Volume 7, Issue 2, February 2017 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Comparative Study of S and Wavelet Transforms in the Detection and Classification of Electrical Disturbances Using SVM

Ing. Miguel Angel Socorro*
Industrial Computers & UCI
Habana, Cuba

Dr. Antonio Cedeño
Industrial Computers & UCI
Habana, Cuba

DOI: 10.23956/ijarcsse/V7I2/01222

Abstract— The detection and classification of disturbances in electrical signals requires the use of methods and tools that allow a precise and complex analysis of the disturbed signal. This paper makes a comparative study of two tools, used in the analysis of electrical perturbations, known as Discrete Wavelet Transform (DWT) and Discrete Stockwell Transform (DST), both combined with Vector Support Machines (SVM) as classification technique and automatic diagnosis. In this investigation several electrical perturbations are modelled, which are examined using both transformations. Experiments show the superiority of DST in the detection, localization and classification of perturbations.

Keywords— Detection and classification of perturbations, Support Vector Machine, Stockwell Transform, Wavelet Transform.

I. INTRODUCTION

An electrical energy system is composed of different dynamic and complex elements that interact with each other and are susceptible to disturbances, produced by transient phenomena, which constitute more than 80% of the operating regimes of any electric network [1]. The simple connection or disconnection of the equipment, the variation of the power quantities of a network or the parameters that characterize its components; in addition to the failures produced, whether due to technological or environmental factors, cause abnormal conditions in the electrical networks that affect the quality of energy; with the consequent stress in the equipment giving rise to the heating and the vibrations, which can cause breakdowns and shorten the useful life of them. These abnormal conditions are known as disturbances in the electrical supply and can be grouped into seven different categories [2]:

In order to improve the quality of distributed electrical energy, it is necessary to detect and classify these disturbances. In this sense, different methods have been defined that allow the identification of such disturbances, evaluate the possible causes and to take the necessary measures.

These methods can be divided into two fundamental stages:

1) Detection and extraction of features using signal processing techniques.

One of the classic techniques used for this purpose in the field of signal processing is the Short Time Fourier Transform (STFT), which allows to obtain information about changes in the frequency spectrum of the signal during small intervals of time, determined by the width of the chosen window; but this width is fixed throughout the analysis and affects the resolution in time and frequency. This limitation is enhanced by the use of the Wavelet theory, which allows the analysis of the signal at various scales. But this type of analysis is not capable of providing a direct expression of the frequency components of the spectrum. A tool capable of improving these difficulties is the S or Stockwell Transform, which originated from the STFT and Wavelet theory, and allows a signal analysis at several frequencies [3].

Different works based on techniques to detect and extract features have been published in recent years. The Discrete Wavelet Transform (DWT) is used in [4] and [5] to extract and analyse the energy distribution in disturbed signals at different levels of resolution. The Discrete Stockwell Transform (DST) is used in [6] to extract and analyse statistical parameters of the signal as peaks of low and high frequencies. In [7] the DST is proposed to obtain vectors of instantaneous frequency in the signal and to analyse its energy.

2) Classification of disturbances using artificial intelligence techniques.

Making use of the features extracted from the signal, the type of disturbance is identified. The Appropriate tools to achieve this goal include automatic learning and classification and automatic diagnostics. Among the most commonly used tools for these topics are Artificial Neural Networks (ANN), Probabilistic Neural Networks (PNN) and Vector Support Machines (SVM).

The ANN are a paradigm of learning and automatic processing inspired by the way the biological nervous system works. They constitute a powerful tool for the classification and recognition of patterns, although they may present some disadvantages such as the need of a large number of training samples and the exit of solutions that have fallen at a local optimum [8]. The PNN were derived from Bayesian networks and Fisher's discriminant analysis and in comparison with ANN, are usually more accurate in their solutions, but in turn require more storage space [8].

The SVM are a classification method based on statistical learning theory and unlike ANN and PNN, which use the principle of empirical risk minimization to minimize training error; SVM are based on the principle of structural risk minimization which is equivalent to minimizing the upper limit of the generalization error of the classification model [9]. SVM also have some unique advantages that make them different from other automatic learning methods, especially in the treatment of problems that involve working with high dimensions and problems with few training samples [8]. These characteristics allow them to avoid problems that usually appear in classic classifiers such as curse of dimension and local minimums.

In this work a comparative study of the techniques and algorithms based on the DWT and DST is made to detect and classify affectations in the electrical quality using SVM as classifier.

II. EXTRACTION OF FEATURES IN DISTURBED SIGNALS

The features extracted from a signal are used as the input of a classification system instead of the signal waveform itself, because this usually leads to a much smaller system input. This paper it proposes the use of DWT and DST as tools to process the voltage signal and extract its distinctive features.

A. Discrete Wavelet Transform (DWT)

The continuous Wavelet Transform (CWT) of a signal x(t) is defining as:

$DWT(a,b) = \int_{-\infty}^{+\infty} x(t)\Psi_{ab}(t)dt$, where: $\Psi_{ab}(t) = \frac{1}{\sqrt{a}}\Psi(\frac{t-b}{a})$ with $a, b \in \mathbb{R}; a \neq 0$

The function $\Psi(t)$ is the base function or mother wavelet where a and b are the parameters of expansion and translation respectively. Then if instead dilate and transfer the mother wavelet continuously, are discretized the parameters of dilate and translate using the Nyquist rule, it obtains: $a = a_0^m$ and $b = nb_0a_0^m$, where a_0 and b_0 are fixed constants with $a_0 > 1$ and $b_0 > 0, m, n \in \mathbb{Z}$. This results in the following expression:

$DWT(m,n) = a_0^{-\frac{m}{2}} \sum_{-\infty}^{+\infty} x(t)\Psi(a_0^{-m}t - nb_0)$

A simple choice of the parameters a and b is to make $a_0 = 2$ and $b_0 = 1$. The resulting DWT according to these parameters is known as Diadical Wavelet Transform [10], and can be implemented using a technique known as Multiresolution Analysis (MRA).

The MRA allows making a representation of the signal at several scales and aims to develop this representation in terms of scales and wavelet functions, as shown below:

$x(t) = \sum_n C_{m,n}\emptyset_{m,n}(t) + \sum_m \sum_n d_{m,n}\Psi_{m,n}(t)$

Where $C_{m,n}$ are known as approximation coefficients, $d_{m,n}$ as coefficients of detail and $\emptyset_{m,n}$ the scale function [4].

B. Discrete Stockwell Transform (DST)

The Stockwell transform (ST) was introduced by R. G. Stockwell in [11] as a modification of the STFT with a variable window or an extension of the CWT with a specific mother wavelet multiplied by a frequency factor f.

Given a signal $x(\tau)$, the ST in its continuous form is defined as:

$S(t,f) = \int_{-\infty}^{+\infty} s(\tau)w(\tau - t, f)e^{-j2\pi f\tau}d\tau$

Where w is a window function defined as:

$w(\tau - t, f) = \frac{|f|}{\sqrt{2\pi}}e^{-\frac{f^2(\tau-t)^2}{2}}$

In practical applications where the captured signal is in its discrete form, the discrete version of the ST is obtained by making $f = \frac{n}{NT}$ and $t = KT$ leaving [12]:

$$S\left(KT, \frac{n}{NT}\right) = \sum_{m=0}^{N-1} F\left[\frac{m+n}{NT}\right]e^{-\frac{2\pi^2m^2}{n^2}}e^{\frac{i2\pi mk}{N}^2}$$

Where N is the length of the analyzed signal, the indices k, m, n are equal to $0,1,\ldots,N-1$, T is the time interval between two consecutive samples and $F\left[\frac{m+n}{NT}\right]$ is the Discrete Fourier Transform of the analyzed signal.

This discrete transform returns a matrix of complex numbers, where each row contains the frequency components at the different times sampled and the columns represent the local spectrum at that particular time.

C. Features extraction

The detail and approximation coefficients obtained from the DWT with the multiresolution analysis and the complex matrix obtained from the DST, cannot be used directly as input from a classifier. To reduce the size of these features it is necessary to use a feature extraction method. Among the features most commonly extracted to characterize the disturbances that affect the quality of the energy are:

- Main components of the signal (amplitude, frequency, initial phase).
- Statistics and probability distributions.
- Signal energy.

This paper proposes the use of signal energy as the main feature, since the other features mentioned do not identify the signal well in the presence of noise [4]. To obtain the energy of the signal, the Parseval identity is used, which relates the signal energy in the time domain to the energy of the coefficients of its transform in the frequency domain, as shown below [7]:

$$E_{signal} = \frac{1}{T}\int_0^T |s(t)|^2 dt = \sum_{n=0}^{N} |F_n|^2$$

Where T is the period of the signal, N is its length and F_n is the Fourier coefficients of the signal.

D. Obtaining signal energy using DWT

Once selected the scale function, the wavelet function and the multiresolution analysis have been, the Parseval identity allows relating the energy of the signal to the energy of the approximation and detail coefficients, as shown below [4]:

$$E_{signal} = \sum_n |C_m(n)|^2 + \sum_m \sum_n |d_m(n)|^2$$

This equation reflects the concentration of energy at each scale.

E. Obtaining signal energy using DST

The ST decomposes the signal in terms of its frequency components, obtaining for each frequency f_0 a spectrum of time-dependent values given by $S(t, f_0)$, where S is the ST and t is the time.

For the case of DST, the rows of its matrix represent this spectrum of values. Taking each row as a time-dependent function, for a given $f = \frac{n}{NT}$ frequency component, it is possible to calculate its energy as follows:

$$E_f\left(\frac{n}{NT}\right) = \sum_{K=0}^{N-1} \left|S\left(KT, \frac{n}{NT}\right)\right|^2$$

Where $n = 1 \dots \frac{N}{2}$, N is the length of the signal and E_f is the vector of energies associated with the frequency $\frac{n}{NT}$. The vector E_f contains the energies of the frequency component from the lowest frequency up to half of the signal sampling frequency (Nyquist rule).

III. CLASSIFICATION OF DISTURBANCES

Once the classifier is selected and the characteristics of the signal are extracted, the existing disturbance type is classified. Classification can be performed by introducing vectors to the classifier with the features extracted from the signal previously normalized.

A. Support Vector Machine (SVM)

Support vector machines are a method used for classification and regression's analysis. The goal of SVM is to construct a model using a set of training samples, each marked as belonging to one of the possible classes of classification. Using this model it is possible to predict the class to which a new sample belongs.

Given a set of n training samples samples x_i belonging in one of the following two classes $y = \{-1,1\}$. Using SVM it can be created a model that can optimally separate the samples belonging to one class from the other. To achieve this it looks for a hyperplane in a space where previously the samples have been projected, that separates them optimally. This hipeplane is defined as:

$$p_0: < w, x > + b = 0$$

Where $< w, x >$ denotes the scalar product, w is the normal vector to the hyperplane and b is a real coefficient. This hyperplane of separation is called optimal if its margin defined as the distance between this hyperplane and the closest samples of both classes, is maximum.

This margin can be expressed as:

$$d(x, w, b) = \frac{|< w, x_i > + b|}{\|w\|_2}$$

Then by scaling w and b such that:

$$|< w, x_i > + b| = 1$$

The margin would be defined as:

$$d(x, w, b) = \frac{1}{\|w\|_2}$$

And it is maximum when $\|w\|_2$ is minimum.

Thus, the problem of finding the optimal hyperplane can be solved by minimizing $\frac{1}{2}\|w\|_2^2$, subject to the restriction that the chosen hyperplane can separate the samples of the two classes correctly, that is:

$$y_i(< w, x_i > + b) \geq 1$$

Applying Lagrange's multipliers the proposed optimization problem is reformulated as:

$$L(w, b) = \frac{1}{2}\|w\|_2^2 - \sum_{i=1}^{N} a_i[y_i[< w, x_i > +b] - 1].$$

Where N is the number of training samples and a_i are the Lagrange's multipliers. The optimal solution to this problem is given by a saddle point, where L is minimized with respect to w, maximized with respect to $a_i > 0$ and maximized or minimized with respect to b according to the sign of $\sum_{i=1}^{N} a_i y_i$ [13].

It is possible that this problem has no solution (that is, there is no a separating hyperplane for a given set of samples), so it relaxes the condition allowing the misclassification of some samples by reformulating the problem as [14]:

Minimizing $\left(d(w,\varepsilon)=\frac{1}{\|w\|_2}+C\sum_{i=1}^{N}\varepsilon_i\right)$

Subject to:

$(y_i(<w,x_i>+b)\geq 1-\varepsilon_i),\ \ \varepsilon_i\geq 0\ \ \forall i=1,\dots,N$

Where ε_i is a slack variable and the parameter $C>0$ determines the ratio between the maximization of the margin of the hyperplane and the minimization of the error of classification.

The SVM can also perform a nonlinear discrimination between classes. If the samples in the input space cannot be linearly separated, then it is possible to do a linear discrimination but in a space of larger dimension, to do so, the input samples are mapped to a space of larger dimension by a function of nonlinear mapping; This space is called a characteristic space and is endowed with a scalar product. To achieve this mapping and to avoid explicit treatment of variables in the new space, kernel functions are used instead of the mapping function itself.

Solving the optimization problem gives the following classification function:

$$f(x)=\sum_{i=1}^{N}a_i y_i K(x_i,x)+b$$

Where $K(x_i,x)$ is the kernel function applied to the sample x_i.

Given then an unknown sample x can be classified using this function as follows:

$$x\in\begin{cases}Clase\ 1 & si \quad f(x)>0\\ Clase\ 2 & si \quad f(x)<0\end{cases}$$

IV. DESCRIPTION OF THE DEVELOPED ALGORITHM

For the study and analysis of the behavior of proposed tools and techniques for the classification of perturbations, a set of electrical perturbations are generated, according to the standard defined by the IEEE in [2], by the parametric equations shown in Table I and simulated using the Matlab development environment.

Table I Parametric Equations for Electrical Disturbances [8]

Disturbances	**Equations**	**Parameters**
Normal signal	$s(t)=sen(\omega_0 t)$	$\omega_0=2*\pi*50$
Voltage sags	$s(t)=\left(1-a\left((t-t_1)-(t-t_2)\right)\right)*sen(\omega_0 t)$	$0.1\leq a\leq 0.9, T\leq t_2-t_1\leq 9T$
Voltage swell	$s(t)=\left(1+a\left((t-t_1)-(t-t_2)\right)\right)*sen(\omega_0 t)$	$0.1\leq a\leq 0.8, T\leq t_2-t_1\leq 9T$
Interruption	$s(t)=\left(1-a\left((t-t_1)-(t-t_2)\right)\right)*sen(\omega_0 t)$	$0.9<a\leq 1, T\leq t_2-t_1\leq 9T$
Harmonics	$s(t)=sen(\omega_0 t)+\sum_{i=2}^{n}k_i sen(i\omega_0 t)$	$0.05<k_i\leq 0.3, i=3,5,7$
Transients	$s(t)=sen(\omega_0 t)+\alpha e^{\left(-\frac{(t-t_1)}{\tau}\right)}\left((t-t_1)-(t-t_2)\right)*sen(2\pi f t)$	$0.1\leq a\leq 0.8$, $0.5T\leq t_2-t_1\leq 3T$, $300\leq f\leq 900$
Voltage fluctuations	$s(t)=\left(1+\alpha sen(2\pi\beta t)\right)*\ sen(\omega_0 t)$	$0.1\leq a\leq 0.2, 0.1\leq\beta\leq 0.5$
Sags + harmonics	$s(t)=\left(1-a\left((t-t_1)-(t-t_2)\right)\right)*(sen(\omega_0 t)+\sum_{i=2}^{n}k_i sen(i\omega_0 t))$	$0.1\leq a\leq 0.9, T\leq t_2-t_1\leq 9T$ $0.05<k_i\leq 0.3, i=3,5,7$
Swell + harmonics	$s(t)=\left(1+a\left((t-t_1)-(t-t_2)\right)\right)*(sen(\omega_0 t)+\sum_{i=2}^{n}k_i sen(i\omega_0 t))$	$0.1\leq a\leq 0.8, T\leq t_2-t_1\leq 9T$ $0.05<k_i\leq 0.3, i=3,5,7$

Nine types of different disturbances are considered, including the normal signal without disturbances: normal signal, voltage sags, voltage swell, interruption, harmonics, transients, voltage fluctuations, voltage sags + harmonics and voltage swell + harmonics, denoted as: P1, P2, P3, P4, P5, P6, P7, P8 and P9 respectively. The signals are simulated for one second at a fundamental frequency of 50 Hz.

Using the Parseval's theorem the energy in each signal is mathematically expressed in terms of the wavelet coefficients and in terms of the spectrum of values of the S transform. According to this representation, the energy concentration of the signal in each wavelet scale and in each frequency component given by the DST is obtained. In this article, the energy's difference (ΔE) contained in the wavelet coefficients at each scale between the disturbed signal (DES_e) and the normal signal (NES_e) is adopted as vector of features of the disturbed signal; as well as the energy difference in the spectrum of values of the DST, associated with each frequency component, between the disturbed signal (DES_f) and the normal signal (NES_f).

This vector of features x is mathematically defined as:

$$x_w = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} DES_{e1} \\ \vdots \\ DES_{en} \end{bmatrix} - \begin{bmatrix} NES_{e1} \\ \vdots \\ NES_{en} \end{bmatrix} = \begin{bmatrix} \Delta E_1 \\ \vdots \\ \Delta E_n \end{bmatrix}, \; x_s = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} DES_{f1} \\ \vdots \\ DES_{fn} \end{bmatrix} - \begin{bmatrix} NES_{f1} \\ \vdots \\ NES_{fn} \end{bmatrix} = \begin{bmatrix} \Delta E_1 \\ \vdots \\ \Delta E_n \end{bmatrix}$$

Where x_w is the energy vector for the DWT and x_s is the energy vector for the DST.

The mother wavelet used is the Daubechies with length 0, 1, 4 and 6 (haar, db1, db4, db6) in 8 levels of decomposition. For classification with SVM, the libSVM library developed in [15] is used. This library is trained with 100 energy samples for each type of signal generated, for a total of 900 training samples and as kernel function is chosen the Gaussian kernel proposed in [16] and [17].

V. EXPERIMENTATION AND RESULTS

To initiate the experiment a total of 200 samples are generated during one second for each one of the nine proposed disturbances; for a total of 1800 samples to be classified differents from the 900 used to train the libSVM library.

Then, a vector of features based in energy for each signal generated is extracted. Each vector is extracted, firstly from the wavelet decomposition of the signal with mother wavelet Haar, db1, db4, db6 and then from the spectrum of values of the DST present in the time-frequency matrix calculated by this.

Table II shows the number of well classified samples for each disturbance analysed with either DWT or DST and the percentage that this represents.

Table II disturbances well classified and per cent that it represents

	P1	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**	**P8**	**P9**	**%**
db0	200	186	192	75	200	175	200	197	193	89.9
db1	200	184	189	71	200	154	200	198	191	88.2
db4	200	164	199	168	200	174	200	196	198	94.4
db6	200	178	200	125	200	168	200	194	197	92.3
ST	200	188	197	176	200	184	200	194	200	96.6

As can be seen in Table II, the number of well classified interrupts (P4) tends to be lower than the others; this is because the interruptions can be confused with the voltage drops (P2). The best percentage of classification is obtained by the DST, followed by the wavelet decomposition with mother wavelet db4. Figure 1 shows these results.

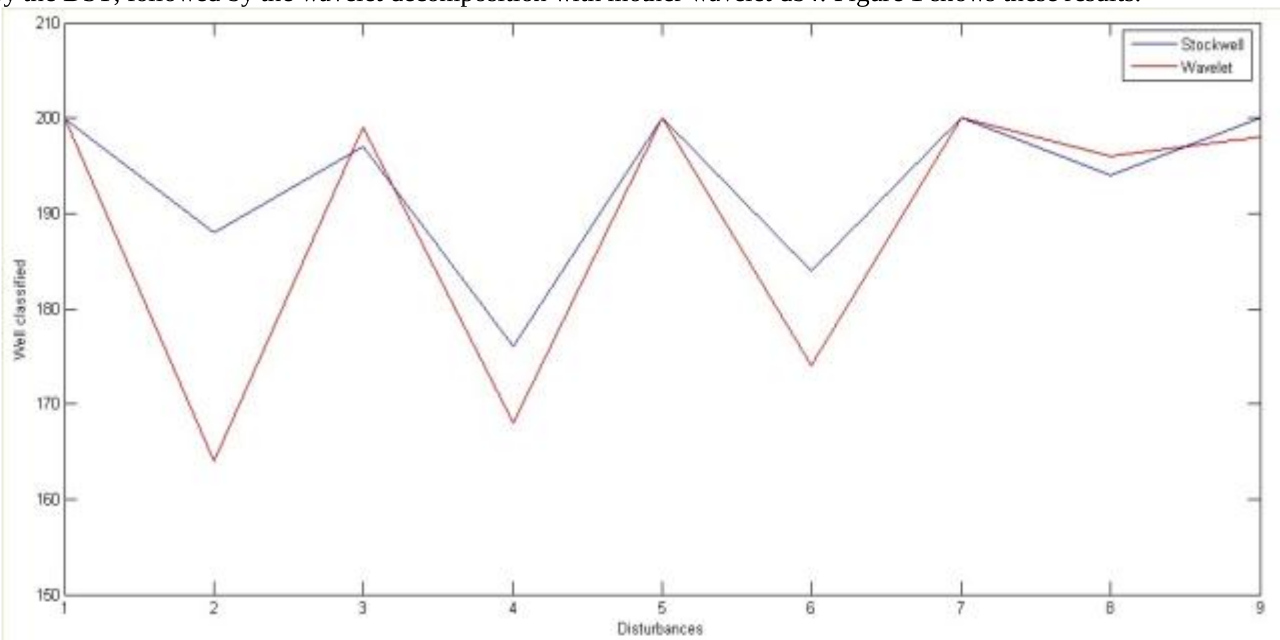

Fig. 1 Representation of perturbations well classified using DST (blue) and DWT (red).

Another aspect to be compared is the effectiveness in the detection and temporal localization of the disturbances, which can be realized visually from the time-frequency representation offered by the DST, unlike the DWT in that a simple visual inspection of its representation in the time-scale domain of the signal, is not always sufficient to detect or temporarily locate a disturbance.

Below are images of representations in the time-frequency domains and time-scale domains of three different perturbations including the normal signal.

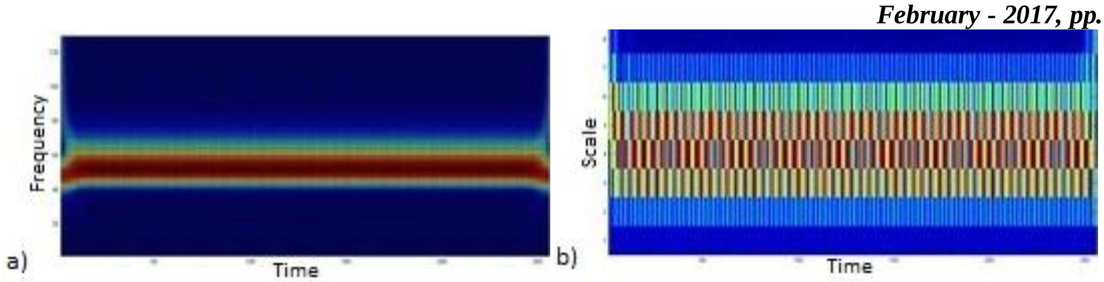

Fig. 2 Normal signal. a) Time-frequency representation using DST. b) Time-scale representation using DWT.

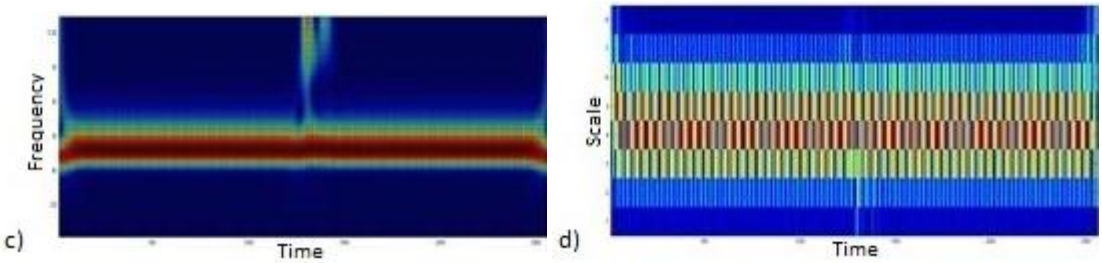

Fig. 3 Transient. a) Time-frequency representation using DST. b) Time-scale representation using DWT.

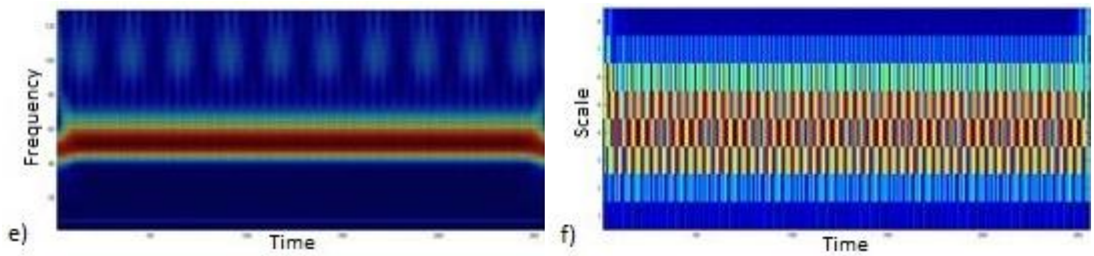

Fig. 4 Harmonics. a) Time-frequency representation using DST. b) Time-scale representation using DWT (b).

As can be seen in Figures 3 and 4 clause (a), it is possible to locate and recognize the type of disturbance (transient and harmonic) with a simple visual inspection of the time-frequency representation generated by the DST, whereas in the time-scale representations of the DWT is much more difficult to achieve this visual identification as can be seen in Figures 3 and 4 clause (b) and in other cases it is not possible to do so [18].

VI. CONCLUSIONS

In this paper, a comparative study of two methods used in the analysis of electrical disturbances was carried out. These methods are based on the use of mathematical transforms such as DWT and DST, in conjunction with SVM. As can be seen from the data obtained in Table II, the technique based in DST showed better results in the classification of electrical disturbances with 96.6% accuracy.

The DST also offered advantages in the detection and temporal location of the disturbances as shown in the time-frequency graphs of Figures 4 and 6, where the start and end moments of each disturbance could be clearly seen due to the changes of frequencies in the signal with respect to time.

REFERENCES

[1] Bertha M Soriano González, Frank Grau Merconchini, and Llohandry Cruz Carnero, "Estudio de fenómenos transitorios en circuitos eléctricos de primer orden," Ingeniería Energética, pp. 295-304, 2014.

[2] IEEE-Std-1159, "IEEE Recommended Practice for Monitoring Electric Power Quality," New York, 2009.

[3] David Jin and Lin Sally, Advances in Electronic Engineering, Communication and Management Vol. 1. Beijing: Springer, 2012.

[4] Weilin Li, Antonello Monti, and Ferdinanda Ponci, "Fault Detection and Classification in Medium Voltage DC Shipboard Power Systems With Wavelets and Artificial Neural Networks," vol. 63, no. 11, 2014.

[5] Sebastian Praveen and Antony Pramod, "A Neural Network Based Power Quality Signal Classification System using Wavelet Energy Distribution," 2015.

[6] Xiong Shicheng, Xia Ii, and Bu Leping, "An Effective S-transform Feature Extraction Method for Classification of Power Quality Disturbance Signals," 2015.

[7] Ameen M. Gargoom, Nesimi Ertugrul, and Wen L. Soong, "Automatic Classification and Characterization of Power Quality Events," vol. 23, no. 4, 2008.

[8] Jianmin Li, Zhaosheng Teng, Qiu Tang, and Junhao Song, "Detection and Classification of Power Quality Disturbances Using Double Resolution S-Transform and DAG-SVMs," vol. 65, no. 10, 2016.

[9] Gustavo Camps-Valls and Lorenzo Bruzzone, Kernel Methods for Remote Sensing Data Analysis. Reino Unido: WILEY, 2009.

[10] David Walnut, An Introduction To Wavelet Analysis. Boston: Birkhauser, 2002.
[11] R. Stockwell, L. Mansinha, and R. Lowe, "Localization of the Complex Spectrum: The S Transform," IEEE Transactions on Signal Processing, pp. 998-1001, 1996.
[12] Xiong Shicheng, Xia Li, and Bu Leping, "An effective S-transform feature extraction method for classification of power quality disturbance Signals," 2015.
[13] Shigeo Abe, Support Vector Machines for Pattern Classification. London: Springer, 2010.
[14] Johan A. K Suykens, Jony Van Gestel, Jos De Brabanter, and Joos Vandewalle, Least Squares Support Vector Machines. London: World Scientific Publishing, 2002.
[15] Chih-Chung and Lin, Chih-Jen Chang, "LIBSVM: A library for support vector machines," vol. 2, no. 3, 2011.
[16] David De Yong, Sudipto Bhowmik, and Fernando Magnago, "An effective Power Quality classifier using Wavelet Transform and Support Vector Machines," vol. 42, no. 15, 2015.
[17] Okelola Muniru Olajide, "Detection and Classification of Power Quality Event using Discrete Wavelet Transform and Support Vector Machine," vol. 4, no. 6, 2015.
[18] Mohammad Javad. Dehghani, "Comparison of S-transform and wavelet transform in power quality analysis.," vol. 3, no. 2, 2009.

