Machine Intelligence Research Labs
Global Network for Innovation and Research Excellence

A Machine Intelligence Approach for Classification of Power Quality Disturbances

B K Panigrahi[1], V. Ravi Kumar Pandi[1], Ajith Abraham[2] and Swagatam Das[2]

[1]*Department of Electrical Engineering, IIT, Delhi,*
bkpanigrahi@ee.iitd.ac.in
[2]*Machine Intelligence Research Labs, MIR Labs*
ajith.abraham@ieee.org
[3]*Department of Electrical and Communication Engineering, Jadavpur University, Kolkata, India*
swagatamdas19@yahoo.co.in

Abstract

This paper presents the combination of advanced signal processing techniques and the machine intelligence approach to classify the power quality events. The Wavelet Transform (WT) and the S – Transform (ST) are utilized to extract the important useful features of the disturbance signal. The features extracted by using the above approaches are used to train a PNN classifier for automatic classification of the PQ disturbances. Eleven types of power quality disturbances are considered for the classification purpose. The simulation results show that the combination of S-Transform and PNN is an effective method to detect and classify different power quality disturbances.

1. Introduction

In recent years, power quality has become a significant issue for both utilities and customers. Power quality issues [1] and the resulting problems are the consequences of the increasing use of solid state switching devices, non-linear and power electronically switched loads, unbalanced power systems, lighting controls, computer and data processing equipments as well as industrial plant rectifiers and inverters. These electronic type loads cause quasistatic harmonic dynamic voltage distortions, inrush, pulse type current phenomenon with excessive harmonics and high distortion. A power quality (PQ) problem usually involves a variation in the electric service voltage or current, such as voltage dips and fluctuations, momentary interruptions, harmonics and oscillatory transients causing failure or mal-operation of the power service equipment. Hence to improve power quality, fast and reliable detection of the disturbances and the sources and causes of such disturbances must be known before any appropriate mitigating action can be taken.

However, in order to determine the causes and sources of disturbances, one must have the ability to detect and localize these disturbances. In the current research trends in power quality studies, Wavelet transform (WT) [2-4] is widely used in analyzing non-stationary signals for power quality assessment [5,6] . In order to identify the type of disturbance present in the power signal more effectively, several authors have presented different methodologies based on combination of wavelet transform (WT) and artificial neural network (ANN) [7] .Using the multiresolution properties of WT [6], the features of the disturbance signal are extracted at different resolution levels and are used to train different ANN algorithms. By this method, it is possible to extract important information from a disturbance signal and determine the type of disturbance that has caused a power quality problem to occur. Gaing [8] demonstrated the classification of 7 types of PQ events by using wavelets and probabilistic neural network (PNN). Energy distribution at 13 decomposition levels of wavelet and time duration of each disturbances are taken as features and these 14 features are applied to PNN for classification with increased memory and computational overhead due to the large number of features. S- Transform [9] has been successively applied to the analysis [10] of different power quality disturbances. A comparative study for the classification of power quality signals has been recently reported using WT [11] and ST [12].

In this paper, we attempted to decompose the power system disturbance signal up to 13 level of decomposition as mentioned in [8] using Wavelet Transform. In this paper the energy of the dialation coefficients at different frequency subbands are considered as the features to identify the disturbance.

The same approach is also extended for the feature selection using ST. The features obtained for 11 classes of power quality disturbances are are separated as training and testing data sets. The classification is performed using a PNN.

2. Introduction to DSP Techniques

(A)Wavelet transform

The Discrete Wavelet Transform (DWT) is a special case of the WT that provides a compact representation of a signal in time and frequency that can be computed efficiently. The DWT is calculated based on two fundamental equations: the scaling function $\phi(t)$, and the wavelet function $\psi(t)$, where

$$\phi(t) = \sqrt{2}\sum_{k} h_k\,\phi(2t-k) \tag{1}$$

$$\psi(t) = \sqrt{2}\sum_{k} g_k\,\phi(2t-k) \tag{2}$$

These functions are two-scale difference equations based on a chosen scaling function (mother wavelet), with properties that satisfy the following conditions

$$\sum_{k=1}^{N} h_k = \sqrt{2}$$

$$\sum_{k=1}^{N} h_k h_{k+2l} = 1 \qquad if\ l = 0$$
$$= 0 \qquad if\ l \in Z,\quad l \neq 0 \tag{3}$$

The discrete sequences h_k and g_k represent discrete filters that solve each equation, where $g_k = (-1)^k h_{N-1-k}$. The scaling and wavelet functions are the prototype of a class of orthonormal basis functions of the form

$$\phi_{j,k}(t) = 2^{\frac{j}{2}}\phi(2^j t - k);\quad j,k \in Z \tag{4}$$

$$\psi_{j,k}(t) = 2^{\frac{j}{2}}\psi(2^j t - k);\quad j,k \in Z \tag{5}$$

where the parameter j controls the dilation or compression of the function in time scale and amplitude. The parameter *k* controls the translation of the function in time. *Z* is the set of integers.

Once a wavelet system is created, it can be used to expand a function *f(t)* in terms of the basis functions

$$f(t) = \sum_{l\in Z} c(l)\phi_l(t) + \sum_{j=0}^{J-1}\sum_{k=0}^{\infty} d(j,k)\ \psi_{j,k}(t) \tag{6}$$

where, the coefficients $c(l)$ and $d(j,k)$ are calculated by inner product as

$$c(l) = \langle \phi_l \mid f \rangle = \int f(t)\phi_l(t)dt \tag{7}$$

$$d(j,k) = \langle \psi_{j,k} \mid f \rangle = \int f(t)\psi_{j,k}dt \tag{8}$$

The expansion coefficients $c(l)$ represent the approximation of the original signal *f(t)* with a resolution of one point per every 2^J points of the original signal. The expansion coefficients $d(j,k)$ represent details of the original signal at different levels of resolution. $c(l)$ and $d(j,k)$ terms can be calculated by direct convolution of *f(t)* samples with the coefficients h_k and g_k , which are unique to the specific mother wavelet chosen.

The WT can be implemented with a specially designed pair of FIR filters called a quadrature mirror filters (QMFs) pair. QMFs are distinctive because the frequency responses of the two FIR filters separate the high- and low-frequency components of the input signal. The dividing point is usually halfway between 0 Hz and half the data sampling rate (the Nyquist frequency). The outputs of the QMF filter pair are decimated (or de-sampled) by a factor of two. The low-frequency (low-pass) filter output is fed into another identical QMF filter pair. This operation can be repeated recursively as a tree or pyramid algorithm, yielding a group of signals that divides the spectrum of the original signal into octave bands with successively coarser measurements in time as the width of each spectral band narrows and decreases in frequency. The tree or pyramid algorithm can be applied to the WT by using the wavelet coefficients as the filter coefficients of the QMF filter pairs as shown in the Fig. 1. In WT multi-resolution algorithm, same wavelet coefficients are used in both low-pass (LP) and high-pass (HP) filters. The LP filter coefficients are associated with the scaling function , and the HP filter is associated with the wavelet function . Fig. 1 shows the tree algorithm of a multi resolution WT for a discrete signal sampled at 3200 Hz. The outputs of the LP filters are called the approximations (A), and the outputs of the HP filters are called the details (D). In wavelets applications, different basis functions have been proposed and selected. Each basis function has its

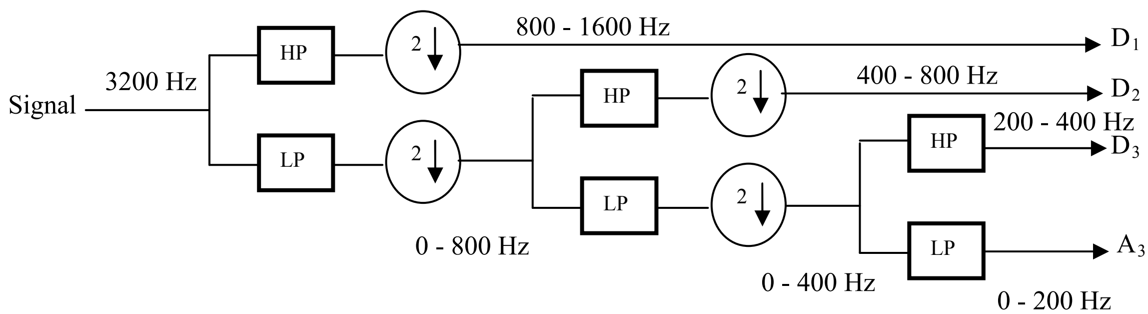

Figure 1. Wavelet Decomposition

feasibility depending on the application requirements. Daubechies wavelet family is one of the most suitable wavelet families in analyzing power system transients In the present work, the db4 wavelet has been used as the wavelet basis function.

(B) S - Transform

Wavelet transform addresses the problem of resolution by introducing a dilation (or scale) parameter d . The continuous wavelet transform (CWT) $W(\tau,d)$ of a function $y(t)$ is given as

$$W\left(\tau,d\right)=\int_{-\infty}^{\infty} y(t)w(t-\tau,d)dt \tag{9}$$

where $W(\tau,d)$ is the scale replica of a fundamental wavelet. The spectral information in the signal $y(t)$ is extracted through the correlation or convolution with $W(\tau,d)$. The dilation parameter d determines the width of the wavelet $W(\tau,d)$ and thus controls the resolution. The wavelet transform is displayed in state space defined by the dilation d and translation τ . The spectral information of the signal $y(t)$ can be obtained from the state space representation. Stockwell *et al.* [9-10] proposed a new windowed Fourier transform called the S transform, as an extension to the ideas of the Gabor transform and the wavelet transform. The S - transform of a signal $x(t)$ is defined as

$$S(\tau,f)=\int_{-\infty}^{\infty} x(t)g(\tau-t)\exp(-j2\pi f\tau)\,d\tau \tag{10}$$

where

$$g(t)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{t^2}{2\sigma^2}\right) \tag{11}$$

and

$$\sigma(f)=\frac{1}{|f|} \tag{12}$$

Combining equation (10) to (12) gives

$$S(\tau,f)=\int_{-\infty}^{\infty} x(\tau)\frac{|f|}{\sqrt{2\pi}}\exp\left(-\frac{(\tau-t)^2 f^2}{2}\right)\exp(-j2\pi ft)\,dt \tag{13}$$

In equation (13), S denotes the S-transform of x, which is a continuous function of time t and the frequency is denoted by f ; and the quantity τ is a parameter which controls the position of the Gaussian window on the t-axis. The scaling property of the Gaussian window is similar to that of the scaling property of continuous wavelets, because one wavelength of the Fourier frequency is always equal to one standard deviation of the window. The S-transform, however, is not a wavelet transform, because the oscillatory part of the S-transform is provided by the complex Fourier sinusoid, which does not translate with the Gaussian window when τ is changed. As a result, the shapes of the real and imaginary parts of the S-transform change as the Gaussian window translates in time. True wavelets do not have this property because their entire waveform translates in time with no change in shape. Thus, the S-transform is conceptually a hybrid of short-time Fourier analysis and wavelet analysis, containing elements of both but falling entirely into neither category.

We can define $A(\tau,f)=\left|S(\tau,f)\right|$ is the amplitude of the S-spectrum and $\theta(\tau,f)=\arctan\left(\frac{\mathrm{Im}[S(\tau,f)]}{\mathrm{Re}[S(\tau,f)]}\right)$ is the phase of the S-spectrum. The phase of S-spectrum is an improvement on the wavelet transform in that the average of all the local spectra does indeed give the same result as the Fourier transform.

3. Feature Extraction

(a) Feature Extraction for WT

The detailed co-efficient D_{ij} at each decomposition level is used to extract the features. Statistical features like energy, mean, standard deviation, Shannon entropy, skew ness and kurtosis of the decomposition coefficients D_{ij} are calculated by using the following equations

$$\text{Energy } ED_i = \sum_{j=1}^{N} \left|D_{ij}\right|^2 \quad i = 1,2 \ldots\ldots l \qquad (14)$$

where i=1,2,…….*l* is the wavelet decomposition level from level 1 to level *l*. N is the number of coefficients of detail at each decomposition level. The term $E(D_{ij}-\mu_i)^k$ in skewness and kurtosis calculation is the expected value of the quantity, also called as k^{th} moment about the mean. Thus for a ‘ *l*’ level decomposition the feature vector adopted is of length ‘6**l*’ i.e., 78 features are there for decomposition up to 13^{th} level and is denoted by

(a) Feature Extraction for ST

Feature extraction is done by applying equation 14 to the contours of the S-matrix as well as directly on the S-matrix. These features have been found to be useful for detection, classification of the PQ disturbance signals. The power signal is normalized with respect to a base value, which is the normal value without any disturbance. The features are specified as follows:

Feature 1: Energy of the Magnitude contour (Magnitude contour corresponding to the maximum magnitude of the S- matrix at each sample, hence it reflects the fundamental of the signal).

Feature 2: Energy of the frequency contour. (This reflects the frequency content of the signal particularly for harmonic, and transients)

Feature 3: Standard deviation of the Phase contour.

Eleven types of power quality disturbances are simulated and the features of all the types of disturbances are extracted from the S – matrix. The signals are generated using MATLAB [15]. The sampling frequency is 64x50 i.e. 3.2kHZ. Thus each cycle of the signal contains 64 samples. The total signal length in each case is 10 cycles having 640 sample points.

4. Classification of PQ Disturbances Using Probabilistic Neural Network

The PNN model is one among the supervised learning networks, and has the following features distinct from those of other networks in the learning processes [16].

- It is implemented using the probabilistic model, such as Bayesian classifiers.
- A PNN is guaranteed to converge to a Bayesian classifier provided that it is given enough training data.
- No learning processes are required.
- No need to set the initial weights of the network.
- No relationship between learning processes and recalling processes.
- The difference between the inference vector and the target vector are not used to modify the weights of the network.

The learning speed of the PNN model is very fast, making it suitable for fault diagnosis and signal classification problems in real time. Fig. 2 shows the architecture of a PNN model that is composed of the radial basis layer and the competitive layer.

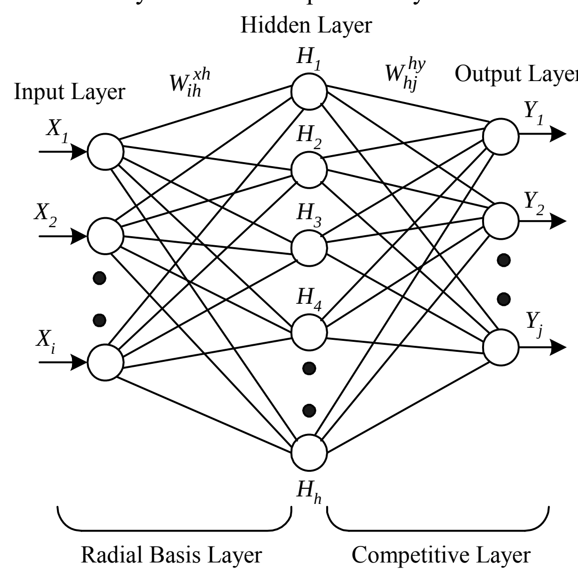

Figure 2. Architecture of PNN

In the signal classification application, the training examples are classified according to their distribution values of probabilistic density function (PDF), which is the basic principle of the PNN. A simple PDF is as follows:

$$f_k(X) = \frac{1}{N_k} \sum_{j=1}^{N_k} \exp\left(- \frac{\left\| X - X_{kj} \right\|}{2\sigma^2} \right) \qquad (15)$$

The output vector H of the hidden layer in the PNN is as below:

$$H_h = \exp\left(\frac{-\sum_i (X_i - W_{ih}^{xh})^2}{2\sigma^2}\right) \quad (16)$$

$$net_j = \frac{1}{N_j}\sum_h W_{hj}^{hy} H_h \quad \text{and} \quad N_j = \sum_h W_{hj}^{hy},$$

$$net_j = \max_k (net_k) \text{ then } y_j = 1, \text{ else } y_j = 0 \quad (17)$$

where

i	number of input layers;
h	number of hidden layers;
j	number of output layers;
k	number of training examples;
N	number of classifications (clusters);
σ	smoothing parameter (standard deviation);
X	input vector;

$\|X - X_{kj}\|$ is Euclidean distance between the vectors X and, X_{kj}, i.e. $\|X - X_{kj}\| = \sum_i (X - X_{kj})^2$;

W_{ih}^{xh} connection weight between the input layer X and the hidden layer H;

W_{hj}^{hy} connection weight between the hidden layer H and the output layer Y.

The learning and recalling processes of the PNN for classification problems can be found in [11].

5. Test Results

Eleven classes (C1-C11) of different PQ disturbances are taken for classification and they are named as

C1→	Normal
C2→	Pure Sag
C3→	Pure Swell
C4→	Momentary Interruption
C5→	Harmonics
C6→	Sag with Harmonic
C7→	Swell with Harmonic
C8→	Flicker
C9→	Notch
C10→	Spike
C11→	Transient

Based on the feature extraction by the S-Transform method, 3 dimensional feature sets for training and testing are constructed. The dimensions here describe the different features derived from S-Transform. All the data sets of features for various classes are applied to PNN for automatic classification of PQ events. To illustrate the nature of the feature sets for all the 11 classes, Figs. 3 and 4are presented here. Fig. 3 illustrates feature 1 vs. feature 2, Fig. 4 represents the feature 2 vs. 3 .It is well visualized that some of the classes have distinct features, where as some have overlapping.

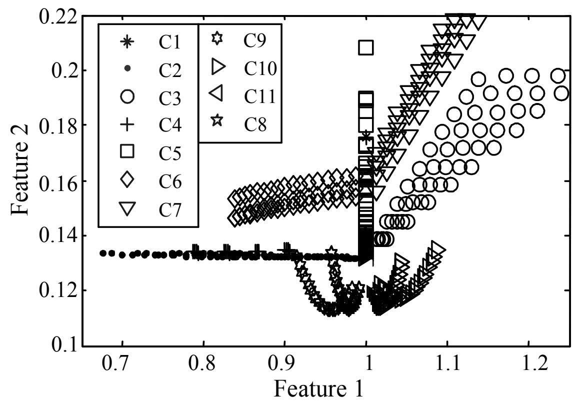

Figure 3. Feature 1 vs. feature 2

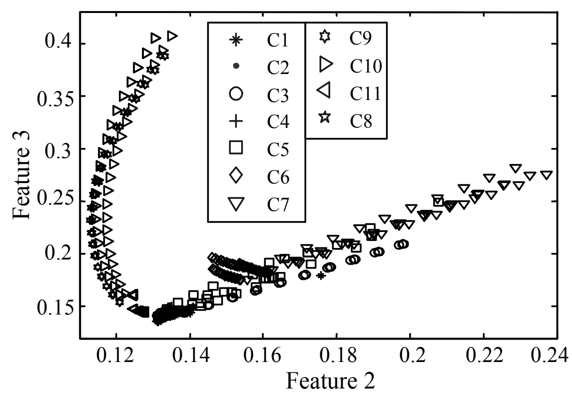

Figure 4. Feature 2 vs. feature 3

To evaluate the performance of PNN, its results are compared with FFML. Total 1100 data's (including all classes) are taken for training and 50 data's of each class are considered for testing. Each data consists of 3 features as stated earlier. Table I shows the classification results with PNN. The diagonal elements represent correctly classified PQ classes. The off-diagonal elements represent the misclassification. The overall accuracy is calculated by finding the average of all diagonal elements and it is found to be 98.5% for ST (Table I) and 96.9% for WT (Table II).

In an electrical power distribution network, the practical data consists of noise; therefore, the proposed approach has to be analyzed under noisy environment. Gaussian white noise is widely considered in the research of power quality issues. We have obtained the noisy signals for all the 11 classes having different signal to noise ratio (SNR). The features for the noisy signal are extracted by using both the WT and ST as described in section 3. It is observed that the proposed approach using ST based features to classify the different power quality disturbances works satisfactorily in comparison to WT based techniques. How ever once the WT based denoising [14] technique is applied for the preprocessing of the signal prior to the feature extraction, the classification accuracy of WT based techniques improved significantly. The classification accuracy for the noisy cases are reported in Table 3.

Table I. Classification results of PNN for ST

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
C1	**99**	0	0	0	1	0	0	0	0	0	0
C2	0	**99**	0	1	0	0	0	0	0	0	0
C3	1	0	**99**	0	0	0	0	0	0	0	0
C4	0	2	0	**98**	0	0	0	0	0	0	0
C5	0	0	0	0	**98**	0	0	2	0	0	0
C6	1	0	0	0	1	**98**	0	0	0	0	0
C7	0	1	0	0	0	0	**99**	0	0	0	0
C8	1	0	0	0	1	0	0	**97**	0	0	1
C9	1	0	0	0	0	0	0	0	**99**	0	0
C10	1	0	0	0	0	0	0	0	1	**98**	0
C11	0	0	0	0	0	0	0	0	0	0	**100**
Overall Accuracy: 98.5%											

Table 2. Classification results of PNN for WT

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
C1	**97**	0	1	0	1	0	0	0	0	0	1
C2	0	**95**	0	1	0	2	0	2	0	0	0
C3	2	0	**97**	0	0	1	0	0	0	0	0
C4	0	2	0	**98**	0	0	0	0	0	0	0
C5	0	0	0	0	**98**	0	0	2	0	0	0
C6	2	0	0	0	2	**96**	0	0	0	0	0
C7	0	2	0	0	0	0	**98**	0	0	0	0
C8	2	0	0	1	2	0	0	**95**	0	0	1
C9	1	0	0	1	0	0	0	0	**97**	0	1
C10	2	0	0	0	0	0	0	0	1	**97**	0
C11	1	0	0	0	0	0	0	0	0	1	**98**
Overall Accuracy: 96.9%											

Table 3. Classification accuracy in percentage for noisy signals

Methods used	SNR 40 dB	SNR 30 dB	SNR 20 dB
ST	98.13	96.45	95.13
WT	91.79	90.16	85.72
WT with denoising	95.6	94.78	92.78

6. Conclusions

This paper describes the most recent and advanced signal processing techniques known as Wavelet Transform and S – Transform to extract useful and important features from the various power quality disturbances. Once the feature set is formed it is the prime duty to classify the disturbance accurately so that the disturbance can be identified and necessary mitigation action can be initiated. We have used PNN for the classification purpose and it was observed that the combination of ST and PNN achieves a higher accuracy as compared to the combination of WT and PNN in case of pure as well as noisy signals.

References

[1] Understanding Power Quality , Math H.J. Bollen, IEEE Press

[2] Daubechies, I. (1990) The wavelet transform, time/frequency location and signal analysis. *IEEE Transactions on Information Theory*, **36**, 961-1005.

[3] Mallat, S. G. (1989) A theory of multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **11**(7), 674-693.

[4] Meyer, Y. (1992) *Wavelets and Operators*. Cambridge University Press, London, U.K.

[5] S. Santoso., E. J. Powers, W. M. Grady, and P. Hofmann, "Power quality assessment via wavelet transform analysis", *IEEE Transactions on Power Delivery*, vol. 11, no. 2, pp. 924-930, April 1996.

[6] A. M. Gaouda, M.M.A. Salama, M. K. Sultan and A.Y. Chikhani, **"**Power Quality Detection and Classification Using Wavelet-Multiresolution Signal Decomposition", *IEEE Transactions on Power Delivery*, vol. 14, no. 4, pp. 1469-1476, Oct. 1999.

[7] Santoso, S., E. J. Powers, W. Grady, and A. Parsons (1997a) Power quality disturbance waveform recognition using wavelet-based neural classifier, Part 1: theoretical foundation. The 1997 IEEE/PES Winter Meeting, New York, NY, U.S.A.

[8] Z. L. Gaing, "Wavelet-Based Neural Network for Power Disturbance Recognition and Classification", *IEEE Trans. on Power Delivery*, vol. 19, no.4, pp. 1560-1568, Oct. 2004.

[9] R. G. stockwell, L. Mansinha and R. P. Lowe, "Localization of the Complex Spectrum: The S-Transform", *IEEE Trans. on Signal Processing*, vol. 44, no.4, pp. 998-1001, April. 1996.

[10] C. R. Pinnegar and L. Mansinha, "The S-Transform with windows of arbitrary and varying shape", Geophysics, vol. 68, no. 1, pp. 381-385, Jan./Feb. 2003.

[11] P. K. Dash, B. K. Panigrahi and G. Panda, "Power Quality analysis using S-Transform", *IEEE Trans. on Power Delivery*, vol. 18, no. 2, pp. 406-411, April 2003.

[12] Murat Uyar, Selcuk Yildirim and Muhsin Tunay Gencoglu, " An effective Wavelet based feature extraction method for classification of power quality disturbance signals", Electric power System Research, Vol. 78, 2008, pp. 1474 – 1755.

[13] Murat Uyar, Selcuk Yildirim and Muhsin Tunay Gencoglu, "An expert system based on S-transform and neural network for automatic classification of power quality disturbances", Expert Systems with Application, doi. 10.1016/j.eswa.2008.07.030

[14] Hong-Tzer Yang and Chiung-Chou Liao, "A de-noising scheme for enhancing wavelet-based power quality monitoring system**",** *IEEE Transactions on Power Delivery,* Vol. 16, No. 3, pp. 353 - 360, July 2001.

[15] MATLAB, Math Works, Inc., Natick, MA, USA, 2000.

[16] D. F. Specht, "Probabilistic neural networks", *Neural Networks*, vol. 3, no. 1, pp. 109-118, 1990.