Abstract—In this paper some patterns based on discrete Wavelet transform are studied for detection and identification of both, low frequency disturbances, like flicker and harmonics, and high frequency disturbances, such as transient and sags. Daubichies4 Wavelet function is used as a base function to detect and identify due to its frequency response and time localization information properties. Based on these patterns, power quality disturbances are automatically classified by support vector machines (SVM). Thus, Radial Base Function (RBF) was used as a kernel, because RBF requires only two parameters ( and C) and cross validation technique and grid search were used in this work. SVM exhibit a good performance as classifier (90 percent of success for most disturbances) in spite of similitude between some disturbance patterns.

Index Terms—Bayes, discrete Wavelet transform, flicker, Fourier transform, harmonics, monitoring, neural networks, power quality, support vector machines, transients, voltage sags, voltage swells.

I. INTRODUCTION

Electromagnetic disturbances cause big economic losses for industry and residential users. Because of this, analysis and characterization of those phenomena are needed. Therefore, monitoring of power quality (PQ) disturbances of electrical energy is fundamental to offer solutions to industrial and to electrical areas. One of the most used signal processing techniques for power quality monitoring is Fourier Transform (FT), which is appropriate for PQ disturbance monitoring in stationary state, but it has limitations in the pursuit of non-stationary disturbances such as sags, transients (oscillatory or impulsive). Due to this, Wavelet Transform (WT) processing technique is proposed for power quality monitoring given its time-frequency multiresolution analysis property.

WT properties, like limited effective time duration, band pass spectrum, waveform similar to disturbance and orthogonality, allow locating information in time and frequency domains. Thus, it is possible to obtain high correlation when PQ disturbances occur and decompose these events into different components without energy aliasing between them.

There are several studies [1]-[7] where WT is used for detecting and identifying disturbances with Wavelet function Daubichies 4. Likewise, neuronal networks have been used to classify different disturbances from its WT. Waveforms of those disturbances are synthesized from mathematical models [8], [9]. Reference [10] shows a method of PQ disturbances detection and classification based on heuristic rules. However, there are no references about using other techniques of classification like Bayesian or support vector machines (SVM) for PQ disturbances.

In this article, mathematical concepts of Discrete Wavelet Transform (DWT) are described. The properties that make DWT effective for this study are also discussed. After this, strategies for PQ disturbances detection and identification by using DWT are studied. Strategies used for automatic classification of these disturbances are also presented. Finally, results of simulation and conclusions of this investigation are shown.

II. DISCRETE WAVELET TRANSFORM

Definitions and properties of DWT are described in this section. Here, advantages and disadvantages compared to other algorithms for detection and classification of PQ disturbances - such as Fourier Transform (FT) and Short-Time Fourier Transform (STFT) - are considered.

FT is not well adapted for the analysis of non-stationary signals. Thus, intervals (windows) are taken from their signal in order to analyze their frequency components [11]; that is, the signal \( x(t) \) is multiplied by a window function which moves in time \( h(t-b) \) and then FT is calculated, as it is expressed in (1):

Paper received on May 22, 2007. This work was supported by Grupo de Investigación en Sistemas de Energía Eléctrica (GISEL) from Escuela de Ingenierías Eléctrica, Electrónica y Telecomunicaciones (E3T) at Universidad Industrial de Santander (UIS).

Valdomiro Vega García is scholarship holder of Engineering Masters of Universidad Industrial de Santander, Bucaramaga, Colombia; e-mail: valdomirovega@ieee.org, valdomirovega@hotmail.com.

César A. Duarte G. is a full-time assistant professor of Universidad Industrial de Santander, Bucaramaga, Colombia; e-mail: cedagua@uis.edu.co.

Gabriel Ordóñez Plata is a full-time titular professor of Universidad Industrial de Santander, Bucaramaga, Colombia; e-mail: gaby@uis.edu.co.
\[(G \cdot x)(b, \xi) = \int_{-\infty}^{\infty} x(t) h(t - b) e^{-i2\pi dt} \quad (1)\]

Here, \( \xi \) represents the windowed signal frequencies whose time location is determined by \( b \). This analysis is limited by the time duration of \( h(t) \) and its frequency bandwidth. Therefore, (1) only allows the study of a fixed interval of the transient disturbance, but it is not possible to know its location, since the windows, both time and frequency, are wide constant.

In order to improve the results obtained from FT and STFT, a dynamic scheme is necessary where, in the same coordinates system, the width of time and frequency windows can be varied simultaneously preserving resolution in both domains (time and frequency). This characteristic is reached by means of the time-frequency multiresolution analysis that WT makes.

The Continuous Wavelet Transform (CWT) is defined in (2),[13]:
\[(W_c x)(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \psi \left( \frac{t-b}{a} \right) dt = \langle x(t), \psi_{a,b}(t) \rangle \quad (2)\]

Where \( \langle , \rangle \) denotes the inner product operation; \( \psi(t) \) is the "mother Wavelet function" or Wavelet of analysis of CWT. In (2) is considered that Wavelet function is a real value signal. The time location is determined by the term:
\[\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi \left( \frac{t-b}{a} \right) \quad (3)\]

Where \( \psi_{a,b}(t) \) is a set of Wavelets generated from the "mother Wavelet function" \( \psi(t) \), which expands and attenuates, or compresses and amplifies as \( a \) increases or diminishes, respectively. In addition, \( \psi(t) \) moves in the time domain as \( b \) changes.

The DWT is obtained by considering that parameters of scaling \( a \) and shifting \( b \) take discrete values: \( a = a_0 \cdot k \), \( b = k b_0 \), \( d_0 \), with \( j, k \in Z \) and \( a_0 > 1 \), \( b_0 > 0 \). By replacing these values in (2) we get [13]:
\[(W_c x)(j,k) = \frac{1}{\sqrt{a_0}} \int_{-\infty}^{\infty} x(t) \psi \left( a_0^{-j} t - k b_0 \right) dt = \langle x(t), \psi_{j,k}(t) \rangle \quad (4)\]

Where the set of Wavelet functions \( \psi_{j,k}(t) \) is given by:
\[\psi_{j,k}(t) = \frac{1}{\sqrt{a_0}} \psi \left( a_0^{-j} t - k b_0 \right) \quad (5)\]

However, for some appropriate "mother Wavelet functions" and factors \( a_0 \) and \( b_0 \), it is possible to express \( x(t) \) like a linear combination of Wavelet functions \( \psi_{j,k}(t) \), scaled and shifted. [13]:
\[x(t) = \sum_{j=2, k \in Z} (W_c x)(j,k) \psi_{j,k}(t) \quad (6)\]

The most used case for calculating DWT is the dyadic scale with \( a_0 = 2 \), \( b_0 = 1 \). In order to reconstruct \( x(t) \) according to (6) it is required that functions \( \psi_{j,k}(t) \) be orthonormal [11], [13].

### III. Algorithms

In this section the basic algorithms of signal decomposition and reconstruction by means of the Wavelet function are described.

The decomposition scheme is conformed by low-pass and a high-pass FIR filters, with impulse responses \( a_m \) and \( b_m \), respectively, followed by a two-decimation process. Therefore, if the samples \( c_n \) of the signal are at the entrance of filters, the coefficients of approximation \( c_{n,j} \) will be obtained at the output of low-pass filter and so will be the detail coefficients \( d_{n,j} \) at the output of high-pass filter [11]:
\[\begin{align*}
  c_{n,j} &= \sum_{m} a_{m-2j} c_{n,m} \\
  d_{n,j} &= \sum_{m} b_{m-2j} c_{n,m}
\end{align*} \quad (7)\]

With these Wavelet coefficients it is possible to reconstruct the signal by inserting zeros between samples. Then, these sequences are processed using low-pass and high-pass FIR filters, with impulse responses \( p_m \) and \( q_m \), respectively:
\[\begin{align*}
  c_{n,j} &= \sum_{m} (p_{k-2m} c_{n-1,m}) + \sum_{m} (q_{k-2m} d_{n-1,m})
\end{align*} \quad (8)\]

In Fig. 1, schemes of decomposition and reconstruction are shown.

This set of filters is determined by the "mother Wavelet function" selected for the transformation. The expression given in (6) is equivalent to (8), but (8) is used for signal samples decomposition.

In Fig. 2, magnitude of frequency responses for decomposition and reconstruction filters obtained from Wavelet function Daubechies 4 are shown.

### IV. Disturbances Detection and Identification

In this section the strategies to detect and identify PQ disturbances with DWT are analyzed. In [14]-[16] these disturbances are classified as: electromagnetic transient, flicker, sags (dips), swells, unbalances, interruptions, notching and frequency variations.

Each one of these disturbances can be detected by using details sequence from the first wavelet decomposition level
(Fig 1). By applying again the decomposition scheme to an approximation sequence, it is possible to find new details with smaller frequency span than that from the detail sequence previously calculated. Thus, different detail levels of the signal (or frequency intervals) can be analyzed.

In Fig. 3, the detection of sag on a 60 Hz signal by using the first detail level with Wavelet function daubechies4 "db4" is illustrated. Fig. 4 shows a set of disturbances studied in this paper and Fig. 5 displays their respective detail sequences from the first level. Notice that in Fig. 5 it is possible to detect the beginning and/or the end of each disturbance in the first level of detail. The reason for this is that both the beginning and the end of the disturbance contains high frequencies, which are detected mainly in the first level of detail.

On the other hand, it is possible to identify disturbances using Wavelet coefficients energy. Signal energy can be calculated from DWT coefficients in each decomposition level. Therefore, it is possible to know signal energy distribution in the frequency span of each decomposition level. Depending on sampling frequency (Fs) and Wavelet function (db4), the bandwidth of each decomposition level is determined. Frequency intervals of approximation (A) and detail (D) sequences for Wavelet function Daubechies 4 can be considered ideally, until the fourth decomposition level, as it is displayed in Fig. 6.

Reference [1] proposes a disturbance identification strategy that calculates energy distribution deviation for each decomposition level; that is, energy of coefficients of each level of detail is calculated (which is equivalent to the detail sequence energy), for both the pure sinusoidal signal and the signal with disturbances, then they are compared by the following expression:

\[ dp(j)(\%) = \left( \frac{En_{\text{dist}}(j) - En_{\text{ref}}(j)}{En_{\text{ref}}(m)} \right) \times 100 \]  

Where \( dp(j)(\%) \) is the deviation between the energy of the signal with disturbances \( En_{\text{dist}}(j) \) and its corresponding fundamental sinusoidal signal energy \( En_{\text{ref}}(j) \), at each wavelet transform decomposition level \( j \). \( En_{\text{ref}}(m) \) is the greatest value of the fundamental sinusoidal signal energy which may corresponds to a different level \( m \).

However, [1] does not propose any method to classify the different types of disturbances.

This identification strategy, described above, has been adopted in this article, since it allows to obtain disturbance patterns with a low degree of resemblance between them (such as it appears in Fig. 7 by evaluating (9)), which is desirable for their classification. Nevertheless, it must be noticed that swell and flicker patterns have similar characteristics. Likewise, when a given disturbance is shifted in time domain its pattern magnitude reveals significant variations. This is explained by the shift invariant property of WT.
30 Hz and disturbances in this work have no relevant information in this interval.

V. DISTURBANCES CLASSIFICATION

Artificial Neuronal Networks (ANNs), Fuzzy Logic or the combination of them have been proposed for PQ disturbances classification in [2], [10] and [12], among others. On the other hand, Bayesian technique [17] and support vector machines (SVM) have been used as pattern classifiers ([18],[19]), but not specifically for PQ events.

In this study, 4 classification techniques were implemented in order to automatically classify disturbances by using their patterns based on WT. These techniques are: multilayer perceptron (MLP) and kohonen ANNs, Bayes and MSV.

An ANN was trained with the following parameters: MLP network, feedforward performance function, tangent sigmoid activation function, 3 hidden layers, one exit layer and [8 6 4 1] neurons by layer. The 8 input are patterns based on 8 WT decomposition levels. 5 disturbance types sampled at a rate of 128 s/c (Fs=60*128 Hz). Each one of the 5 disturbance categories were sampled at a number between 1 and 5 that allows classifying these 5 disturbance types (randomly selected from the signals database) were studied. The output is a number between 1 and 5 that allows classifying these 5 disturbance types. ANN parameters were determined according to [20]. Network training was made with 4 800 inputs (600 inputs for each input layer neuron) and 600 outputs. Training validation was accomplished with 1 600 inputs (200 for each neuron at entrance layer) and 200 outputs.

In this work, a database with 19 430 synthetic signals was generated. A number of variations were considered to characterize different disturbances (magnitude, starting point, duration, frequency, etc.) [14]-[16].

Signals used for training, validating and evaluating ANN performance were selected from this signal database. However, each signal was used only once. The ANN previously mentioned is known as a supervised learning ANN because it is necessary to know the output corresponding to each input element.

On the other hand, there are non-supervised learning networks. In these networks, the output of every input set is not previously known, nor are the attributes that will be used to classify the disturbances. Because of this, non-supervised learning networks are very useful as category generators (clustering) [21].

A Kohonen network was trained in this work. Kohonen network belongs to competitive networks category or self organizing maps (SOM), that is, non-supervised learning network type. These networks have a two-layer architecture (input-output) (a single connections layer), linear activation functions and unidirectional information flow (cascade networks). This model is called LVQ (Learning vector quantization) [21]. This network was trained with the same input of that of the previous network, but output data set was not necessary. The maximum allowed error was set to 0,001 and 800 iterations were executed.

Bayes decision technique is the base of statistical methods for patterns recognition. This technique considers $K$ classes denominated $w_k$, and one input vector $X$. $P(w_k/X)$ is the *aposteriori* probability that can be calculated with Eq. (10), [17]:

$$P(w_k/X) = \frac{P(X/w_k) \cdot P(w_k)}{P(X)}$$

(10)

$P(w_k)$ is $w_k$ class probability (apriori probability). $P(X/w_k)$ is $X$ probability distribution conditioned to a particular $w_k$ class. $P(X)$ is $X$ probability distribution.

The decision rule can be stated as: “the correct class is the one that displays the greatest *aposteriori* probability”. From (10) a discriminating function given in (11) was implemented in this work.

$$g(X) = P(X/w_k) \cdot P(w_k)$$

(11)

This simplification is possible since $P(X)$ is equal to 20% because there are 5 classes (disturbances).

In recent years, SVM have shown good performance in patterns classification and recognition [18],[19]. In order to understand the way it operates, consider a data set distributed in two categories as it is shown in Fig. 8. The linear SVM look for a hyper-plane in such a way that the greatest number of points of the same category are located at the same hyper-plane side, whereas the distance (margin) of such categories to the hyper-plane is the greatest [18], [19], [21].

![Fig. 8 Hyperplane of separation](image)

There is only one optimal separation hyper-plane (OSH), so the distance from OSH to the closest training pattern (support vector) is the maximum [18], [19], [21]. In order to carry out pattern linearization and to make the pattern classification easier, a Radial Base Function (RBF) $k(x,y) = \frac{1}{2\sigma^2} e^{-\frac{||x-y||^2}{2\sigma^2}}$ was used as a kernel. RBF only requires a parameter ($\sigma$). In this work, crossed validation technique and the grid search were used [18], [19]. Parameters of penalty $C = 2^{22}$ and $\sigma = 0,7071$ were obtained. Classification results appear in Table II.

VI. SIMULATION RESULTS

In order to prove the effectiveness of each proposed detection and identification scheme, 200 disturbances of 5 types (randomly selected from the signals database) were analyzed. Success percentages are shown in Table I and II. Each one of the 5 disturbance categories were sampled at a 128 s/c (Fs=60*128 Hz).
A PQ disturbance detection and identification technique was implemented. This technique combines advantages of disturbances identification strategy based on DWT, with the advantages of the ANNs and SVM to classify information automatically. Once the disturbance is detected, it is possible to locate it from the detail sequence at first decomposition level (Fig. 5).

A database of 19,430 synthetic signals was generated, with different disturbances and different signal variations for training, validating and evaluating each classification scheme. The success percentage obtained in the evaluation of the strategy of detection, identification and classification, for most of the disturbances categories was better than 80% and 90%, in spite of shifting no-invariant property of WT. This makes changes in pattern magnitude when disturbance shifts in time domain.

SVM could be the best classifier for patterns obtained in this work. Though, ANNs (supervised) display good performance. Since no classifier is completely efficient when patterns of different disturbances are very similar, it is necessary to use a classification strategy that considers other signal parameters.

Bayes technique decision is not a good classifier for the patterns used in this work because input data are not close to a normal distribution.

VIII. REFERENCES

IX. BIOGRAPHIES

Valdomiro Vega García. Electrical engineer UIS, Master in Engineering (Candidate) UIS. Scholarship holder, Bucaramanga, Colombia. Student Member IEEE. Part-time Professor of Escuela de Ingenierías Eléctrica, Electrónica y Telecomunicaciones Universidad Industrial de Santander. GISEL researcher. Areas of Work: Signal processing and Power quality email: valdomirovega@ieee.org valdomirovega@hotmail.com

César Antonio Duarte Gualdrón. Electrical engineer UIS, Master in Electrical Power UIS, Bucaramanga, Colombia. Assistant Professor of Escuela de Ingenierías Eléctrica, Electrónica y Telecomunicaciones Universidad Industrial de Santander. GISEL researcher. Areas of Work: Signal processing, Power quality and education based on competences. E-mail: cedagua@uis.edu.co.

Gabriel Ordóñez Plata. Electrical engineer, UIS, Bucaramanga, Colombia. Doctorate Industrial Engineer UPCO, Madrid, Spain. Titular Professor Escuela de Ingenierías Eléctrica, Electrónica y Telecomunicaciones Universidad Industrial de Santander. GISEL director. Areas of work: Signal processing, electrical measurements, Power quality, technological management and education based on competences. Email: gaby@uis.edu.co.