We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

4,800

Open access books available

122,000

International authors and editors

135M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities

CLARIVATE ANALYTICS

BOOK CITATION INDEX

INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Wavelet and PCA to Power Quality Disturbance Classification Applying a RBF Network

Giovani G. Pozzebon[1], Ricardo Q. Machado[1], Natanael R. Gomes[2],
Luciane N. Canha[2] and Alexandre Barin[2]
[1]São Carlos School of Engineering, Department of Electrical Engineering,
University of São Paulo
[2]Federal University of Santa Maria
Brazil

1. Introduction

The quality of electric power became an important issue for the electric utility companies and their customers. It is often synonymous with voltage quality since electrical equipments are designed to operate within a certain range of supply specifications. For instance, current microelectronic devices are very sensitive to subtle changes in power quality, which can be represented as a disturbance-induced variation of voltage amplitude, frequency and phase (Dugan et al., 2003).

Poor power quality (PQ) is usually caused by power line disturbances such as transients, notches, voltage sags and swells, flicker, interruptions, and harmonic distortions (IEEE Std. 1159, 2009). In order to improve electric power quality, the sources and causes of such disturbances must be known. Therefore, the monitoring equipment needs to firstly and accurately detect and identify the disturbance types (Santoso et al., 1996). Thus, the use of new and powerful tools of signal analysis have enabled the development of additional methods to accurately characterize and identify several kinds of power quality disturbances (Karimi et al., 2000; Mokhtary et al., 2002).

Santoso et al. proposed a recognition scheme that is carried out in the wavelet domain using a set of multiple neural networks. The network outcomes are then integrated by using decision-making schemes such as a simple voting scheme or the Dempster-Shafer theory. The proposed classifier is capable of providing a degree of belief for the identified disturbance waveform (Santoso et al., 2000a, 2000b). A novel classification method using a rule-based method and wavelet packet-based hidden Markov models (HMM) was proposed bay Chung et al. The rule-based method is used to classify the time-characterized-feature disturbance and the wavelet packet-based on HMM is used for frequency-characterized-feature power disturbances (Chung et al., 2002). Gaing presented a prototype of wavelet-based network classifier for recognizing power quality disturbances. The multiresolution-analysis technique of discrete wavelet transforms (DWT) and Parseval's theorem are used to extract the energy distribution features of distorted signals at different resolution levels. Then, the probabilistic neural network classifies these extracted features of disturbance type identification according to the transient duration and energy features (Gaing, 2004). Zhu et

al. proposed an extended wavelet-based fuzzy reasoning approach for power quality disturbance recognition and classification. The energy distribution of the wavelet part in each decomposition level is calculated. Based on this idea, basic rules are generated for the extended fuzzy reasoning. Then, the disturbance waveforms are classified (Zhu et al., 2004). Further on, Chen and Zhu presented a review of the wavelet transform approach used in power quality processing. Moreover, a new approach to combine the wavelet transform and a rank correlation is introduced as an alternative method to identify capacitor-switching transients (Chen & Zhu, 2007).

Taking into account these ideas, this chapter proposes the application of a different method of power quality disturbance classification by combining discrete wavelet transform (DWT), principal component analysis (PCA) and an artificial neural network in order to classify power quality disturbances. The method proposes to analyze seven classes of signals, namely Sinusoidal Waveform, Capacitor Switching Transient, Flicker, Harmonics, Interruption, Notching and Sag, which is composed by four main stages: (1) signal analysis using the DWT; (2) feature extraction; (3) data reduction using PCA; (4) classification using a radial basis function network (RBF). The MRA technique of DWT is employed to extract the discriminating features of distorted signals at different resolution levels. Subsequently, the PCA is used to condense information of a correlated set of variables into a few variables, and a RBF network is employed to classify the disturbance types.
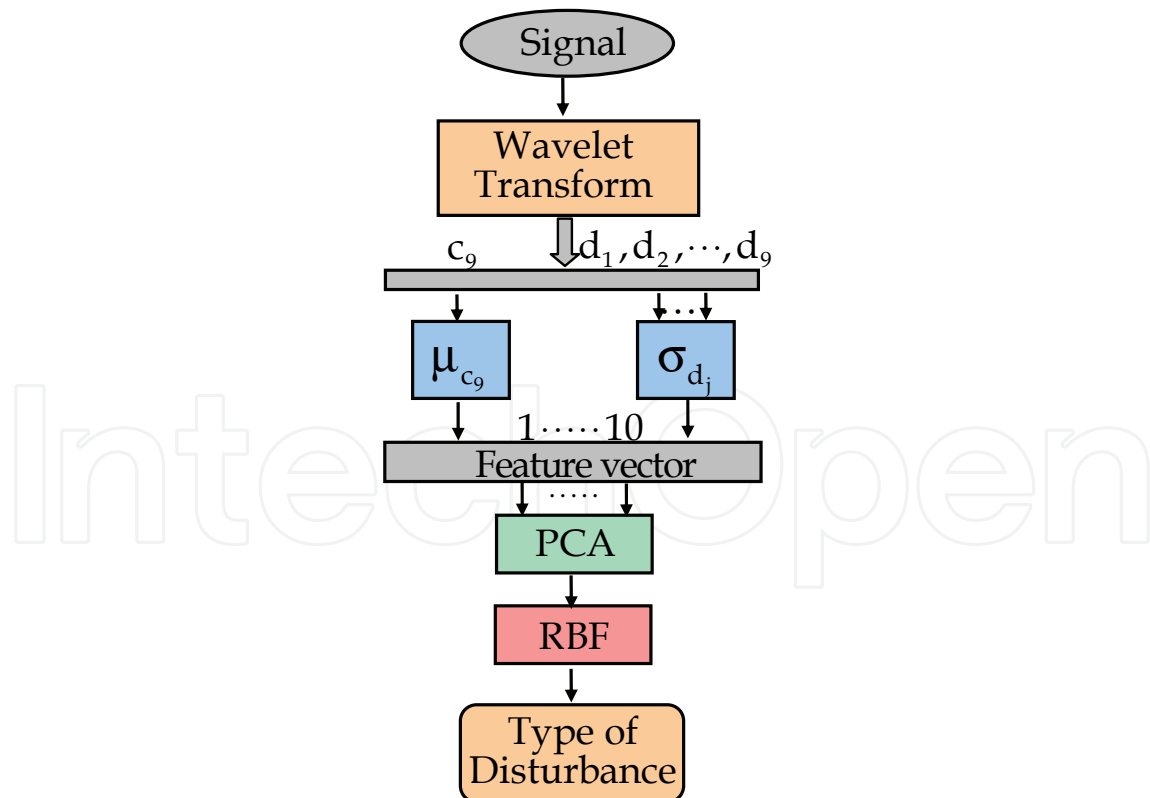
2. Proposed classification scheme

Fig. 1. Diagram of the proposed classification scheme

This section presents the mainframe of the scheme proposed in this paper using the wavelet transform, principal components and neural networks to classify PQ disturbances. The

proposed scheme diagram is shown in Fig. 1. Initially, the input signals are analyzed using the discrete wavelet transform tool, which employs two sets of functions called scaling functions and wavelet functions associated with low pass and high pass filters, respectively. Then, the signal is decomposed into different resolution levels aiming to discriminate the signal disturbances. The output of the DWT stage is used as the input to the feature extraction stage, on which the featuring signal vectors are built. In order to reduce the amount of data, the PCA technique is applied to the feature vector in order to concentrate the information from the disturbance signal and to reduce the amount of the training data used, consequently minimizing the number of input RBF neurons while maintaining the recognition accuracy. Finally, an RBF network is employed to perform the disturbance type classification.

As aforementioned, in the introduction, this work proposes to analyze seven classes of different types of PQ disturbances as follows: Pure sine (C_1); Capacitor switching (C_2); Flicker (C_3); Harmonics (C_4); Interruption (C_5); Notching (C_6); Sag (C_7). The databases used for training and evaluation of the proposed system and classification algorithms were performed in Matlab®. The tools used in this approach are presented in sequence.

2.1 The wavelet transform and multiresolution analysis

The DWT is a versatile signal processing tool that has many engineering and scientific applications (Barmada et al., 2003). One area in which the DWT has been particularly successful is transient analysis in power systems (Santoso et al., 2000a, 2000b; Yilmaz et al., 2007), used to capture the transient features and to accurately localize them in both time and frequency contexts. The wavelet transform is particularly effective in representing various aspects of non-stationary signals such as trends, discontinuities and repeated patterns, in which other signal processing approaches fail or are not as effective. Through wavelet decomposition, transient features are accurately captured and localized in both time and frequency contexts.

A wavelet is an effective time–frequency analysis tool to detect transient signals. Its features of extraction and representation properties can be used to identify various transient events in power signals. The discrete wavelet transform analyzes the signal at different frequency bands with different resolutions by decomposing the signal into a coarse approximation and detail information (Chen & Zhu, 2007). This capability to expand function or signal with different resolutions is termed as Multiresolution Analysis (MRA) (Mallat, 1989). The DWT employs two sets of functions called scaling functions, $\phi_{j,n}[t]$, and wavelet functions, $\psi_{j,n}[t]$, which are associated with low-pass and high-pass filters, respectively. The decomposition of the signal into the different frequency bands is simply obtained by successive high-pass and low-pass filtering of the time domain signal. The discrete forms of scale and wavelet functions are, respectively, defined as follows.

$$\phi_{j,n}[t] = 2^{\frac{j}{2}} \sum_{n} c_{j,n} \phi[2^j t - n] \tag{1}$$

$$\psi_{j,n}[t] = 2^{\frac{j}{2}} \sum_{n} d_{j,n} \psi[2^j t - n] \tag{2}$$

Where c_j and d_j are the scaling and wavelet coefficients indexed by j, and both functions must be orthonormal.

The wavelet and scaling functions are used to perform simultaneously a multiresolution analysis decomposition and reconstruction of the signal. The former can decompose the original signal in several other signals at different resolution levels. From these decomposed signals, the original signal can be recovered without losing any information. Therefore, in power quality disturbance signals, the MRA technique discriminates the disturbances from original signals, and then they can be analyzed separately (Debnath, 2002). The recursive mathematical representation of MRA is as follows:

$$V_j = W_{j+1} \oplus V_{j+1} = W_{j+1} \oplus W_{j+2} \oplus \cdots \oplus W_{j+n} \oplus V_n \tag{3}$$

Where: V_{j+1} is the approximate version of a given signal at scale j+1; W_{j+1} is the detailed version displaying all transient phenomena of the given signal at scale j+1; symbol $\oplus$ denotes an orthogonal summation; and n represents the decomposition level.

Since $\psi_{j,n} \in W_j$ it follows immediately that $\psi_{j,n}$ is orthonormal to $\phi_{j,n}$ because all $\phi_{j,n} \in V_j$ are $V_j \perp W_j$. Also, because all V_j are mutually orthogonal, it follows that the wavelets are orthonormal across scaling. A detailed approach about this theory can be found in (Mallat, 1989; Strang & Nguyen, 1997; Debnath, 2002).

From the engineering point of view, DWT is a digital filtering process in the time domain, by discrete convolution, using the analyses Finite-Impulse-Response (FIR) filters h and g, followed by a down sampling of two. The filter g(k) can generate a detailed version of the signal, while h(k) produces an approximate version of the signal. In the DWT, the resulting coefficients from the low-pass filtering process can be processed again as entrance data for a subsequent bank of filters, generating another group of approximation and detail coefficients.

The schematic diagram in Fig. 2 shows two decomposed levels of DWT. The input signal f(t) is split into the approximation $c_{j,k}$ and the detail $d_{j,k}$ by a low-pass and a high-pass filters named h(k) and g(k), respectively. Both, output approximation and detail are decimated by 2. In a practical approach, a DWT depends on:

- the original signal, f(t);
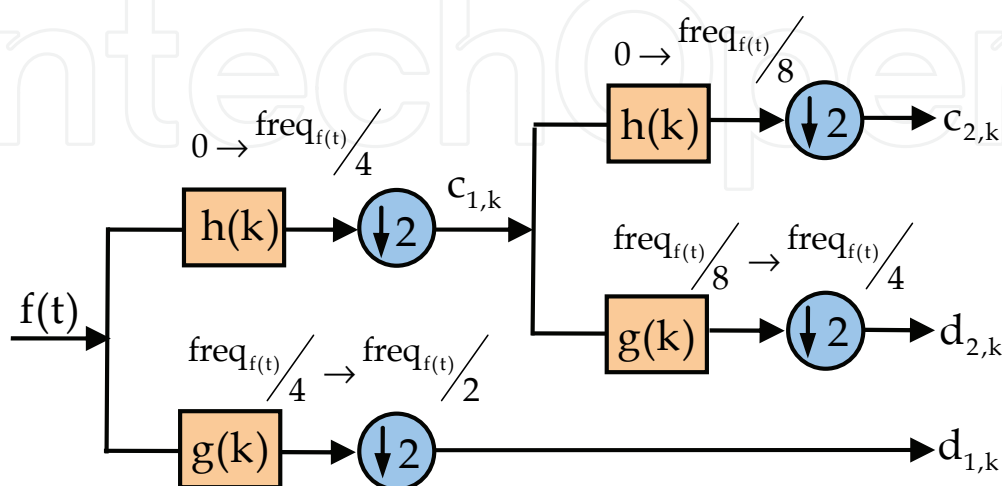- the low-pass filter, h(k), used;
- the high-pass filter, g(k), used.

Fig. 2. Decomposition of f(t) into 2 scales

There are several families of wavelet functions which contain filters of several supports (filter size). However, in this application, as in (Zhu et al., 2004; Chen & Zhu, 2007), the Daubechie "db4" wavelet was adopted to perform the DWT. As can be seen in (Daubechies, 1992), the "db4" is a wavelet of support four, i.e., each filter has four coefficients, $h = \begin{bmatrix} h_0 & h_1 & h_2 & h_3 \end{bmatrix}$. The analysis filter h is always the QMF (quadrature mirror filter) pair of g. Therefore, from the high-pass filter g the low-pass filter h is obtained, inverting the order and putting the negative sign alternately as follows, $g = \begin{bmatrix} h_3 & -h_2 & h_1 & -h_0 \end{bmatrix}$. The coefficients of the "db4" wavelet filters, h and g are presented as follows (Daubechies, 1992).

$$h = \begin{bmatrix} \dfrac{1+\sqrt{3}}{4\sqrt{2}} & \dfrac{3+\sqrt{3}}{4\sqrt{2}} & \dfrac{3-\sqrt{3}}{4\sqrt{2}} & \dfrac{1-\sqrt{3}}{4\sqrt{2}} \end{bmatrix} \tag{4}$$

$$g = \begin{bmatrix} \dfrac{1-\sqrt{3}}{4\sqrt{2}} & \dfrac{-3+\sqrt{3}}{4\sqrt{2}} & \dfrac{3+\sqrt{3}}{4\sqrt{2}} & \dfrac{-1-\sqrt{3}}{4\sqrt{2}} \end{bmatrix} \tag{5}$$

In practice, the previous filters are the only elements required to calculate the DWT of any signal. As previously mentioned, it is a digital filtering process in the time domain by discrete convolution. In (6) and (7) the relations from the level c_j to the next level, c_{j+1} and d_{j+1} are given. This relation involves the filters h and g. For specific filter, these equations allow to find the wavelet coefficients using samples of the signal f(t), once the samples are the initial coefficients c_j (Mix & Olejniczak, 2003).

$$c_{j+1}(k) = \sum_n h(n - 2k)c_j(n) \tag{6}$$

$$d_{j+1}(k) = \sum_n g(n - 2k)c_j(n) \tag{7}$$

Approximation and detail coefficients are down-sampled by 2 in each decomposition level. According to the Nyquist theorem cited in (Mallat, 1989), the maximum frequency of an original signal f(t) sampled at $freq_{f(t)}$ Hz is ($freq_{f(t)}/2$) Hz. Therefore, the maximum frequencies $freq_{Level}$ of signals c_j and d_j at each resolution level, are given by (8), where $freq_s$ is the sampling frequency.

$$freq_{Level} = \frac{freq_s}{2^{Level}} \tag{8}$$

The sampling frequency and amplitude for all types of disturbances considered in this approach are 15.36 kHz (256 samples per period – fundamental frequency of 60 Hz) and 1 p.u., respectively. The wavelet transform is applied to perform a 9-level decomposition of each discrete disturbance signal to obtain the detailed version coefficients ($d_1 - d_9$), and the approximated version coefficient (c_9). In this experiment, the adopted frequency bandwidths at each decomposition level are shown in Table 1.

Following, it is presented an example of a simple algorithm in MatLab® to demonstrate how the DWT is applied in practice (Mix & Olejniczak, 2003).

*% with f as the original signal; N the length of f; and **g** and **h** the wavelet filters*:
% for one decomposition level, the following is done:

```
f0=conv(g,f);              % convolution of g with f.
f1=conv(h,f);              % convolution of h with f.
f0=f0(1,2:N+1);            % eliminate first value.
f1=f1(1,2:N+1);
f0=reshape(f0,2,N/2);      % down sampling.
f1=reshape(f1,2,N/2);
c1=f0(1,:);                % the output c of length N/2.
d1=f1(1,:);                % the output d of length N/2.
```

This example code produces decomposed signals (c₁ and d₁) at level 1, which are the approximated and detailed version of the original signal f, respectively. In this example the signal f contains N samples, but the two derived signals c_1 and d_1 contain N/2 samples due to down-sampling.

Level	Parameter	Frequency band (Hz)	Harmonics included
9	$c_{9,k}$	0 – 15	-
9	$d_{9,k}$	15 – 30	-
8	$d_{8,k}$	30 – 60	1st
7	$d_{7,k}$	60 – 120	1st – 2nd
6	$d_{6,k}$	120 – 240	2nd – 4th
5	$d_{5,k}$	240 – 480	4th – 8th
4	$d_{4,k}$	480 – 960	8th – 16th
3	$d_{3,k}$	960 – 1920	16th – 32th
2	$d_{2,k}$	1920 – 3840	32th – 64th
1	$d_{1,k}$	3840 – 7680	64th – 128th

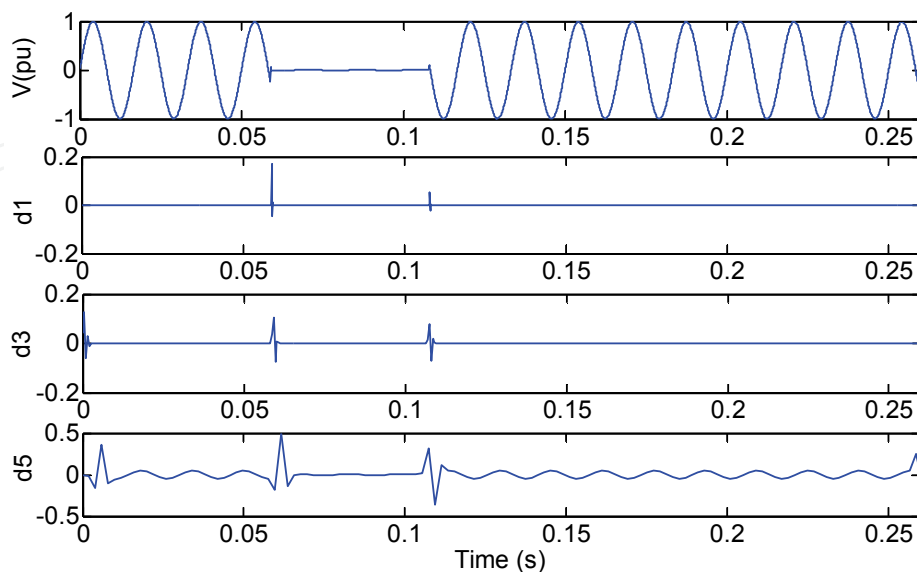Table 1. Scale to Frequency Range Conversion Based on 60 Hz Power Frequency

Fig. 3. The voltage interruption signal and detail coefficients: first decomposed level (d1), third decomposed level (d3), and fifth decomposed level (d5)

Figures 3, 4 and 5 show the decomposition plots of the signals using the "db4" wavelet. An example of interruption disturbance and the detail coefficients d1, d3 and d5 are shown in Fig. 3. Fig. 4 shows an oscillatory transient disturbance due to a capacitor switching transient, and detail coefficients d1, d2 and d3, respectively. Fig. 5 depicts a sag disturbance and detail coefficients d1, d3 and d5, respectively.
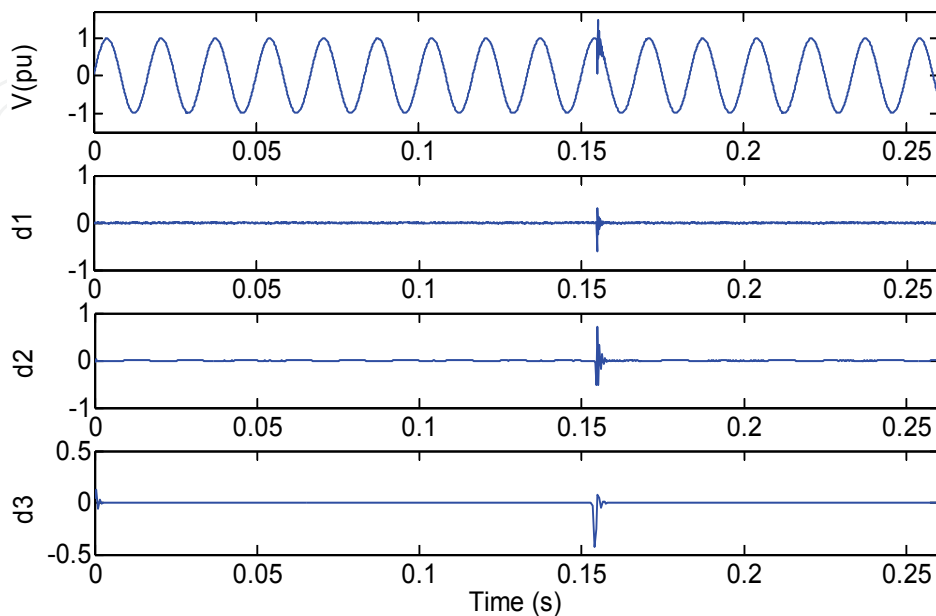

Fig. 4. The capacitor switching signal and detail coefficients: first decomposed level (d1), second decomposed level (d2), and third decomposed level (d3)
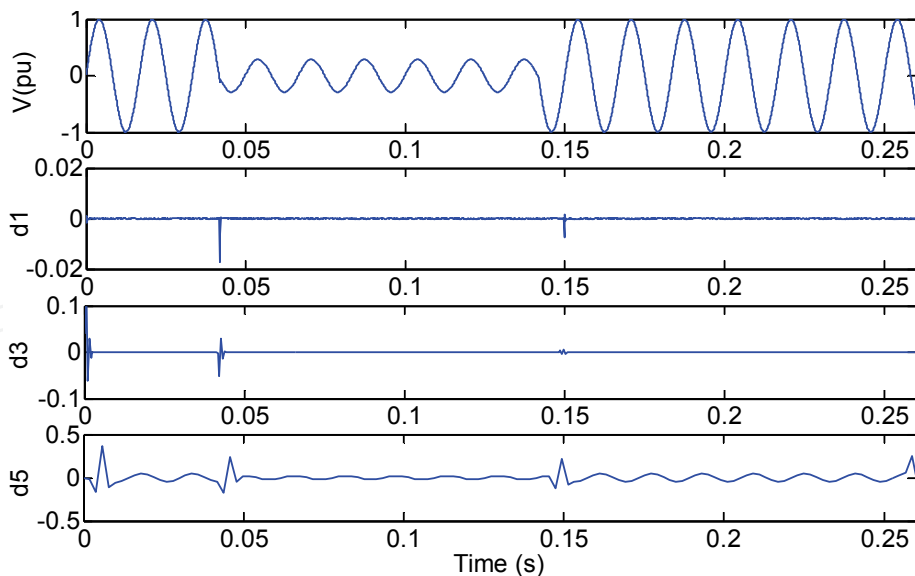

Fig. 5. The voltage sag and detail coefficients: first decomposed level (d1), third decomposed level (d3), and fifth decomposed level (d5)

2.2 Feature extraction
Using the DWT technique to analyze a signal through the level of MRA will generate severe variation in coefficients d_j. Therefore, applying the standard deviation at different resolution

levels one can quantify the variation magnitude of the signals (Oleskovicz et al., 2006; Kanitpanyacharoean & Premrudeepreechacharn, 2004). The standard deviation is defined as in (9).

$$\sigma_X = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}} \tag{9}$$

where X_i represents the detail coefficient data; $\overline{X}$ is the data average; and n is the amount of data.

Figure 6 shows a comparison of the two standard deviation curves along the 9 decomposition levels. One of these curves is from the analysis of a signal containing a sag disturbance, and the other curve is from a sinusoidal waveform analysis. By observing Fig. 6 one can notice there was a decrease in the coefficients of the 8th decomposition level, which concentrates most of the fundamental frequency components. While Fig. 7 shows the standard deviation curve of the detail coefficients from MRA analysis of an oscillatory transient signal caused by capacitor switching compared with a pure sinusoidal analysis. Differently from Fig. 6, which presents variations at the 8 level, Fig. 7 displays variations at the 3 and 4 levels, which concentrate high frequencies.
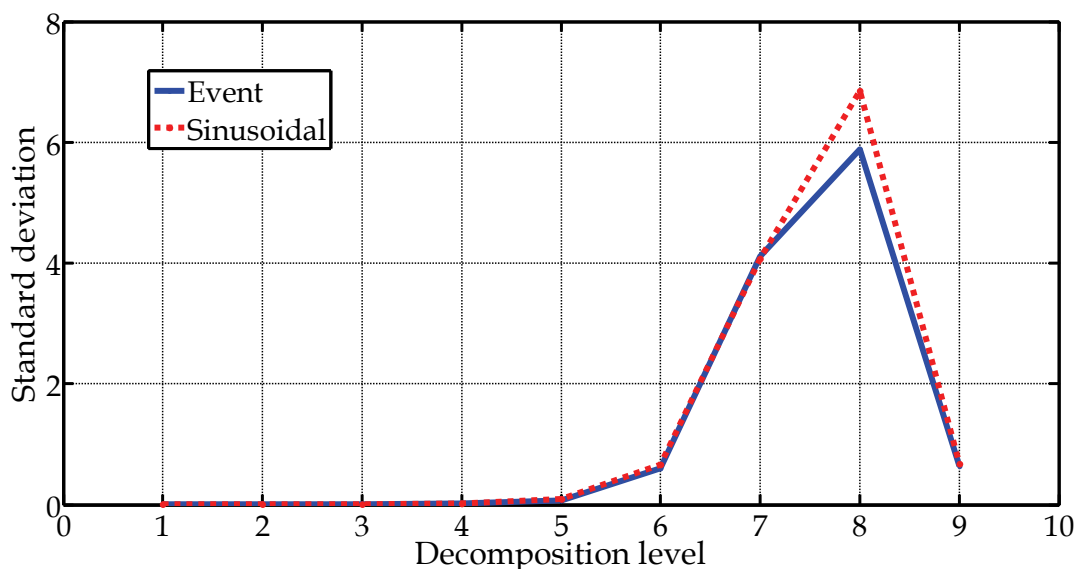

Fig. 6. Standard deviation curves – sag disturbance analysis compared with a sinusoidal waveform analysis

Several simulation tests were performed in order to determine the number of levels, hence reaching better results with 9 decomposition levels. With a lower number of levels some important features can be suppressed at lower frequencies. On the other hand, a higher number of levels can generate non-representative features since there is not sufficient information at the last levels due to the number of signal data considered in this paper. Additionally, in order to take into account these features that originated from the low frequency components (smaller than 60 Hz), this paper includes the approximated version, c_9, of the original disturbance signal f(t). Experimentally, better results were obtained when using the mean of c_9 rather then the standard deviation. The average of data is defined by (10).

$$\mu_X = \frac{1}{n}\sum_{i=1}^{n} X_i \tag{10}$$

where X_i represents the approximate coefficients and n is the total of data.

The features extracted from the standard deviation and the average value will compound the feature vector, $\mathbf{x} = \begin{bmatrix} \sigma_{d_1} & \sigma_{d_2} & \cdots & \sigma_{d_L} & \mu_{c_L} \end{bmatrix}^T$, of each signal, where L is the number of MRA levels.
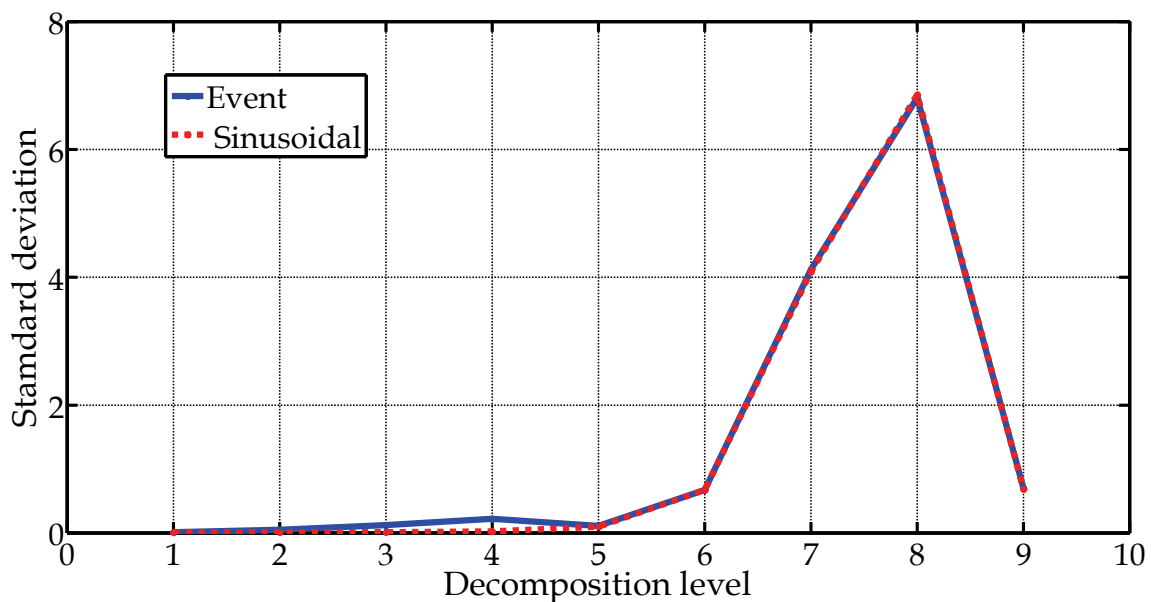

Fig. 7. Standard deviation curves – oscillatory transient disturbance analysis compared with a sinusoidal waveform analysis

2.3 Principal component analysis

The principal component analysis is a statistical technique whose purpose is to condense information from a large set of correlated variables into fewer variables ("principal components"), while preserving the variability that is in the data set (Jolliffe, 2002). Each component contains new information about the data set, and it is ordered in such a way that the first few components account for most of the variability.

The PCA transforms a random vector $\mathbf{x} \in \mathbb{R}^m$ into another vector $\mathbf{y} \in \mathbb{R}^n$ (for $n \leq m$), projecting $\mathbf{x}$ into n orthogonal directions with more variance. Generally, most of the data variance is explained by a reduced number of components, so it is possible to reject other components without losing relevant information.

Let $\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_M \end{bmatrix}_{N \times M}$ be a matrix formed by M feature vectors $\mathbf{x}$, previously defined, where $\mathbf{x}_1$, $\mathbf{x}_2$, $\cdots$, $\mathbf{x}_M$ are the M observation disturbance signals, and N=10 is the dimension of each feature vector. As in (Jolliffe, 2002; Castells et al., 2007), derivation of the principal components is based on the assumption that $\mathbf{x}$ is a random process characterized by the correlation matrix $\mathbf{R}_x = E[\mathbf{x}\mathbf{x}^T]$. Then, the principal components of $\mathbf{x}$ are obtained from the application of an orthonormal linear transformation $\mathbf{\Phi} = [\varphi_1 \varphi_2 \cdots \varphi_N]$ to $\mathbf{x}$,

$$\mathbf{w} = \mathbf{\Phi}^T \mathbf{x}, \tag{11}$$

so that, the elements of the principal component vector $\mathbf{w} = \begin{bmatrix} w_1 & w_2 & \cdots & w_N \end{bmatrix}^T$ become mutually uncorrelated. The first principal component is obtained as a scalar product $w_1 = \varphi_1^T \mathbf{x}$, where vector φ_1 is chosen so that the variance of w_1, is

$$E[w_1^2] = E[\varphi_1^T \mathbf{x}\mathbf{x}^T \varphi_1] = \varphi_1^T \mathbf{R_x} \varphi_1 , \tag{12}$$

maximally subject to the constraint $\varphi_1^T \varphi_1 = 1$. The maximal variance is obtained when φ_1 is the normalized eigenvector corresponding to the largest eigenvalue of $\mathbf{R_x}$, denoted by λ_1. The resulting variance is:

$$E[w_1^2] = \varphi_1^T \mathbf{R_x} \varphi_1 = \lambda_1 \varphi_1^T \varphi_1 = \lambda_1 , \tag{13}$$

The second principal component w_2 is obtained by choosing φ_2 as the eigenvector corresponding to the second largest eigenvalue of $\mathbf{R_x}$, and so on until the variance of $\mathbf{x}$ is completely represented by $\mathbf{w}$. Hence, to obtain the whole set of N different principal components, the eigenvector equation for $\mathbf{R_x}$ needs to be solved, according to (14).

$$\mathbf{R_x \Phi} = \mathbf{\Phi \Lambda} , \tag{14}$$

where $\mathbf{\Lambda}$ denotes a diagonal matrix with the eigenvalues $\lambda_1, \ldots, \lambda_N$. Since $\mathbf{R_x}$ is rarely known in practice, the $N \times N$ sample correlation matrix, is defined by

$$\hat{\mathbf{R}}_x = \frac{1}{M} \mathbf{X}\mathbf{X}^T , \tag{15}$$

replacing $\mathbf{R_x}$ when the eigenvectors are calculated in (13).

As previously mentioned, to obtain the principal components, it is necessary to calculate the eigenvalues and eigenvectors of $\hat{\mathbf{R}}_x$, which is the covariance matrix of the $N \times M$ data matrix $\mathbf{X}$. Therefore, once matrix $\mathbf{\Phi}$ is known it is possible to calculate the principal components. In this experiment, the 4 first principal components were adopted, which are w_1, w_2, w_3 and w_4, since their contributions correspond to more than 90% of the data variability, i.e., $\beta_P > 90\%$, defined according to (16) as

$$\beta_P = \frac{\sum\limits_{k=1}^{P} \lambda_k}{\sum\limits_{k=1}^{N} \lambda_k} 100\% , \tag{16}$$

where λ_k are the eigenvalues of $\hat{\mathbf{R}}_x$, $N = 10$ and $P = 4$.

Note that the PCA result is a set of independent elements that could be used as an orthonormal basis to approximate the original feature set extracted from the analyzed signals. It is possible, through a linear combination of such components, to generate those features used to characterize the analyzed signals. In this paper, the contributions of the first 4 principal components, resulting from the PCA technique, is 97.72% as the canonical weights to feature vectors in the training data set. Fig. 8 shows the principal components and their respective percentage of variances.
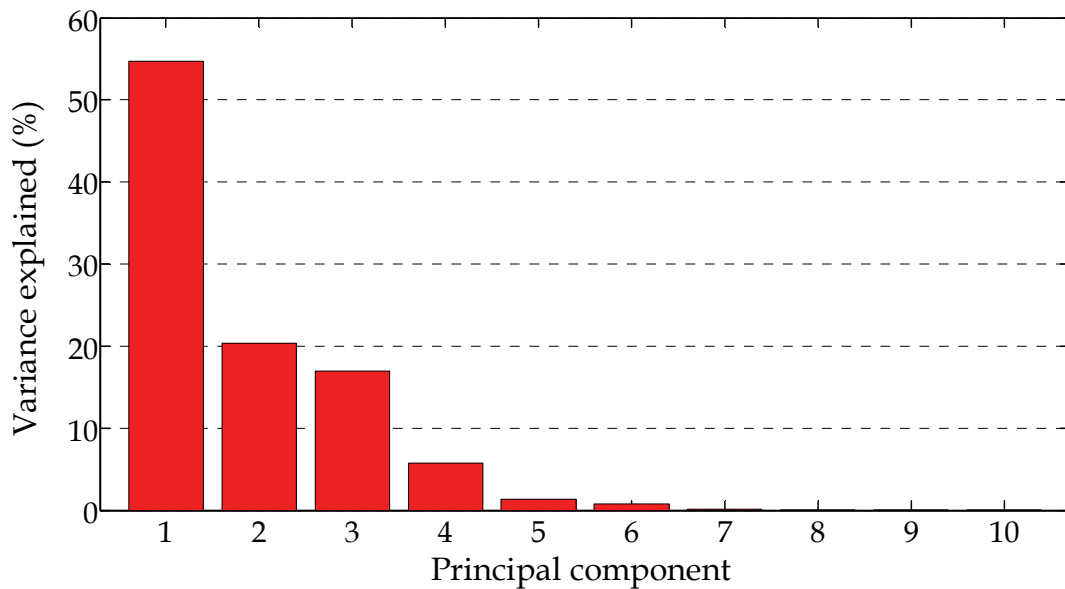
Fig. 8. Variance explained by principal components

The $\mathbf{W}$ matrix containing the principal components is applied to perform a new data set to represent each disturbance signal. These new set is calculated as in (17).

$$\mathbf{Y} = \mathbf{W}_{N \times P}{}^{T} \mathbf{X}_{N \times M} \, , \qquad (17)$$

where $\mathbf{X}_{N \times M}$ is the original data matrix, and $\mathbf{W}_{N \times P}$ is a matrix containing the P first components. Hence, $\mathbf{Y}$ is a $P \times N$ matrix representing a new uncorrelated data set.

2.4 The radial basis function network

The PQ disturbances classification is the last stage of the algorithm. In order to execute such classifications this paper proposes a radial-basis function (RBF) network. The classifications are performed by the RBF into two ways: one by not applying PCA algorithm and the other applying PCA algorithm so as to validate this proposed method.
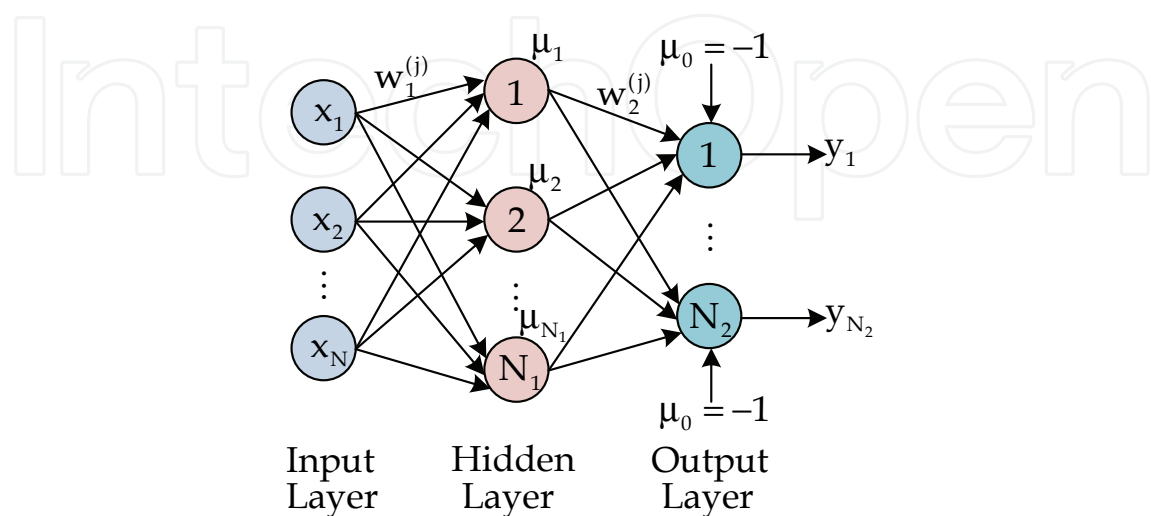

Fig. 9. A typical radial basis function network

Construction of a RBF network, in its most basic form, involves three layers with entirely different roles. The input layer is made up of source nodes connecting the network to its environment. The second layer, the only hidden layer in the network, applies a nonlinear transformation from the input layer space to the hidden layer space; in most applications the hidden space is of high dimensionality. The output layer is linear, supplying the network response to the activation pattern (signal) applied to the input layer (Fyfe, 1996). In other words, the hidden neuronal function forms a basis for the input vectors and output neurons merely by calculating a linear combination of the hidden neuron's output. A typical RBF network is shown in Fig. 9.

In pattern classification problems, the separation of classes in the RBF is accomplished with radial-basis functions. The maximum distance criteria are specified so as to obtain the classification. An often-used set of basis functions is the Gaussian functions whose mean and variance may be determined to a certain degree by the input data (Haykin, 1998).

2.4.1 The RBF training process

The RBF training consists of two stages, i.e., the hidden layer training process followed by the output layer training process. The RBF configuration for both classification procedures is as follows.

- Neurons at the input layer: feature vector lengths (4 or 10).
- Neurons at the hidden layer: 45.
- Neurons at the output layer: 7.
- Output learning rate: 10^{-4}.
- Output accuracy: 10^{-8}.

For the hidden layer training a non-supervised method was used, i.e., a typical grouping algorithm. The objective is to adjust the center of each Gaussian function in areas of input vectors tending to form groups. This procedure can be done by means of vector quantization using the "K-means" algorithm (Tou & Gonzalez, 1974; Haykin, 1998).

Training the output layer was only executed after of the function base parameters executed in the previous stage were determined. The training group of the output layer is formed by input/output pairs $\{\mu(k), d(k)\}$, whose vectors are specified after training the first layer. The output layer was trained using the "generalized delta rule" algorithm, which adjusts the network weights minimizing the quadratic error between $\mu(k)$ and $d(k)$ (Haykin, 1998). The Gaussian function usually used as an activation function for neurons of the hidden layer has the following form:

$$\mu_j = \exp\left(-\frac{\left(x(k) - w_1^{(j)}\right)^T \left(x(k) - w_1^{(j)}\right)}{2\sigma_j^2}\right),\tag{18}$$

where $j = 1 \cdots N_1$ are the inputs, $w_1^{(j)}$ are the synaptic weights connecting the input layer to the hidden layer, and σ_j^2 are the variances between $x(k)$ and $w_1^{(j)}$ calculated during the training of the hidden layer. Then, the maximum answer of each neuron should happen when $x(k)$ is very close to $w_1^{(j)}$. The answer decreases as $x(k)$ moves away from $w_1^{(j)}$, making the Gaussian cone narrower, as the variance σ_j^2 becomes smaller. Additionally, these cones have an identical output for inputs that are at a fixed distance of the Gaussian

function center. Finally, the outputs y_j of the exit layer neurons were obtained in the following way:

$$y_j = g\left(\sum_{i=0}^{N_1} w_{2ji}\mu_i\right) = \sum_{i=0}^{N_1} w_{2ji}\mu_i, \tag{19}$$

where $j = 1\cdots N_1$, y_j is the output of the jth neuron of the output layer, w_{2ji} is the synaptic weights matrix of the second layer, $g(\cdot)$ is the linear function activation, and μ_i is the output value of the ith neuron of the first neural layer (added by the threshold), being $\mu_0 = -1$.

3. Computational results

The classification results are presented in this section and the proposed method using the PCA algorithm is evaluated by considering two situations. Firstly, the RBF training and the PQ disturbance classification was performed without using the PCA algorithm. Then, in the second situation, the PCA algorithm was applied to perform the RBF training and classification. The processing time was obtained, as well as the number of training epochs, the mean squared error (MSE) and the classification accuracies in both situations so as to compare the proposed classification performance. Table 2 shows an overview of the RBF's computational performance obtained in the training and classification procedures. Table 3 shows the classification results of the first situation when the classification process was performed without using the PCA. The classification results with the PCA are shown in Table 4. In this situation the feature vector dimension was reduced, preserving those features with more variability, as presented in section II. In both tables each row represents a disturbance to be classified and each column shows which category that disturbance was classified in.

	PCA	Without PCA
Neurons at the input layer	4	10
Training epochs	48385	49065
Mean squared error	0.1223	0.1276
Training time	243.3 s	203.5 s
Classification time	44.9 ms	45.0 ms

Table 2. RBF classification performances

	C_1	C_2	C_3	C_4	C_5	C_6	C_7	(%)
Sinusoidal	25	0	0	0	0	0	0	100
Transient	0	25	0	0	0	0	0	100
Flicker	5	0	18	0	0	1	1	72
Harmonics	0	0	0	25	0	0	0	100
Interruption	0	0	0	0	23	0	2	92
Notching	2	0	0	0	0	23	0	92
Sag	1	0	0	0	1	0	23	92
Overall accuracy: 92.57 %								

Table 3. Classification results using PCA algorithm

	C_1	C_2	C_3	C_4	C_5	C_6	C_7	(%)
Sinusoidal	25	0	0	0	0	0	0	100
Transient	2	23	0	0	0	0	0	92
Flicker	6	1	16	0	0	1	1	64
Harmonics	0	0	0	25	0	0	0	100
Interruption	0	0	0	0	23	0	2	92
Notching	2	0	0	0	0	23	0	92
Sag	1	0	0	0	1	0	23	92
Overall accuracy: 90.29 %								

Table 4. Classification results without using PCA algorithm

4. Final comments

The main goal of this approach is to accomplish classifications of PQ disturbances, which are responsible for degrading the quality of electric power systems, by applying a RBF network. To best adapt the PQ disturbance signals, several stages of preprocessing were necessary. For instance, by applying the wavelet transform it was possible to analyze signals at different ranges of frequencies and to maintain the time resolution. This makes wavelet transformation a highly applicable tool in PQ analysis as broached in this paper. This is so because several disturbances happen with different characteristics in time and frequency, which are difficult to analyze with simple Fourier transform.

Among the several forms of analysis for the DWT coefficients, this paper proposes the calculation of the standard deviation curve of details and the average of approximation coefficients as a complement to this analysis. By observing Figs. 6 and 7, variations of the decomposition levels are clearly observable according to the disturbance contained in the signals.

The PCA tool has shown to be an effective and very interesting tool when it is applied to the pre-processing classification. By applying the PCA it was possible to reduce the amounts of data inserted into the RBF network, and it was also possible to choose which data were more significant and best representative of the disturbances. Although with a reduction of 60% in the feature vectors, the principal components could represent 97.72% of the information that characterized the disturbances. In addition, the method using PCA to perform a disturbance classification did not increase the processing time of the RBF network, as shown in Table 2.

Finally, taking into account the amount of disturbance classes considered in this paper, the classification results applying the proposed PCA method were better than those obtained when PCA was not applied. Using the same RBF configuration, the classification using PCA obtained 92.57% of success, while the classification not using PCA obtained only 90.29%.

As shown in Table 3, three disturbance classes were classified with 100% of success. The sinusoidal signals (without any type of disturbance), were inserted in the classification to discriminate those signals containing disturbances from those without any disturbances. Therefore, the signals containing some disturbance type were classified as not containing disturbance. This can be due to a very low disturbance intensity, which might have characterized the signals as pure sinusoidal signals.

5. Conclusion

In this work, wavelet and neural network methodology was employed to extract features and classify disturbance signals, respectively. Different from other works, which adopted similar approaches; this paper is concerned in eliminating redundant information from the feature space, therefore, reducing the number of input data and maintaining the processing time. The PCA technique was used to reduce feature space dimension, allowing the implementation of a neural network with a lower number of neurons, hence performing a better classification process.

By comparing the classification results and processing time from two systems discussed in this work, it was possible to demonstrate the validity of the PCA technique, which was able to classify the disturbances with an average accuracy of 92.57%.

6. Acknowledgment

7. References

Barmada, S., Landi, A., Papi, M. & Sani, L. (2003). Wavelet multiresolution analysis for monitoring the occurrence of arcing on overhead electrified railways, *Journal Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, Vol.217, No.3, (2003), pp.(177-187), ISSN 2041-3017.

Castells, F., Laguna, P., Bollmann, A. & Roig, J. M. (2007). Principal component analysis in ECG signal processing, *EURASIP Journal on Advances in Signal Processing,* Vol.2007, No.1, (January 2007), pp. (98-98), ISSN 1110-8657.

Chen, S. & Zhu, H. Y. (2007). Wavelet transform for processing power quality disturbances, *EURASIP Journal on Advances in Signal Processing,* Vol.2007, No.1, (January 2007), pp. (176-176), ISSN 1110-8657.

Chung, J., Powers, E. J., Grady, W. M. & Bhatt, S. C. (2002). Power disturbance classifier using a rule-based method and wavelet packet-based hidden markov model, *IEEE Transaction on Power Delivery*, Vol.17, No.1, (January 2002), pp. (233–241), ISSN 0885-8977.

Daubechies, I. (1992). *Ten Lectures on wavelet*, Society for Industrial and Applied Mathematics (SIAM), ISBN 0-89871-274-2, Philadelphia, PA, USA.

Daubechies, I. (1996). Where Wavelets Come From? – A Personal Point of View, *Proceedings of the IEEE*, Vol.84, No.4, (April 1996), pp. (510-513), ISSN 0018-9219.

Debnath, L. (2002). *Wavelet Transforms and Their Applications*. (1st Edition), Birkhäuser: Boston, ISBN 0-8176-4204-8, Boston, USA.

Dugan, Roger C., McGranaghan, Mark F., Santoso, Surya & Beaty, H. Wayne (2003). *Electrical Power Systems Quality* (2nd Edition), McGraw-Hill, ISBN 978-0-07-138622-7.

Fyfe, C. (1996). *Artificial Neural Networks*, (Edition 1.1), Department of Computing and Information Systems, The University of Paisley, pp. 113-126.

Gaing, Z.-L. (2004). Wavelet-based neural network for power disturbance recognition and classification, *IEEE Transaction on Power Delivery*, Vol. 9, No.4, (October 2004), pp. (1560-568), ISSN 0885-8977.

Haykin, S. (1998). *Neural Networks: A Comprehensive Foundation,* (2nd Edition), Prentice Hall, ISBN 0132733501, Saddle River, NJ, USA.

IEEE Standard 1159-2009 (2009). *IEEE Recommended Practice for Monitoring Electric Power Quality, (Revision of IEEE Std 1159-1995)* pp. (c1-81), June 26 2009, doi: 10.1109/IEEESTD.2009.5154067.

Jolliffe, I. T. (2002). *Principal Component Analysis,* (2nd Edition), Springer, ISBN 0-387-95422-2, New York, NY, USA.

Kanitpanyacharoean, W. and Premrudeepreechacharn, S. (2004). Power Quality Problem Classification Using Wavelet Transformation and Artificial Neural Networks, *Power Systems Conference and Exposition, 2004. IEEE PES*, ISBN 0-7803-8718-X, Vol.3, pp. (1496-1501), October 2004.

Karimi, M., Mokhtari, H. & Iravani, M.R. (2000). Wavelet Based On-Line Disturbance Detection for Power Quality Applications, *IEEE Transactions on Power Delivery*, Vol. 15, No.4, (October 2000), pp. (1212– 220), ISSN 0885-8977.

Mallat, S. G. (1989). A theory for multiresolution signal decomposition: the wavelet representation, *IEEE Transaction on Pattern Analysis and Machine intelligence,* Vol.11, No.7, (July 1989), pp. (674–693), ISSN 0162-8828.

Mix, D. F & Olejniczak, K. J. (2003). *Elements of wavelets for engineers and scientists.* (1stEdition), John Wiley & Sons, Inc., ISBN 0-471-46617-4, Hoboken, New Jersey,USA.

Oleskovicz, M., Coury, D. V., Carneiro, A. M. F, Arruda, E., Filho, O. D. &. Souza, S. A (2006). Estudo comparativo de ferramentas modernas de análise aplicadas à qualidade da energia elétrica, *Revista Controle & Automação,* Vol.17, No.3, (September 2006), pp. (331–341), ISSN 0103-1759.

Santoso, S., Powers, E. J., Grady, W. M. & Hofmann, P. (1996). Power quality assessment via wavelet transformation analysis, *IEEE Transaction on Power Delivery,* Vol.11, No.2, (April 1996), pp. (924–930), ISSN 0885-8977.

Santoso, S., Powers, E. J., Grady, W. M. & Parsons, A. C. (2000a). Power quality disturbance waveform recognition using wavelet-based neural classifier—Part 1: Theoretical foundation, *IEEE Transaction on Power Delivery,* Vol. 15, No.1, (January 2000), pp. (222–228), ISSN 0885-8977.

Santoso, S., Powers, E. J., Grady, W. M. & Parsons, A. C. (2000b). Power quality disturbance waveform recognition using wavelet-based neural classifier—Part 2: Application, *IEEE Transaction on Power Delivery,* Vol. 15, No.1, (January 2000), pp. (229–235), ISSN 0885-8977.

Strang, G. & Nguyen, T. (1997). *Wavelets and Filter Banks.* (2nd Edition), Wellesley-Cambridge Press, ISBN 0-9614088-7-1, Wellesley, USA.

Tou, J. T. & Gonzalez, R. C. (1974). *Pattern Recognition Principles*, Addison-Wesley, Reading, MA, USA.

Yilmaz, A. S., Subasi, A., Bayrak, M., Karsli, V.M. & Ercelebi, E. (2007). Application of lifting based wavelet transforms to characterize power quality events, *Energy Conversion and Management*, Vol.48, No.1, (January 2007),pp. (112-123), ISSN 0196-8904.

Zhu, T. X., Tso, S. K., & Lo, K. L. (2004). Wavelet-based fuzzy approach to power-quality disturbance recognition, *IEEE Transaction on Power Delivery,* Vol.19, No.4, (October 2004), pp. (1928–935), ISSN 0885-8977.
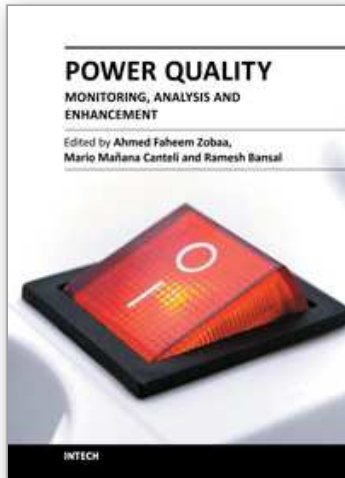
Power Quality â€"Â Monitoring, Analysis and Enhancement
Edited by Dr. Ahmed Zobaa

This book on power quality written by experts from industries and academics from various counties will be of great benefit to professionals, engineers and researchers. This book covers various aspects of power quality monitoring, analysis and power quality enhancement in transmission and distribution systems. Some of the key features of books are as follows: Wavelet and PCA to Power Quality Disturbance Classification applying a RBF Network; Power Quality Monitoring in a System with Distributed and Renewable Energy Sources; Signal Processing Application of Power Quality Monitoring; Pre-processing Tools and Intelligent Techniques for Power Quality Analysis; Single-Point Methods for Location of Distortion, Unbalance, Voltage Fluctuation and Dips Sources in a Power System; S-transform Based Novel Indices for Power Quality Disturbances; Load Balancing in a Three-Phase Network by Reactive Power Compensation; Compensation of Reactive Power and Sag Voltage using Superconducting Magnetic Energy Storage; Optimal Location and Control of Flexible Three Phase Shunt FACTS to Enhance Power Quality in Unbalanced Electrical Network; Performance of Modification of a Three Phase Dynamic Voltage Restorer (DVR) for Voltage Quality Improvement in Distribution System; Voltage Sag Mitigation by Network Reconfiguration; Intelligent Techniques for Power Quality Enhancement in Distribution Systems.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Giovani G. Pozzebon, Ricardo Q. Machado, Natanael R. Gomes, Luciane N. Canha and Alexandre Barin (2011). Wavelet and PCA to Power Quality Disturbance Classification Applying a RBF Network, Power Quality â€"Â Monitoring, Analysis and Enhancement, Dr. Ahmed Zobaa (Ed.), ISBN: 978-953-307-330-9, InTech, Available from: http://www.intechopen.com/books/power-quality-monitoring-analysis-and-enhancement/wavelet-and-pca-to-power-quality-disturbance-classification-applying-a-rbf-network

INTECH
open science | open minds

www.intechopen.com