

# Time-Frequency Transforms for Classification of Power Quality Disturbances

Alejandro Rodríguez, Jose A. Aguado, Jose J. López,  
Francisco Martín, Francisco Muñoz and Jose E. Ruiz  
*University of Malaga  
Spain*

## 1. Introduction

The usual operations on the distribution network such as switching loads and circuits, the proliferation of power electronic equipment and non-linear loads and the distributed generation with renewable energy are several of the most common causes that are leading to an increasing polluted power system in terms of voltage signal distortion.

One way of improving the power quality (PQ) parameters consists of analyzing these disturbances efficiently and understanding them deeply (Dugan, 2000) and PQ monitoring is one major task in order to achieve it. PQ monitoring is not an easy task usually involving sophisticated hardware instrumentation and software packages. Many recent approaches in PQ monitoring try to achieve it through the automated classification of different disturbances.

The different approaches in this field lead their efforts in two directions, the main parts that form an automated classification as depicted in Fig. 1. The first make focus to obtain a

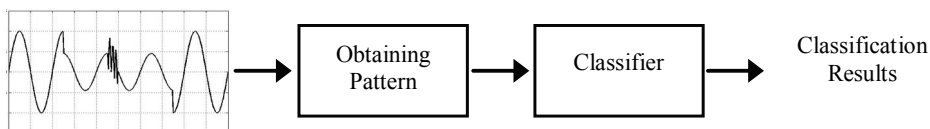


Fig. 1. Automated classification scheme

suitable pattern that allow distinguish clearly each disturbance, by the use of signal processing tool. Among existing signal processing tools the Fourier Transform (FT) results inadequate for analysis of non-stationary events and Short-Time Fourier Transform (STFT) although improves this drawback, it does not achieve good resolution in both time and frequency. Nowadays time-frequency transforms are used to get feature extraction, such as Wavelet transform (WT) (Santoso, 1994) and S-transform (ST) (Dash, 2003). WT extracts information from the signal in time and frequency domains, simultaneously, and provide greater resolution in time for high frequency components of a signal and greater resolution in frequency for the low frequency components of a signal. The ST can conceptually be

interpreted as a hybrid of STFT and WT. It uses variable window length and, by using the FT kernel, it can preserve phase information during the decomposition (Stockwell et al., 1996). The frequency-dependent window produces higher frequency resolution at lower frequencies, while at higher frequencies, sharper time localization can be achieved.

The second approach is oriented to use a classifier able to assign each disturbance correctly in its class, so the most of the artificial intelligent techniques have been combined with WT or ST, as Artificial Neural Networks (ANN), Decision Tree Fuzzy Logic, Hidden Markov Model, Support Vector Machines, etc.

In this chapter two different classification systems have been developed, using the WT and ST for pattern extraction, and an ANN as classifier algorithm.

The features obtained from WT are not completely distinctive and it is necessary to add features that give clear information about the signal magnitude. The real mean squared (RMS) value of the voltage signal have been obtained to achieve it.

On the other hand, the features obtained from ST analysis are sufficient to achieve a pattern that can properly classify the disturbances. In order to increase accuracy, simplicity and reliability, this chapter proposes a reduced and simple set of features extracted from the ST. Even in the presence of complex disturbances with different levels of noise, these features characterize the signals in a suitable way.

This chapter is organized as follows. In the second section the time-frequency transforms used in power quality are presented, particularly the WT and ST, and the obtaining pattern using these two transforms. A brief description of ANN used as classifier algorithm used in this chapter is given in section three. In section fourth the classification results obtained using the resulting classification system are presented. These systems are checked by signals obtained from electric power simulation in section fifth. Finally, conclusions are presented.

## 2. Time-frequency transforms and obtaining pattern

### 2.1 Fourier transform

The most used classical signal processing is the FT. This transform represents a signal as a sum of sinusoidal terms of different frequencies, named the frequency spectrum. This technique is suitable for stationary signals, but it is not efficient when the signal contents short-term transient disturbances.

In order to solve this drawback, a technique based on the FT is applied to short time intervals. This method is known as "Short Time Fourier Transform" (STFT), and consists of analyzing by the FT a sliding window of the signal. It is not possible reaching a good resolution in time and frequency simultaneously, and therefore it is necessary to adopt a compromise solution between the frequency and time resolutions.

The STFT is obtained by choosing a sliding window (short time interval) where the FT is applied. For narrow windows (or short time intervals) a good time resolution is obtained, suitable for short-term transients; on the other side a relative wide window enables a good frequency resolution but gives inaccurate time resolution.

Therefore, the problem lies in the fact that the window width is a parameter that must be fixed before analyzing the signal, before knowing what resolution is more suitable. At last, the total number of operations for computing STFT is  $N \cdot \log N$ .

### 2.2 Wavelet transform

Wavelet analysis is a powerful tool widely used in many scientific applications, especially in transient, non stationary, or time-varying situations.

A wavelet can be considered as a small wave which has its energy concentrated in time, and fits certain mathematic properties (Burrus et al., 1998). Fig. 2 shows three examples of wavelet.

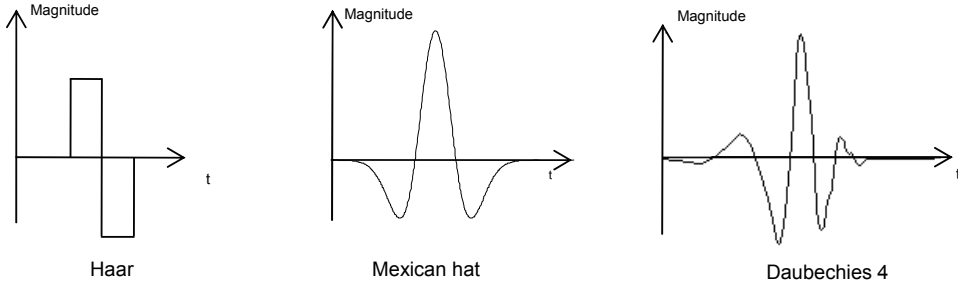


Fig. 3. Three examples of wavelet

The mother wavelet  $\psi(t)$  can be scaled and translated in the time, generating a family of functions named the wavelet expansion set (Wavelet system):

$$\psi_{j,k}(t) = 2^{j/2} \psi(2^j t - k) \tag{1}$$

A function  $f(t)$  can be expressed as a linear decomposition of this wavelet system as follows:

$$f(t) = \sum c_{j,k} \psi_{j,k}(t) , j, k \in Z \tag{2}$$

where  $j$  and  $k$  are integer indices,  $c_{j,k}$  are real coefficients, and  $\psi_{j,k}(t)$  is the expansion set. The set of expansion coefficients  $c_{j,k}$  is called the *wavelet spectrum* of the function  $f(t)$ .

If the wavelet expansion system is orthogonal, then:

$$j \neq l \text{ or } k \neq m \Rightarrow \langle \psi_{j,k}(t), \psi_{l,m}(t) \rangle = 0 \tag{3}$$

where  $\langle \rangle$  denotes the inner product defined as:

$$\langle \psi_{j,k}(t), \psi_{l,m}(t) \rangle = \int \psi_{j,k}(t) \cdot \psi_{l,m}(t) dt \tag{4}$$

When the expansion wavelet system is orthogonal the Wavelet spectrum can be computed as follows:

$$c_{j,k} = \int \psi_{j,k}(t) \cdot f(t) \cdot dt \tag{5}$$

The index  $j$  is related with the frequency and the index  $k$  with time. In practical applications a minimum frequency have to be established.

For a mother wavelet a basic scaling function  $\phi(t)$  is defined. A set of scaling functions is defined in terms of integer translates of  $\phi(t)$  by

$$\phi_k(t) = \phi(t - k) \tag{6}$$

A two-dimensional family of scaling functions can be defined by scaling and translation by

$$\varphi_{j,k}(t) = 2^{2j} \varphi(2^j t - k) \tag{7}$$

Let consider the minimum frequency is the corresponding to the value  $j=J_0$ . The equation (2) becomes to

$$f(t) = \sum_k \sum_j c_{j,k} \Psi_{j,k}(t) = \sum_k a_{J_0,k} \varphi_{J_0,k}(t) + \sum_{j=J_0}^{\infty} \sum_k d_{j,k} \Psi_{j,k}(t), \quad t \in R \tag{8}$$

where  $J_0$  is an integer. Equation (4) is a linear combination of wavelet coefficients,  $(a_{j_0,k}, d_{j,k})$ , a set of functions  $\varphi_{j_0,k}(t)$ , called *scaling functions* and  $\Psi_{j,k}(t)$ , called *wavelet functions*. Coefficients  $a_{j_0,k}$  and  $d_{j,k}$  are the *Discrete Wavelet Transform (DWT)* of  $f(t)$ , and can be calculated as:

$$a_{J_0,k} = \langle f(t), \varphi_{J_0,k}(t) \rangle \quad d_{j,k} = \langle f(t), \Psi_{j,k}(t) \rangle \tag{9}$$

Equation (4) can be truncated for  $j=J-1$ , obtaining:

$$f(t) = \sum_{k=0}^{2^{J_0}-1} a_{J_0,k} \varphi_{J_0,k}(t) + \sum_{j=J_0}^{J-1} \sum_{k=0}^{2^j-1} d_{j,k} \Psi_{j,k}(t), \quad t \in R \tag{10}$$

The first summation in (6) is a broad representation of  $f(t)$  that has been expressed as a linear combination of  $2^{J_0}$  translations of the *scaling function*,  $\varphi_{j_0,0}$ . The second summation contains the *details* of  $f(t)$ . For each level  $j$ , a linear combination of  $2^j$  translations of the *wavelet function*,  $\psi_{j_0}$ , are added to obtain a more accurate approximation of  $f(t)$ .

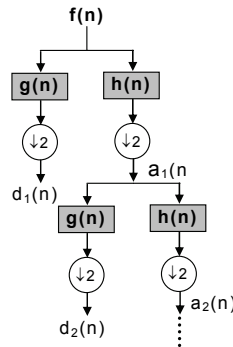


Fig. 2. Mallat algorithm schematics

The Mallat algorithm (Mallat, 1999) has been used in the practical implementation of DWT. The DWT acts as two FIR (Finite Impulse Response) quadrature filters defined by two sequences  $h(n)$  and  $g(n)$ .  $h(n)$  is a high frequency filter and  $g(n)$  is a low frequency filter. Both filter have the same cut frequency  $f_N/2$ , where  $f_N$  is the Nyquist frequency. Therefore the function  $f(n)$  is split in two parts, the high frequency part  $d_1$  that contains the higher octave and is called *detail function*, and the low frequency part  $a_1$ , that contains the frequencies lower than  $f_N/2$ , and is called *smoothed function*. Decimation by 2 is done for eliminating redundant information. The algorithm is iterated for  $a_1$ , obtaining a second level detail function  $d_2$  and a second level smoothed function  $a_2$ , that is again splitted, obtaining a series of detail and broad functions. The total number of operations for computing DWT is  $N$ .

The original function  $f(n)$  is split into a series of detail functions  $d_1, d_2, \dots, d_k$ , and a smoothed function  $a_k$ , that correspond to the frequencies:

$d_1: f_N - f_N/2; d_2: f_N/2 - f_N/4; \dots; d_k: f_N/2^k - f_N/2^{k+1}$ ; and  $a_n$  contains the frequencies lower than  $f_N/2^{k+1}$ .

Fig. 3 and Fig. 4 show an example of the application of the Mallat multi-resolution algorithm to the wavelet spectrum computation of a signal with a sag to 40%.

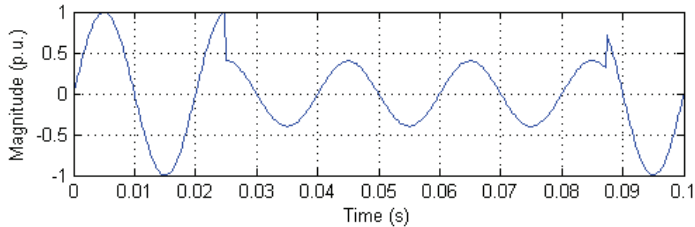


Fig. 3. Voltage sag to 40% and 400 samples length

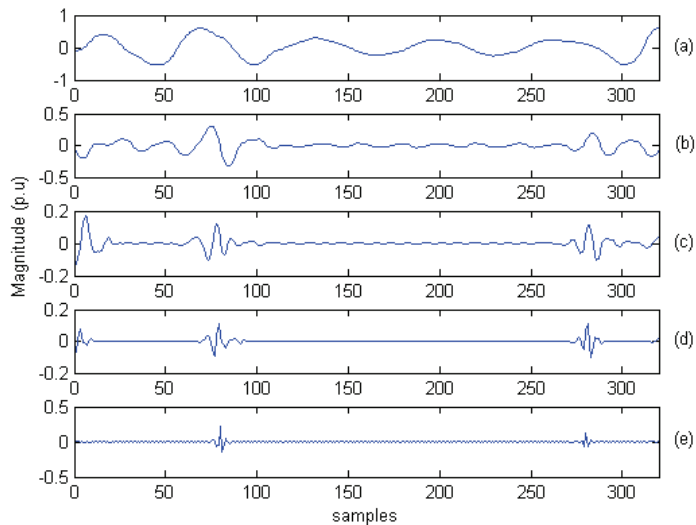


Fig. 4. Voltage sag detail levels wavelet analysis, from d1 to d5, (a) to (e), respectively

It is important to note that the wavelet is not a single specified function but a framework within which can design different wavelets. In this work, Daubechies 3 (db3) has been used as wavelet mother, performing 5 levels of decomposition. A voltage sag signal is shown in Fig. 3, and its wavelet transform decomposition, using db3, can be seen in Fig. 4.

### 2.3 Wavelet based pattern

The pattern used in this work is based on Parseval's Theorem. This Theorem states that the energy of a signal  $f(t)$  remains the same whether it is computed in a signal domain (time) or in a transform domain (frequency) as follows:

$$E_{\text{signal}} = \frac{1}{T} \int_0^T |f(t)|^2 dt = \sum_{n=0}^N |F[n]|^2 \quad (11)$$

where  $T$  and  $N$  are the time period and the length of the signal, respectively, and  $F(n)$  is the Fourier transform of the signal. In the case of the DWT, the signal is decomposed in terms of bands of frequencies, thus the energy of a signal can be given as:

$$E_{\text{DWT}} = \frac{1}{T} \int_0^T |f(t)|^2 dt = \sum_{l=-\infty}^{\infty} |a(l)|^2 + \sum_{j=0}^{\infty} \sum_{k=-\infty}^{\infty} |d_j(k)|^2 \quad (12)$$

with the energy in the expansion domain partitioned in time by  $k$  and scaled by  $j$ . The sampling frequency of the signals is 3.2 kHz, which is equal to  $64f_1$  where  $f_1$  is the power system frequency (50 Hz in Europe). The wavelet spectrum contains the total information of the original waveform as shown in Table 1.

DWT coefficients	Frequency band
$d_1$	$32f_1 \div 16f_1$
$d_2$	$16f_1 \div 8f_1$
$d_3$	$8f_1 \div 4f_1$
$d_4$	$4f_1 \div 2f_1$
$d_5$	$2f_1 \div f_1$

Table 1. Frequency band information contained in the wavelet spectrum.

In certain signals an energy-based pattern is not completely discriminatory because the energy of the magnitude disturbance depends on the depth of the disturbance and its duration in time.

The spectra of normal signals, sags and swells do not contain energy in the bands up to the fundamental frequency, as these signals only differ in the magnitude of the fundamental frequency component, and present very few energy in the high frequencies.

Therefore, it is necessary to provide a feature based on the magnitude of the signal. The RMS value is a widely accepted tool that provides information about how much the magnitude of the voltage changes. It is a fast and simple algorithm (13) that requires very few computational resources.

$$V_{\text{RMS}} = \sqrt{\frac{1}{N} \sum_{n=1}^N (f(n))^2} \quad (13)$$

where  $f(n)$  is the signal of length  $N$ .

The digital measurement instruments perform the computation of this parameter from instantaneous values of the samples, choosing a temporary window depending on the frequency of the steady state signal. If the RMS values are updated when a new sample is acquired, the method is called RMS continuum. If the RMS values are updated at a certain interval, usually half cycle, then it is called RMS (1/2). In this work the RMS (1/2) has been computed, selecting the maximum and minimum values. These values provide a feature based on the signal magnitude, as shown in Table 3.

Signal	RMS voltage values	
	Max	Min
Perfect signal	1.001	0.999
Normal to 0.91 in 4.5 cycles	1.001	0.909
Sag to 0.8 in 1 cycle	1.001	0.799

Table 3. Maximum and minimum RMS voltage value for different signals

Therefore the pattern is made up in two stages. The first part has five values, the energy of the signal in each frequency band of wavelet decomposition (feature extraction). These values have been normalized with reference to the values obtained from an ideal sinusoidal signal. The second part of the pattern is formed for two values the maximum and minimum RMS value calculated directly (feature selection) from each signal.

## 2.4 S-transform

The S-Transform (Stockwell et al., 1996) is a time-frequency transform generated by the combination of WT and STFT. The S-transform  $s(\tau, f)$  of the signal  $x(t)$  is defined as follows:

$$S(\tau, f) = \int_{-\infty}^{\infty} X(t) g_f(\tau - t) \exp(-j2\pi ft) dt \quad (14)$$

where

$$g_f(\tau - t) = \frac{|f|}{\sqrt{2\pi}} \exp\left(-\frac{f^2(\tau - t)^2}{2}\right) \quad (15)$$

$g_f(\tau - t)$  is the Gaussian window function,  $\tau$  is a shift parameter for adjusting the position in the time axis and  $f$  is the scale parameter.

$X(f)$  is defined as the Fourier transform of  $x(t)$ . The relationship between the S-transform and the Fourier transform is:

$$X(f) = \int_{-\infty}^{\infty} S(\tau, f) dx \quad (16)$$

The discrete ST is defined by:

$$S\left(\frac{n}{NT}, jT\right) = \sum_{m=0}^{N-1} X\left(\frac{m+n}{NT}\right) \cdot \exp\left(-\frac{2\pi^2 m^2 k^2}{n^2}\right) \exp\left(\frac{i2\pi nj}{N}\right) \quad (17)$$

In order to obtain the ST, the FFT of the original signal is computed [17]. The total number of operations for computing ST is  $N \cdot (N + \log N)$ .

The multi-resolution ST is a complex matrix whose rows and columns values are frequency and are time values, respectively. Each column represents the local spectrum in time. Frequency-time contours having the same amplitude spectrum are also obtained. This information is used to detect and characterize power disturbance events.

A mesh-dimensional output of the ST yields frequency-time, amplitude-time and frequency-amplitude plots. Examples of multi-resolution ST analysis for signals containing harmonics

and oscillatory transients are presented in Fig. 2 and 3, respectively. Each figure plots the disturbance signal (a), the time-frequency contours (b), and 3-D mesh giving amplitude, frequency and time plots (c).

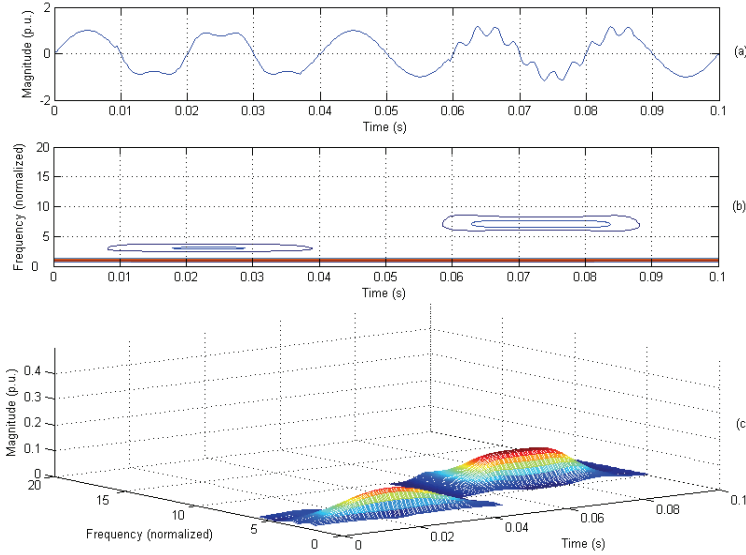


Fig. 2. Signal with 3rd and 7th harmonic content (a). S-transform Time-Frequency contours (b). 3D mesh Time-Frequency-Amplitude (c).

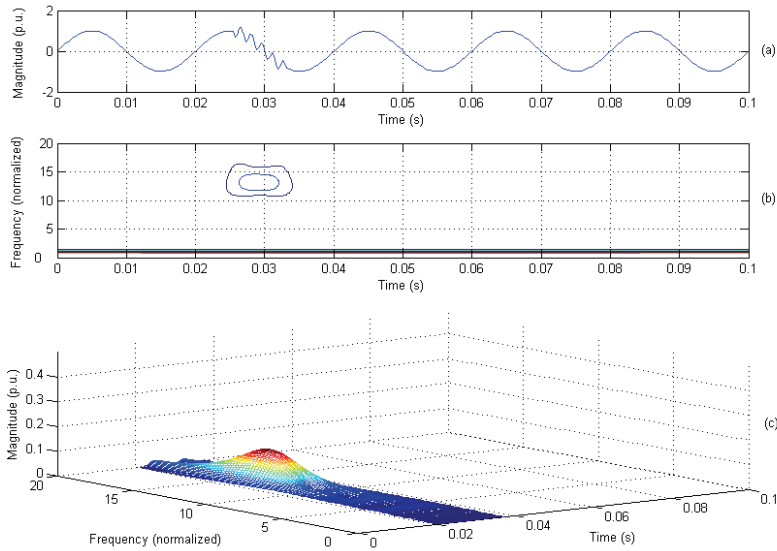


Fig. 3. Oscillatory transient (a). S-transform Time-Frequency contours (b). 3D mesh Time-Frequency-Amplitude (c).

## 2.4 S-transform based pattern

An efficient pattern can be defined from observation of ST contours. Below some examples of signals with different disturbances are analyzed in order to illustrate the pattern proposed in this approach.

A sinusoidal signal without any disturbance is depicted in Figure 4 (a). The fundamental frequency contour (b) shows a horizontal line. Other frequency contours (c), (d) and (e), corresponding to 150, 250 and 700 Hz show the non-existence of these frequencies in the signal.

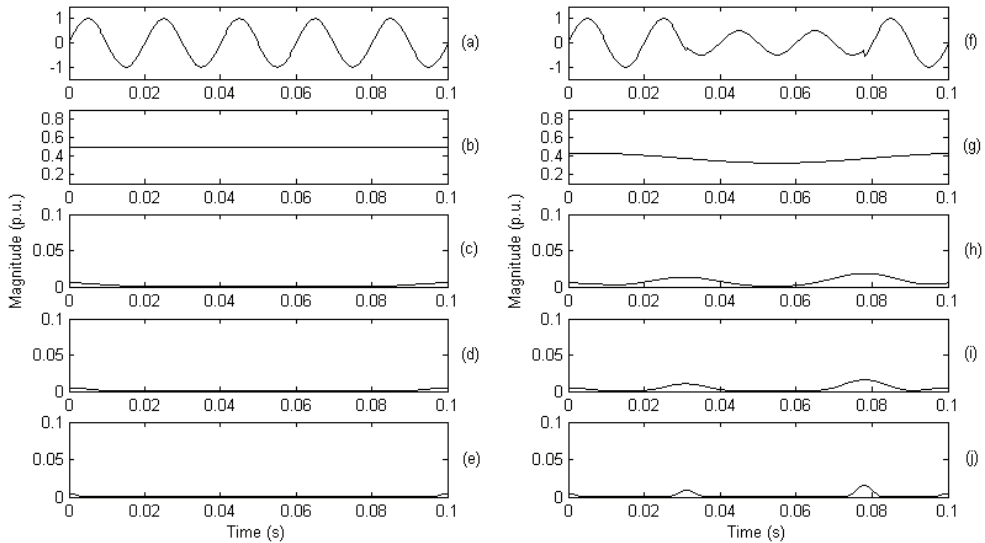


Fig. 4. Pure sinusoidal signal and 50 Hz, 150 Hz, 250 Hz and 700 Hz ST contours (a) to (e), and voltage sag, 50 Hz, 150 Hz, 250 Hz and 700 Hz ST contours from (f) to (j), respectively.

A sinusoidal signal with voltage sag and several frequency contours are depicted in Fig. 4 (f-j). The 50 Hz contour (g) is observed to decrease its value during the voltage sag. In the case of an interruption (not shown), a similar behaviour can be observed, but with a deeper diminution. The frequency contours corresponding to other frequencies (h-j) present a low energy value at the beginning and end of the sag.

An oscillatory transient disturbance is depicted in Figure 6 (a) and the 150 Hz, 250 Hz, 350 Hz and 700 Hz contours in (b) to (e), respectively. A big amount of energy can be noted in the contour corresponding to the frequency present in the transient signal, 700 Hz in this particular case.

Figure 6 (f) depicts a sinusoidal signal with a third and a fifth harmonic simultaneously, and several frequency contours. It clearly shows a big amount of energy in the contours corresponding to the frequencies present in the signal, 150 and 250 Hz, Fig. 6 (h-i).

The examples shown above illustrate the way of taking advantage of this particularity of the ST in order to search for specific frequency disturbances such as any order harmonics or particular oscillatory transients. The results shown above have been taken into account in the election of the characteristic features.

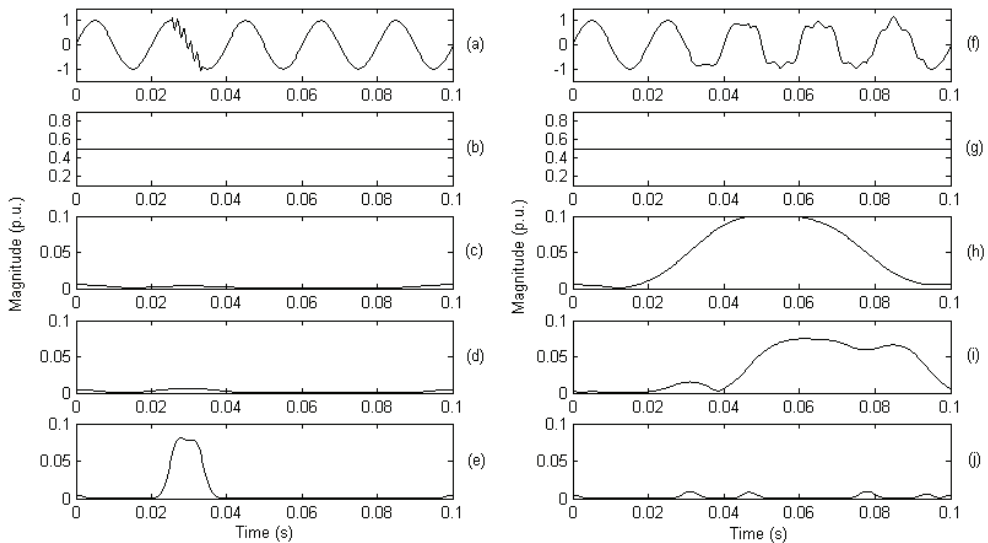


Fig. 4. Oscillatory transient, 50 Hz, 150 Hz, 250 Hz and 700 Hz ST contours (a) to (e), respectively, and signal with third and fifth harmonic component, 50 Hz, 150 Hz, 250 Hz and 700 Hz ST contours from (f) to (j), respectively.

The fundamental frequency contour has proven to contain valuable information about sags, swells and interruptions. Hence, the mean value of the 50 Hz contour has been taken as a distinctive feature. But this value does not clearly discriminate among sags and interruptions, and therefore the minimum value of the 50 Hz contour, which gives an idea of the severity of the disturbance, has been taken as the second feature.

In order to discriminate disturbances with presence of harmonics, the energy of the third, fifth and seventh harmonic (150 Hz, 250 Hz and 350 Hz) contours are used as distinctive features. This approach has been restricted to these frequencies although it could be extended to other harmonics within the Nyquist condition.

The sum of the energies from 600 Hz to 1600 Hz (Nyquist frequency) has also been taken as another characteristic feature. A high value of this energy gives information related to high frequency transient events.

A summary of the distinctive features used in ST based pattern is listed below:

- F1: Mean of the fundamental frequency contour (50 Hz)
- F2: Minimum of the fundamental frequency contour
- F3: Energy of the 3rd harmonic contour (150 Hz)
- F4: Energy of the 5th harmonic contour (250 Hz)
- F5: Energy of the 7th harmonic contour (350 Hz)
- F6: Sum of energy from 600 to 1600 Hz contours

### 3. Artificial neural network

As algorithm in the automated classification scheme has been selected ANN. Neural networks have emerged as an important tool for classification, and have been successfully applied to a variety of real world classification tasks in industry, business and science

(Widrow et al., 1994). Applications include classification of power quality disturbances (Borras et al. 2001).

An ANN is composed of very simple elements operating in parallel. These elements are inspired by biological nervous systems. As in nature, the network function is determined largely by the connections between elements. An ANN can be trained to perform a particular function by adjusting the values of the connections between elements. The ANNs can approximate any function with arbitrary accuracy, and they are nonlinear models, which makes them flexible in modeling real world complex relationships. Moreover ANNs are data driven self-adaptive methods in that they can adjust themselves to the data without any explicit specification of functional or distributional form for the underlying model (Zhang, 2000). In this chapter two different ANNs have been used, backpropagation and probabilistic.

### 3.1 Backpropagation

Feedforward backpropagation (BP) is a gradient descent algorithm, in which the network weights are moved along the negative of the gradient of the performance function. The term BP refers to the manner in which the performance function is propagated from the output to backward, and feedforward to the direction of the different connection between elements.

Once the network weights and biases have been initialized, the network is ready for training. The training process requires a set of examples of proper network behaviour, network inputs and target outputs. During training the weights and biases of the network are iteratively adjusted to minimize the network performance function. The default performance function for feedforward networks is mean square error, (i.e. the average squared error between the network outputs and the target outputs).

The BP has been set with two layers, one of them is a hidden layer. This structure can uniformly approximate any continuous function (Cybenko, 1989). The number of nodes in the hidden layer has been chosen as a function of number of inputs (Hecht-Nielsen, 1989), i.e. the number of pattern features, according to the expression  $(2n+1)$ , where  $n$  is the number of inputs. So for each pattern used, the hidden layer has different number of nodes, and the output layer only has one. The transfer functions for the hidden and output layer are tansigmoidal and lineal, respectively. The learning ratio has been 0.1, the epoch 3500 and the training algorithm Levenberg-Marquadt.

### 3.2 Probabilistic

The basic principle of probabilistic neural network (PNN) is implemented using the probabilistic model, such as Bayesian classifiers (Specht, 1990). The training examples are classified according to their values of probabilistic density function. When an input is presented, the first layer computes distances from the input vector to the training input vectors, and produces a vector whose elements indicate how close the input is to a training input. The second layer sums these contributions for each class of inputs to produce a vector of probabilities as output. Finally, a complete transfer function on the output of the second layer picks the maximum of these probabilities (Mishra et al., 2008). The most obvious advantage of this network is that training is trivial and instantaneous.

## 4. Classification results

Once the algorithm classifier is selected it has to be trained. At this stage, the classification algorithm is adjusted so that the function that assigns data to its corresponding class has the

lowest possible error based on a particular set of training data. In this way the system learns to match the output of the function with the tag or class for each data.

Once the classifier has been adjusted, the results are validated with a different data set, in order to check the effectiveness of the whole system.

The data set consists 6000 signals, where 10% are used to verify the effectiveness of the network, so 900 and 100 for each class, have been used for training and verification stage, respectively.

In this section the most usual power quality disturbances have been considered, sags, interruptions (int), swells, oscillatory transients (OT), harmonics (harm) and signals considered as normal voltage (normal). These signals are generated following mathematical models presented in (Gargoom et al., 2008), simulated in Matlab [Matlab, 2000] with a fundamental frequency of 50 Hz and a virtual voltage of 1V. The signals have a five cycles length (100 ms) and the sampling frequency is 3.2 kHz.

#### 4.1 Classification using WT based pattern and ANN

The resulting classification system using WT based pattern is presented in Fig. 8. It can be observed that the obtained pattern is composed of a feature selection (RMS value) and feature extraction. Selection chooses distinctive features from a set of candidates, while extraction utilizes some transformations, WT in this case, to generate useful and novel features from the original ones (Xu & Wunsch, 2005).

The classification results obtained are presented in Table 4 and 5, using BP and PNN as algorithm classifier, respectively.

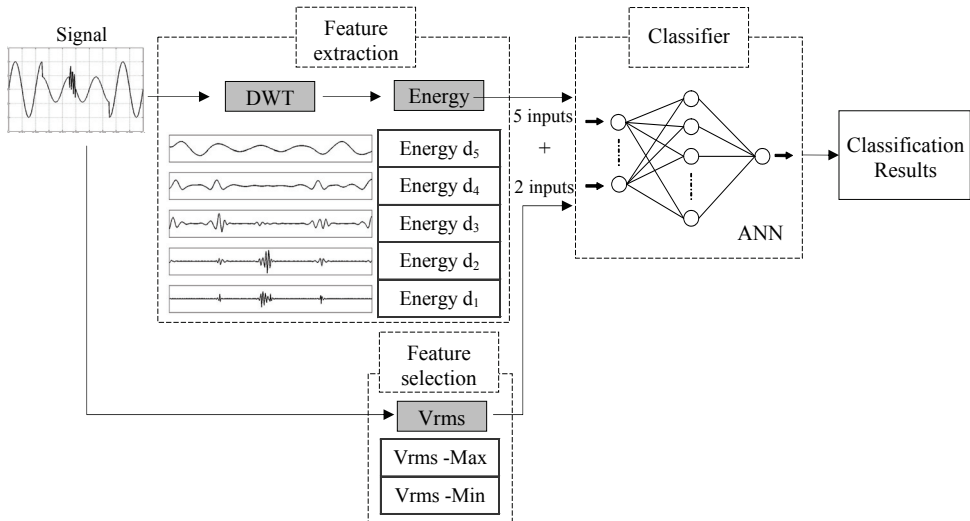


Fig. 8. Resulting classification system using wavelet based pattern and ANN

As can be observed the classification results are slightly better with BP than PNN except with high level of noise, 20dB. For disturbances, the worst results are obtained for sag and interruptions.

	No noise	40 dB	30 dB	20 dB
Normal	0.97	0.99	0.97	0.8
Sag	1	1	0.95	0.91
Int	0.99	0.97	1	0.94
Swell	1	1	0.99	0.98
OT	0.9	0.96	1	1
Harm	1	1	1	0.86
<b>Total</b>	<b>0.977</b>	<b>0.987</b>	<b>0.985</b>	<b>0.915</b>

Table 4. Classification results using WT based pattern and BP

	No noise	40 dB	30 dB	20 dB
Normal	1	0.99	0.99	0.99
Sag	0.91	0.93	0.96	0.96
Int	0.74	0.82	0.92	0.92
Swell	0.96	1	1	1
OT	1	0.96	0.99	0.99
Harm	1	1	0.99	0.99
<b>Total</b>	<b>0.935</b>	<b>0.950</b>	<b>0.975</b>	<b>0.943</b>

Table 5. Classification results using WT based pattern and PNN

**4.2 Classification using ST based pattern and ANN**

The resulting classification system using ST based pattern is presented in Fig. 9. The classification results are presented in Table 6 and 7, using BP and PNN as algorithm classifier, respectively.

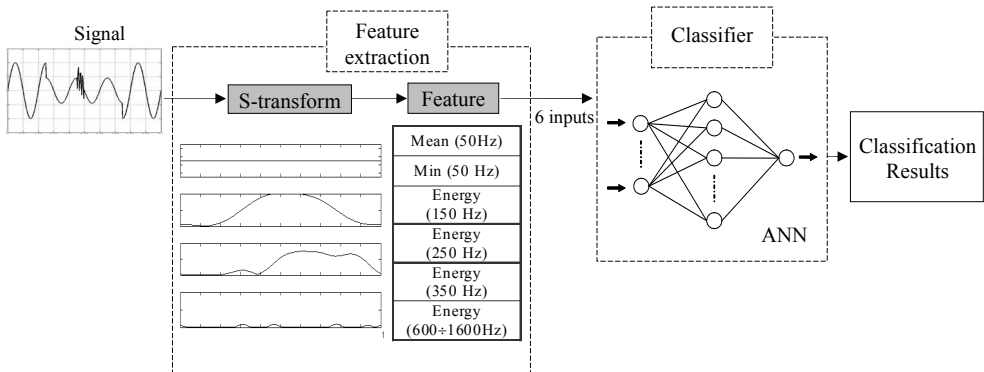


Fig. 9. Resulting classification system using ST based pattern and ANN

	No noise	40 dB	30 dB	20 dB
Normal	1	0.98	0.99	0.94
Sag	0.99	0.96	0.99	0.93
Int	0.79	0.98	1	0.99
Swell	0.97	0.95	0.96	0.99
OT	1	1	1	0.99
Harm	1	1	1	1
<b>Total</b>	<b>0.958</b>	<b>0.978</b>	<b>0.990</b>	<b>0.973</b>

Table 6. Classification results using ST based pattern and BP

	No noise	40 dB	30 dB	20 dB
Normal	1	1	1	0.99
Sag	0.79	0.65	0.57	0.62
Int	0.61	0.73	0.69	0.76
Swell	0.89	0.95	0.7	0.94
OT	0.62	0.97	1	1
Harm	0	0	0.64	1
<b>Total</b>	<b>0.652</b>	<b>0.717</b>	<b>0.795</b>	<b>0.885</b>

Table 7. Classification results using ST based pattern and PNN

The obtained results using ST based pattern are hardly better with BP than PNN. As can be observed the classification results are quite robust against the noise using BP. For disturbances the worst results are obtained for sag and interruptions.

Anyway, the results using ST based pattern and BP are better than WT based pattern and any ANN used, moreover the ST based pattern no needs additional features, as occurs with WT based pattern. On the other hand, the total number of operations to compute ST is greater than WT.

## 5. Checking the resulting classification systems

In this section, a data set of power quality disturbances has been generated using the power network simulation environment PSCAD/EMTDC (PSCAD, 2005). This application is an industry standard simulation tool for studying the transient behaviour of electrical networks.

The aim is to obtain different power quality disturbances of data set used for training and validation, in order to check the implemented system.

Several electrical systems with different events, typical and simple, as faults, switching big loads, non linear loads, etc., have been simulated to obtain different power quality disturbances. Moreover different signals can be obtained from a single system by using a facility of software simulation.

One example consists of a low voltage load fed by two parallel lines through a power transformer as shown in Fig. 6. A three phase fault occurs in the middle of one of the two

parallel transmission lines at a given time. Approximately two cycles after ( $t=0.045$  s) the fault is cleared. In Fig. 7 (a) can be seen the signal generated, where it can be observed an oscillatory transient produced by the fault, voltage sag during two cycles, and other oscillatory transient produced when the fault is cleared.

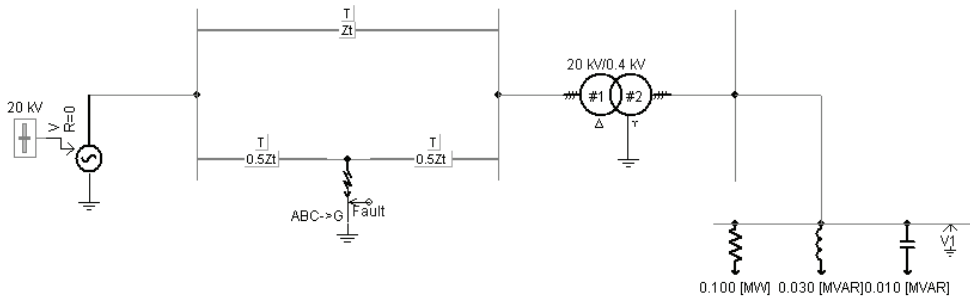


Fig. 6. 3 bus test system scheme in PSCAD

The signals shown in Fig. 7 (b)-(c) belong to initial data set used for training stage of classifier systems. In Fig. 7 (b) it is shown a normal voltage to 0.9, just in the boundary with voltage sag. Fig. 7 (c) depicted a very light oscillatory transient and, just after, a slightly magnitude diminution.

The values of the features in WT based pattern for signals of Fig. 7 (a)-(c) are shown in Table 8, in rows A to C, respectively. As can be observed the signal A(simulation) present a higher value of energy in relative high frequencies  $D1$  ( $32f_1 \div 16f_1$ ) than C, therefore A has an oscillatory transient. The minimum of RMS value for A, lower than B, gives information about the diminution of the magnitude, voltage sag.

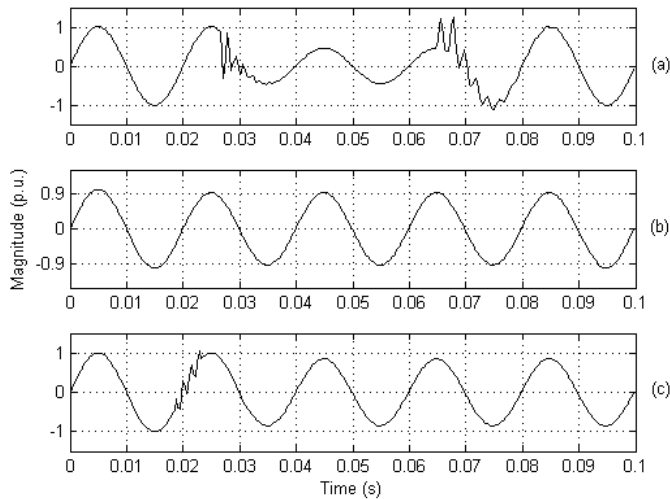


Fig. 7. Signal generated by model simulation (a). Normal voltage to 0.9 (b). Oscillatory transient and a slightly magnitude diminution.

From Table 8 the obtained values of the features for these signals, any efficient classifier can properly assign the disturbance to each corresponding class.

Signals	Energy					RMS	
	D1	D2	D3	D4	D5	Max	Min
A	407.9	336.6	3.164	0.920	0.675	1.043	0.457
B	2.426	1.281	0.959	0.885	0.897	1	0.9
C	351.4	27.22	0.722	0.794	0.801	1.012	0.849

Table 8. Feature values of WT based pattern for different signals

In order to compare the ST based pattern for these same signals are shown several ST contours. In Fig. 8 (a) it is shown the signal from simulation. The fundamental frequency ST contour for this signal and the normal one to 0.9 is showed in (b). The 700 Hz ST contour of the simulation signal and the oscillatory transient is depicted in (c). It can be observed that the simulation signal contours advise the presence of voltage sag, the 50 Hz ST contour for simulation signal (A) is below normal voltage to 0.9 (B), moreover an oscillatory transient because the energy in 700 Hz ST contour for simulation signals (A) is higher than oscillatory transient (C) used as reference. These graphs show that any efficient classifier can properly allocate the disorder to each corresponding class.

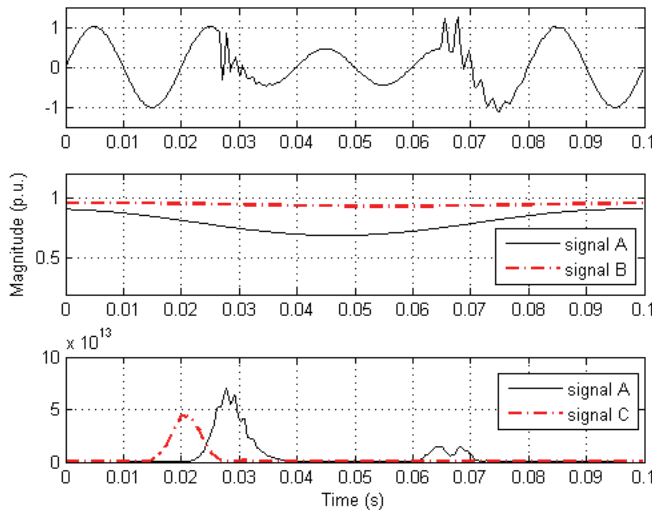


Fig. 8 . Signal generated from simulation (a), comparison of 50 Hz ST contours (b) and comparison of 700 Hz ST contours (c)

Others systems have been simulated using PSCAD and the power quality disturbances obtained are showed in Table 9, which also indicates the classification results using resulting classification systems.

These signals verify the behaviour of the implemented system with power quality disturbances based on electrical models, and therefore close to real ones.

Disturbance	Number	pattern-ANN		
		WT-BP	WT-PNN	ST-BP
Normal	10	0.90	0.90	1
Sag	40	0.875	0.80	0.9
Int	10	0.80	0.70	0.9
Swell	30	0.933	0.866	0.933
OT	30	0.866	0.933	0.933
Harm	30	0.933	0.933	0.933
Total	150	0.893	0.866	0.926

Table 9. Classification results for signals obtained by simulation software

As can be observed the results with signals obtained by electrical networks simulation are worse than those obtained with generated following mathematical models. The BP results do not worsen than PNN. At last the classification system using ST based pattern are better than WT based pattern.

## 7. Conclusion

Two different automated classification systems has been implemented, based on time-frequency transforms, WT and ST, in combination with Artificial Neural Network as algorithm classifier.

The classification results are very similar using two different algorithm classifiers, a BP and PNN, so it can be affirmed that is more important designing a properly and suitable pattern that the election of the algorithm. Any efficient classifier algorithm can do its expected job.

Therefore the extraction and selection features both are very crucial to the effectiveness of classification applications. Suitable features can greatly decrease the workload and simplify the subsequent process. Generally, ideal features should be of use in distinguishing patterns belonging to different class, immune to noise, easy to extract and interpret.

It can be appreciate that the detail signals of Wavelet transform analysis are very similar in the different disturbances, and in some cases almost identical, therefore becomes necessary to add other features, RMS value, which can distinguish magnitude disturbances. The S-transform generates several contours that put in evidence the nature of the disturbance, and these contours have certain characteristics that are suitable for automated pattern recognition.

On the other hand, the computational cost for Wavelet transform is less that for S-transform,  $N$  and  $N \cdot (N + \log N)$  respectively.

The implemented systems have been tested with signals generated with simulation program obtaining similar results.

A further limitation of many of the studies is that both training and testing are based in signals obtained following mathematical models. In this paper signals generated by power network simulation have been used to verify the classification system. In any case it becomes necessary to have power quality disturbances available to all researchers in this field that can be reliably used to compare classification results of each approach and not every researcher uses their own signals, whether synthetic, simulated or acquired.

The implemented system has been tested with signals generated with simulation program obtaining very good results.

## 8. References

- Borras, D.; Castilla, M. Moreno, N. & Montano, J.C. (2001) Wavelet and neural structure: a new tool for diagnostic of power system disturbances. *IEEE on Industry Applications*, Vol. 37 (Jan. 2001) pp. 184-190
- Burrus, C. S.; Gopinath, R. A.; Guo, H. (1998). *Introduction to wavelets and wavelet transforms*, Prentice Hall, ISBN: 0134896009, New Jersey, USA
- Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems*, No. 2 (1989) pp. 303-314
- Dash, P.K.; Panigrahi, B.K. & Panda; G. (2003). Power quality analysis using S-transform. *IEEE Trans. on Power Delivery*, Vol. 18, (Apr. 2003) pp. 406-411
- Dugan, R.S.; McGranaghan, M.F.; Santoso, S. & Beaty, H.W. (2000). *Electrical Power System Quality*, Wiley-Interscience, ISBN 007138622X, New York, USA
- Gargoom, A.M.; Ertugrul, N. & Soong, W.L. Automatic classification and characterization of power quality events. *IEEE Trans. on Power Delivery*, Vol. 23, (Oct. 2008) pp. 2417-2425
- Hecht-Nielsen, R. Counterpropagation networks. *Applied Optics*, No. 26 (Oct. 1989) pp. 4979-4984
- Mallat, S. (1999). *A Wavelet Tour of Signals Processing*, Academic Press, ISBN 0123743702
- MATLAB (2000). Natick, MA: Math Works, Inc., 2000.
- Mishra, S.; Bhende, C.N. & Panigrahi, B.K. (2008). Detection and classification of power quality disturbances using S-transform and probabilistic neural network. *IEEE Trans. on Power Delivery*, (Jan. 2008) Vol. 23.
- PSCAD (2005). Manitoba HVDC Research Centre, Inc., 2005.
- Santoso, S.; Powers, E.J. & Grady, W.M. (1994). Electric power quality disturbance detection using wavelet transform analysis, *Proceedings of IEEE - SP International Symposium on Time- Frequency and Time Scale Analysis*, pp. 166-169
- Specht, D.F. Probabilistic neural networks. *Neural Network*, (1990) Vol. 3, pp. 109-118
- Stockwell, R.G.; Mansinha, L. & Lowe, R.P. (1996). Localization of the complex Spectrum: the S-transform. *IEEE Trans. on Signal Processing*, No. 44, (1996) pp. 998-1001
- Widrow, B.; Rumelhard, D.E. & Lehr, M.A. (1994) Neural networks: Applications in industry, business and science. *Commun. ACM*, Vol. 37, pp. 93-105
- Xu, R. & Wunsch D. (2005). Survey of Clustering Algorithms. *IEEE Trans. on Neural Networks*. Vol. 16, No. 3 (May. 2005) pp. 645-678
- Zhang, G.P. (2000). Neural Networks for Classification: A Survey. *IEEE Trans. on Systems, Man, and Cybernetics - part c: applications and reviews*. Vol. 30, No. 4, (Nov. 2000) pp. 451-462