w and b are derived in such a way that unseen data is classified correctly. This is achieved by maximizing the margin of separation between the two classes.

The SVM method finds the oriented hyperplane that maximizes the closest distance between observations from the two classes (the so-called "margin"), while at the same time minimizes the amount of training errors. In this way, good generalization ability of the resulting function is achieved, and therefore the problem of over fitting is mitigated. According to (Michalik *et al*, 2006) this can be formulated as an optimization problem subjected to the constraint that all training samples are correctly classified. Matlab SVM toolbox has been used.SVM finds the optimal separating hyperplane expressed as $\mathbf{w}^{*\mathrm{T}}.\mathbf{x} + b^* = 0$, that maximizes the margin as well as minimizes the number of misclassified patterns. The optimal weight vector $\mathbf{w}^*$ is given as

$$w^* = \sum_{j=1}^{N} \lambda_j^* x_j \qquad (7)$$

where $\lambda^* = (\lambda_1^*, \lambda_2^*, \lambda_3^* \dots, \lambda_N^*)$ is a soloution of quadratic programming problem. Support vectors are the points x_i with $\lambda^* > 0$. The classification of a new data vector $\mathbf{x}$ involves the evaluation of the decision function

$$y = sign(f(\mathbf{x})) \qquad (8)$$

where, $f(\mathbf{x})$ is the optimal decision boundary derived from the set of training samples and is given by

$$f(\mathbf{x}) = \mathbf{w}^{*\mathrm{T}} \mathbf{x} + b^* \qquad (9)$$

The above equation can be expressed as

$$f(\mathbf{x}) = \left(\sum_{j=1}^{N} \lambda_j^* \mathbf{x}_j \right) \mathbf{x} + b^* \tag{10}$$

Then the class $y \in \{-1, 1\}$ of $\mathbf{x}$ is expressed evaluating the dot product between the data and some elements (support vectors) of the training set S.If data is linear, a separating hyper plane may be used to divide the data. However, most of the practical data is nonlinear and inseparable and therefore, kernels are used to non-linearly map the input data to a high-dimensional space. The new mapping is then linearly separable (Cristianini *et al,* 2000). The same construction discussed above can be extended to the case of non-linear separating surfaces, where each point $\mathbf{x}i$ in the input space is mapped into a higher dimensional space, termed as feature space, $\Phi(\mathbf{x})$ for which optimal separating hyperplane is searched. The mapping is done in such a way that the dot product of two points in the feature space $(\Phi(\mathbf{x}). \Phi(\mathbf{y}))$ can be rewritten as kernel function *K(x, y).* Then the optimal separating hyperplane in the feature space can be written as a non-linear separating surface in the input space.

$$f(\mathbf{x}) = \sum_{j=1}^{N} \lambda_j^* \ K(\mathbf{x}_i, \mathbf{x}) + b^* \tag{11}$$

The above equation is represented as linear combination of kernel functions centered on the support vectors only. Because *f(x)* does not depend on the dimensionality of the feature space, SVM classifier represents an attractive choice for classification problem. Any function as a kernel can be used if it follows Mercer's condition. Polynomial kernel, radial basis function and multi-layer perceptions are some commonly used kernel functions. In this paper radial basis function (RBF) which gives the best result and most widely used kernel is used. It is expressed by the equation

$$K(x, y) = \exp\left(-\frac{\|x - y\|^2}{2\sigma^2} \right) \tag{12}$$

where σ is the width of the Gaussian function.

5. Methodology

The classification structure is composed of preprocessing feature extraction stage and the post processing classification stage, using two type's classifiers namely neural network and SVM based.

5.1 Wavelet Based Feature Extraction

Feature extraction plays a vital role in designing the intelligent system based on pattern recognition. The performance of the classifier is adversely affected if the extracted features are not chosen well. Moreover, to enhance computing efficiency and accuracy of recognizing a PQ event the feature extractor must reduce the original signal to a lower dimension which contains the most useful information of the original signal.

DWT decomposes the input signal into a group of different frequency levels, each corresponding to a particular frequency band. Therefore, the wavelet technique discriminates disturbances from the original signal, and then analyses them separately. The discontinuities in the signal due to disturbances in the form of sharp edges, transitions and jumps are reflected in the higher frequency bands .Thus any change in the smoothness of signal can be detected and localized at the finer resolution level. The wavelets coefficients of the finer resolution level will have high magnitude at the start and the end point of disturbance, when DWT is used to analyze a distorted signal through l level decomposition of MRA. Therefore, the start t_s and the end time t_e easily obtained from the variation in wavelets coefficients and the time duration t_t of the disturbance can hence be calculated as

$$t_t = |t_e - t_s| \tag{13}$$

Next important task after the detection and localization of the PQ event is the extraction of features which can be used for their classification. The squared wavelet coefficients at different resolution levels, the standard deviation and/or mean value and the maximum modulus of the wavelet coefficients at different resolution level have been used. The energy of a distorted signal can be partitioned at different resolution levels in different ways depending on the power quality problems. The energy distribution of a distorted signal can be used as a discriminatory feature for classification. The energy of a distorted signal at a resolution level j is given by

$$P_j = \sum_{i=1}^{N_m} (d_m(i))^2 \tag{14}$$

where N_m is the number of available wavelets coefficient at the resolution level j. The energy distribution of the disturbances at different levels using multi-resolution is unique for a disturbance and can be used as a feature for automatic classification of the power quality events as shown in Fig. 3.

5.2 Classification Stage

In the classification state, the detailed energy distribution features up to 10-level decomposition, of each signal are fed as input to neural network. Three types of well known neural network classifiers, namely, feed forward multilayer back propagation (FFML), learning vector quantization (LVQ), probabilistic neural network (PNN) have been used for comparison. The same data was again tested using SVM classifier.

6. Results and discussion

6.1 Data Generation

The simulation data is generated in MATLAB based on the model in paper (Rodney *et al,* 2010). One pure sine-wave signal (frequency = 50 Hz, amplitude 1.0 p.u) and 480 distorted signals ie 80 signals for each of the six PQ disturbance signals including voltage sag, voltage swell, harmonics, low frequency transient, high frequency transient, voltage fluctuation are generated. The disturbance signals given in Table I show the signal generation models and their controlled parameters. The detailed energy distribution up to 10-level decomposition using db4, of each signal is obtained.

Figs. 4(a)-4(c) show the first three detailed versions of a three-level decomposition and detailed energy distribution of a pure sine wave. Any changes in the pattern of signal can be detected and localized at the finer resolution levels. As far as detection and localization is concerned, the wavelet coefficients of the first finer decomposition level of *f(t)* are normally adequate to detect and localize any disturbance in the signal. As seen in Fig. 4(c), for a pure signal, the set of coefficients at the finer first three-resolution level is zero. Since the signal is smooth the duration of disturbance could not be detected. Any changes in the signal can be detected and localized in time due to changes in the magnitude of these coefficients.

Figs. 4(d)-4(f) show the detailed version of four-level decomposition and the detailed energy distribution ($P_1^D - P_{10}^D$) of a voltage sag signal which is detected and localized at first three finer resolution levels. Various PQ events such as voltage swell, voltage harmonic, low frequency transients and voltage fluctuations along with their detail energy distributions are shown in Fig. 5(a)-5(h).
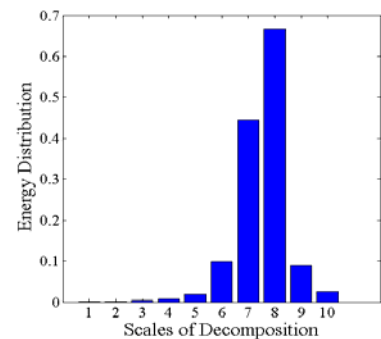
Table 1. Signal Models and their parameters

PQ disturbance	Model	Parameters
Sine -wave	$x(t) = A\sin(\omega t)$	A=1.0
Sag	$x(t) = A(1-\alpha(u(t-t_1)-u(t-t_2)));$ $\quad$ $t_1 < t_2, u(t)=1, t \geq 0$	$0.1 \leq \alpha \leq 0.9$, $T \leq t_2 - t_1 \leq 8T$
Swell	$x(t) = A(1+\alpha(u(t-t_1)-u(t-t_2)));$ $\quad$ $t_1 < t_2, u(t)=1, t \geq 0$	$0.1 \leq \alpha \leq 0.8$, $T \leq t_2 - t_1 \leq 8T$
Harmonics	$x(t) = A[\sin \omega t + \alpha_3 \sin(3\omega t) + \alpha_5 \sin(5\omega t)]$	$0.1 \leq \alpha_3 \leq 0.2$, $0.05 \leq \alpha_5 \leq 0.1$
Transients	$x(t) = \sin(\omega t) + ae^{-k}\sin(b\omega t);$ $\quad$ where $k = \dfrac{t}{\lambda}$ $20 \leq b \leq 80;$ for high frequency transeint $5 \leq b \leq 20;$ $\quad$ for low frequency transeint	$0.1 \leq a \leq 0.9$ for transient period $a = 0$ otherwise $0.1 \leq \lambda \leq 0.2$
Fluctuation	$x(t) = (1 + a\sin(b\omega t))\sin(\omega t)$	$0.1 \leq a \leq 0.2$, $0.4 \leq b \leq 0.6$
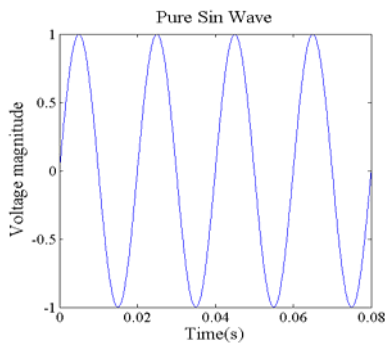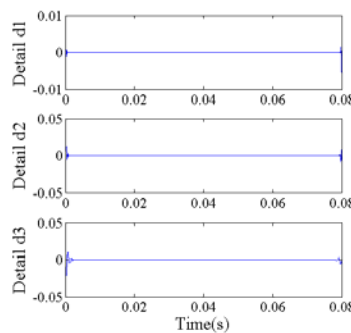
(a) Pure sine wave   (b) Three-level decomposition of (a)   (c) Detailed energy distribution of (a)
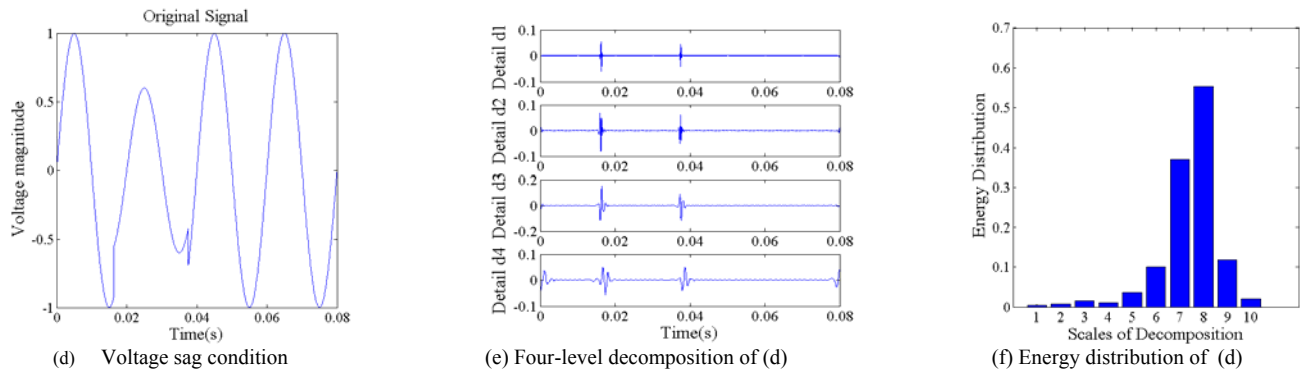
(d)   Voltage sag condition   (e) Four-level decomposition of (d)   (f) Energy distribution of (d)

Fig. 4. MRA decomposition and detailed energy distribution of a pure sine wave and voltage sag

(a) Voltage swell case   (b) Detailed energy distribution of (a)   (c) Voltage harmonic case   (d) Detailed energy distribution of (c)

(e) Low frequency transient case   (f) Detailed energy distribution of (e)   (g) Voltage fluctuation case   (h) Detailed energy distribution of (g)
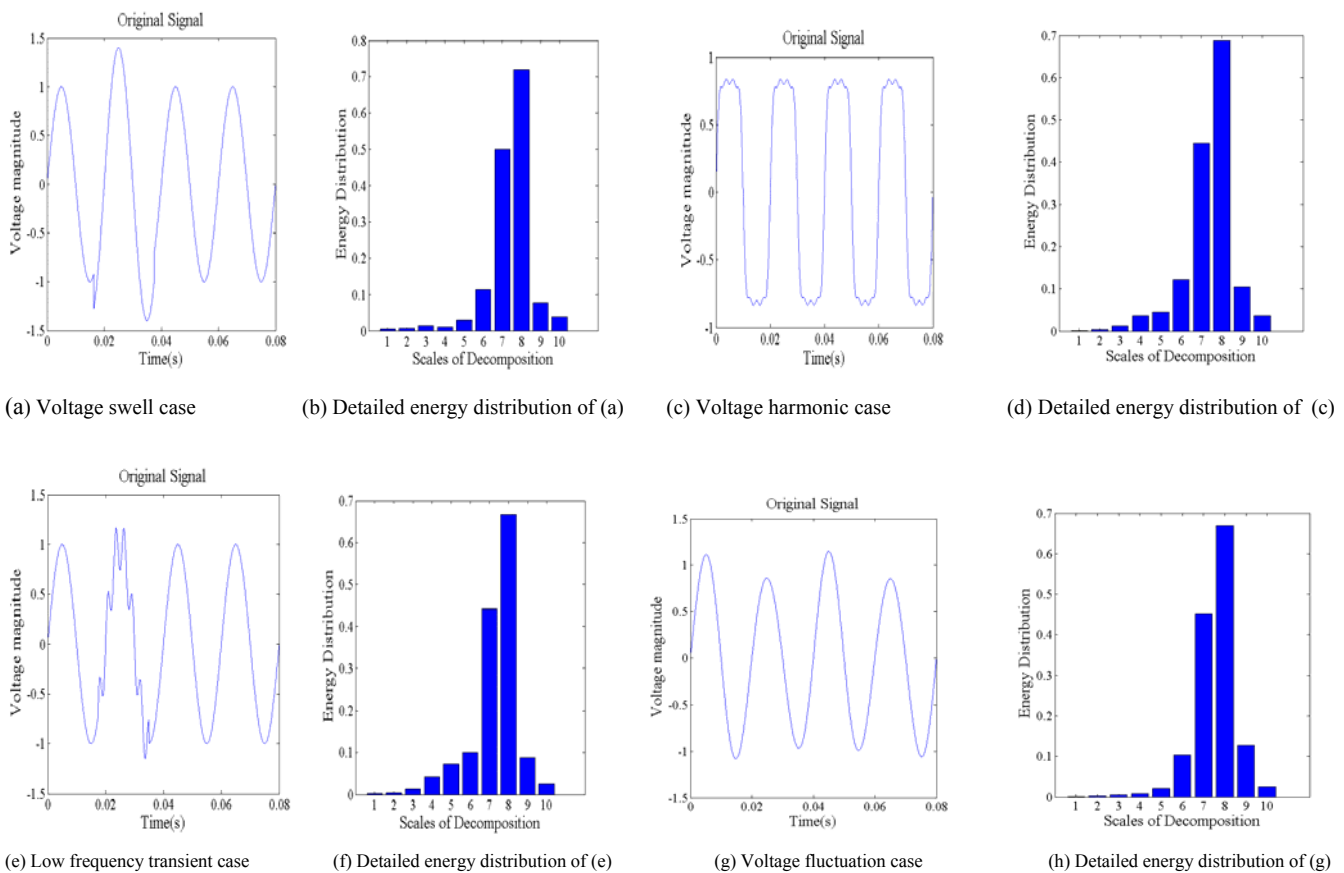
Fig. 5. Variation in detailed energy distribution for different PQ events.

6.2 Simulation Results

Each disturbance signal is decomposed with the Daubechies (db4) wavelet filter for 10 levels. Based on the feature extraction with detailed energy distribution, 10-dimensional feature sets for training and testing data were constructed. The dimensions, here, describe different features resulting from the wavelet transform, that is to say, the total size of training data set is 10×300, where 10 is the dimension of feature size of each case and 300 comes out from 50 cases per class multiplied by 6 classes. For testing, 30 samples per class is taken. Classification results are described in terms of a 6×6 confusion matrix. The diagonal elements represent the correctly classified PQ disturbances. The off-diagonal elements represent the misclassified PQ disturbances. The overall classification accuracy for LVQ, FFML and PNN is 86%, 93.5%, 98.0%, respectively. Hence, PNN can effectively classify different kinds of PQ event as compared to other wavelet based neural networks. The 6x6 confusion matrix of FFML, PNN and LVQ are shown in Table II, Table III and Table IV, respectively. The results of SVM classification is shown in Table V. The simulation results of SVM classifier have verified that the proposed method has good efficiency in classifying six types of PQ disturbances with different characteristics. This is comparable with the percentage accuracy of the techniques based on ANN. The

simulations show that SVM has higher performance than ANN with feed forward multilayer back propagation (FFML), learning vector quantization (LVQ), probabilistic neural network (PNN). The performance of ANN varied obviously when training samples increased. Moreover, the training of ANN took much more time compared with the SVM classifier.

Table 2. Classification result of FFML

	C1	C2	C3	C4	C5	C6
C1	**99**	1	0	0	0	0
C2	6	**85**	9	0	0	0
C4	0	0	0	**95**	4	0
C5	0	0	0	7	**98**	0
C6	0	0	8	0	0	**91**

Overall Accuracy: 93.5%

Table 3. Classification result of PNN

	C1	C2	C3	C4	C5	C6
C1	**98**	0	2	0	0	0
C2	2	**96**	1	0	0	0
C3	0	1	**99**	0	0	0
C4	0	0	0	**100**	0	0
C5	0	0	1	2	**96**	0
C6	0	0	1	0	0	**99**

Overall Accuracy: 98.0%

Table 4. Classification result of LVQ

	C1	C2	C3	C4	C5	C6
C1	**93**	0	7	0	0	0
C2	9	**90**	1	0	0	0
C3	4	5	**90**	0	1	0
C4	0	9	0	**80**	10	0
C5	0	0	1	8	**91**	0
C6	0	3	20	4	12	**72**

Overall Accuracy: 86%

Table 5. Classification result of SVM

Number of test samples	Types of disturbances	Correctly Classified
30	Sag	30
30	Swell	26
30	Harmonics	30
30	High frequency transient	30
30	Low frequency transient	28
30	Fluctuation	29

7. Conclusion

This paper proposes a PQ event classification scheme utilizing a wavelet feature extractor based on detailed energy distribution and a classifier based on neural networks and SVM. The proposed method can reduce the quantity of extracted features of distorted signal without losing its characteristics and thus, requires less memory space and computation time. The performance of three types of neural network is tested for effective classification of PQ events. It is observed that PNN correctly classifies the PQ event with high accuracy PNN and gives the best performance as compared to LVQ and FFML. Therefore, the proposed method can be used as the PQ event classifier. This classification efficiency of SVM is comparable with the percentage accuracy of the techniques based on ANN. The simulations show that SVM has higher performance than ANN with feed forward multilayer back propagation (FFML), learning vector quantization (LVQ), probabilistic neural network (PNN).

References

Angrisani L., Daponte P., D'Apuuo M. and Testa A., 1996, A new wavelet transform based procedure for electrical power quality analysis," *Proceedings of the international Conference on Harmonics and Quality of Power (ICHQP),* Las Vegas, Nevada, USA, pp. 608-614.

Angrisani L., Daponte P.,'Apuzzo M. D, Testa A., 1998, A measurement method based on the wavelet transform for power quality analysis, IEEE Trans. Power Deliv., 990–998.

Arrillaga J., Bollen M.H.J. and Waston N.R., 2000. Power quality following deregulation, *Proc. IEEE.*, vol. 88, no. 2, pp. 246-261.

Arrillaga J., Watson N.R. and Chen.S. 2000. *Power system quality assessment*, John Wiley and Sons Ltd., London.

Bishop C.M., 1995, *Neural networks for pattern recognition,* Oxford University Press, Oxford.

Bishop C. M., 2008, *Pattern Recognition and Machine Learning*, Springer, New York, USA,

Borras D., Castilla M., Moreno N., Montano J.C., 2001, Wavelet and neural structure: a new tool for diagnostic of power system disturbances, *IEEE Trans. Industry Appl.* , 184–190.

Charri O., Meunier M., and Brouaye F., 1996, Wavelets: A new tool for the resonant grounded power distribution system relaying, *IEEE Trans on Power Del.*, vol. 11, pp. 1301-1308.

Cristianini N. and Shawe J. -Taylor, 2000,*An Introduction to Support Vector Machines*, Cambridge, MA: Cambridge Univ. Press .

Dugan Roger.C., McGranaghan M.F., Santoso S. and Beaty H.W., 2003. *Electrical power system quality*, 2nd Edn. McGraw Hill Book Company, New York.

Gaouda A.M., Salama M.M.A., Sultan M.R., Chikhani A.Y., 1999, Power quality detection and classification using wavelet-multiresolution signal decomposition, *IEEE Trans. Power Deliv*. 1469–1476.

Jain S.K., Saxena D., and Singh S.N., 2011. *Adaptive Wavelet Neural Network Based Harmonic Estimation of Single-Phase,* The 2nd AIT-PEA International Conference & Utility Exhibition on Power and Energy Systems: Issues and Prospects for Asia (ICUE 2011), September 28-30, Pattaya, Bangkok.

Karimi M., Mokhtari H., and Iravani M.R., 2000, Wavelet based on-line disturbance detection for power quality application, *IEEE Trans.on Power Del*., vol.15, no. 10, pp. 1212-1220.

Mallat S., 1989, A theory for multiresolution signal decomposition: The wavelet representation, *IEEE Trans. Pattern Anal. Mach. Intell.* Vol.11, no.7, pp 674-693.

Michalik M., Rebizant W., Lukowicz M., Sang-Jae L., Sang- Hee K., 2006, High impedance fault detection in distribution networks with use of wavelet- based algorithm, *IEEE Trans. on Power Del*., vol. 21, pp.1793-1802.

Perunicic B., Mallini M., Wang Z., Liu Y., 1998, Power quality disturbance detection and classification using wavelets and artificial neural networks, Proc. 8th Int. Conf. Harmonics and Quality of Power, Vol. 1, Athens, Greece, pp. 77–82.

Pillay P., Ribeiro P. and Pan Q., 1996, Power quality modeling using wavelets, *IEEE Proceedings of the 7* International Conference on Harmonics and Quality of Power (ICHQP),* Las Vegas, Nevada, USA, , pp. 625-631

Rioul O. and Vetterli M., 1991, Wavelets and Signal Processing, *IEEE Signal Processing Magazine*, pp. 14-38.

Robertson D. C., Camps O.1., Mayer J. S. and Gish W.B., 1996, Wavelet and electromagnetic power system transients, *IEEE Transaction on Power Delivery*, Vol. 11, No 2, pp. 1050-1058.

Rodney H.G. Tan, Ramachandaramurthy V.K., 2010, Numerical Model Framework of Power Quality Events, *European Journal of Scientific Research,* Vol.43, pp.30-47.

Santoso S., Powers E. J., and Grady W. M., 1997 Power quality disturbance data compression using wavelet transform methods, IEEE Trans. on Power Del.,vol.12, pp. 1250-1257.

Santoso S., Powers E. J., Grady W. M., and Hofmann P., 1996, Power quality assessment via wavelet transform analysis," *IEEE Transaction on Power Delivery,* Vol. 11, No 2, pp. 924-930.

Santoso S., Powers E.J., Grady W.M., Parsons A.C., 2000, Power quality disturbance waveform recognition using wavelet-based neural classifier-Part 1: theoretical foundation, *IEEE Trans. Power Deliv*. Vol. 15, No. 1, pp. 222–228.

Tunaboylu N. S. and Collins E. R., 1996, The wavelet transform approach to detect and quantify voltage sags, Proceedings of the 7" International Conference on Harmonics and Quality of Power (ICHQP), Las Vegas, Nevada, USA, pp. 619-623.

Vapnik V. N*., ,* 1998.*Statistical Learning Theory,* New York: Wiley.

Biographical notes

D. Saxena obtained B. Tech. (Hons) in Electrical Engineering from KNIT Sultanpur (UP), India and M. Tech. in Process Control from Netaji Subhas Institute of Technology, New Delhi, India in 1999 and 2003, respectively. Presently, she is working as Associate Professor and Head, Department of Electrical and Electronics Engineering, Invertis Institute of Engineering and Technology, Bareilly. She is registered as PhD candidate in UP Technical University Lucknow, India. Her research interests are power quality, power electronics, control systems and DSP application in power.

K.S. Verma received his B. Tech. (Hons) in Electrical Engineering and M. Tech. in Electrical Engineering (Power Systems) both from KNIT Sultanpur (India) in 1987 and 1998, respectively. He obtained his Ph.D. degree in Electrical Engineering (Power Systems) from Indian Institute of Technology, Roorkee, (India) in 2003. Presently, he is Director KNIT Sultanpur (India). His research interests include FACTS Technology, distributed generation, power system optimization & control, power quality and AI application in power system.