Advanced Signal Processing Techniques for Feature Extraction in Data Mining

Maya Nayak
Professor
Orissa Engineering College
BPUT, Bhubaneswar

Bhawani Sankar Panigrahi
Asst. Professor
Orissa Engineering College
BPUT, Bhubaneswar

ABSTRACT
This paper gives a description of various signal processing techniques that are in use for processing time series databases for extracting relevant features for pattern recognition. In addition to describe the normally used signal processing methods, we also present a novel signal processing technique, which is a modification of the well-known Short-Time Fourier Transform (STFT) and Wavelet Transform for extracting suitable features for both visual and automatic classification of non-stationary time-series databases [3][5]. Although the new signal processing technique is useful for speech, medical, financial, and other types of time varying databases, only the power network disturbance time series is used for computing features for visual and automatic recognition.

General Terms
Signal processing, Time series database, pattern recognition, data mining.

Keywords
Feature extraction, Short Time Fourier Transform (STFT), Wavelet Transform, S-transform, Transient, Swell, and Sag.

1. INTRODUCTION

Spectral analysis using the Fourier Transform is a powerful technique for stationary time series where the characteristics of the signal do not change with time. For non-stationary time series, the spectral content changes with time and hence time-averaged amplitude spectrum found by using Fourier Transform is inadequate to track the changes in the signal magnitude, frequency or phase. The signals with time varying characteristics that occur in electrical networks with power electronically controlled static devices, geophysical systems such as earth tremors, speech and acoustic systems, cardio vascular systems, financial and currency markets are typical cases for consideration of the advanced signal processing techniques, such as Short Time Fourier Transform and more recently Wavelet Transforms for mining time-series databases.Advanced signal processing techniques such as Wavelet transform and interesting modification of wavelet transform approach known as S-Transform are used for extracting pertinent feature vectors from sampled time series data either obtained by simulation of field test. The S-Transform on the other hand produces a time-frequency representation of a time series. It uniquely combines a frequency dependent resolution that simultaneously localizes the real and imaginary spectra. The basic functions for the S-Transform are Gaussian modulated cosinusoids, so that it is possible to use intuitive notions of cosinusoidal frequencies in interpreting and exploiting the resulting time-frequency spectrum. With the advantage of fast loss-less invertibility from time to time-frequency; and back to the time domain, the usage of the S-transform is very analogous to the Fourier Transform. In the case of non-stationary disturbances with noisy data, the S-Transform provides patterns that closely resemble the disturbance type and thus requires simple classification procedure. Further, the S-transform can be derived from the continuous wavelet transform (CWT) choosing a specific mother wavelet and multiplying a phase correction factor. Thus the S-Transform can be interpreted as a phase corrected continuous wavelet transform. The S-Transform approach provides a variable window, inversely proportional to frequency, thus facilitating the capture of both low and high frequency transient's signals of the electricity supply network subjected to disturbances.

2. PHASE CORRECTED WAVELET: S-TRANSFORM

The S-Transform of a time series $y(t)$ is defined as:

$$S(\tau, f) = \int_{-\infty}^{\infty} y(t) g(t-\tau, f) e^{-i2\pi ft} \qquad (2.1)$$

where the Gaussian modulation function $g(\tau, f)$ is given by

$$g(\tau, f) = \frac{|f|}{\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}} \qquad (2.2)$$

with the spread parameter inversely proportional to the frequency

$$\sigma = \frac{1}{|f|} \qquad (2.3)$$

the final expression becomes

$$S(\tau, f) = \int_{-\infty}^{\infty} y(t) \frac{|f|}{\sqrt{2\pi}} e^{-\frac{(\tau-t)^2 f^2}{2}} e^{-i2\pi ft} dt \qquad (2.4)$$

where f is the frequency, t and τ, are both time.

The S-Transform of the continuous time function $y(t)$ can also be defined as a CWT with a specific mother wavelet $W(\tau, d)$ multiplied by a phase factor.

$$S(\tau, f) = e^{i2\pi f\tau} W(\tau, d) \quad (2.5)$$

where the mother wavelet is defined as

$$W(\tau, d) = \int_{-\infty}^{\infty} y(t) w(t-\tau, d) dt \quad (2.6)$$

and $$\psi(t, f) = \frac{|f|}{\sqrt{2\pi}} e^{-\frac{t^2 f^2}{2}} e^{-i2\pi ft} \quad (2.7)$$

where the dilation factor is the inverse of the frequency *f*.The phase factor in equation (2.19) is in fact a phase correction of the definition of the Wavelet Transform [9][10]. It eliminates the concept of "wavelet analysis" by separating the mother wavelet into two parts, the slowly varying envelope (the Gaussian function) that localizes in time, and the oscillatory exponential kernel that selects the frequency being localized. It is the time localizing Gaussian that is translated while the oscillatory exponential kernel remains stationary. By not translating the oscillatory exponential kernel, The S-Transform localizes the real and imaginary components of the spectrum[11] independently, localizing the phase spectrum as well as the amplitude spectrum. This is referred to as absolutely referenced phase information.The choice of windowing function is not limited to the Gaussian function. Other windowing function was tried with success.

The inverse S-Transform is given by

$$y(t) = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} S(\tau, f) d\tau \right] e^{i2\pi ft} df \quad (2.8)$$

and since $S(\tau, f)$ is complex, it can be written as

$$S(\tau, f) = A(\tau, f) e^{i\theta(\tau, f)} \quad (2.9)$$

where $A(\tau, f)$ is the amplitude of the S-spectrum and $\theta(\tau, f)$ is the phase of the S-spectrum. The phase spectrum is an improvement on the wavelet transform in that the average of all the local spectra does indeed give the same result as the Fourier Transform.Equation (2.4) is derived on the assumption that the spread of the Gaussian modulation function is proportional to the inverse of frequency. To increase the frequency resolution, the spread of the Gaussian window $\sigma$is written as

$$\sigma = \frac{\alpha}{|f|} \quad (2.10)$$

and the generalized S-Transform is obtained as

$$S(\tau, f) = \int_{-\infty}^{\infty} y(t) \frac{|f|}{\alpha\sqrt{2\pi}} e^{\frac{(\tau-t)^2 f^2}{2\alpha^2}} e^{-i2\pi ft} dt \quad (2.11)$$

In equation (2.10) α controls the frequency resolution. If α is above 1, the frequency resolution would increase. Likewise if α is below 1, the time resolution improves. The S-Transform is a

linear operation on the signal $y(t)$. If additive noise is added to the signal $y(t)$,it can be modeled as $y_{noisy}(t) = y(t) + \eta(t)$. The operation of the S-Transform leads to

$$S\{y_{noisy}(t)\} = S\{y(t)\} + S\{\eta(t)\} \quad (2.12)$$

The amplitude and phase spectrum of S-transform are given by

$$A = abs(S(n, j)) \quad (2.13)$$
$$\phi(n, j) = ang(\mathrm{Im}(|S(n, j)|) / \mathrm{Re}(|S(n, j)|))$$

The discrete version S-Transform of a signal is obtained as

$$S[j, n] = \sum_{m=0}^{N-1} Y[m+n] W[m, n] e^{i\frac{2\pi mj}{N}} \quad (2.14)$$

where *Y[m + n]* is obtained by shifting the discrete Fourier Transform (DFT) of *y(k)* by *n*, *Y[m]* being given

$$Y[m] = \frac{1}{N} \sum_{k=0}^{N-1} y[k] e^{-j\frac{2\pi mk}{N}} \quad (2.15)$$

and $$W[m+n] = e^{\frac{-2\pi m^2 \beta^2}{n}} \quad (2.16)$$

j, *m* and *n* = 0, 1, ….., N-1. Another version of the discrete S-Transform used for computation

$$S[m, n] = e^{-j\frac{2\pi nm}{N}} \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} x[m+k] w^*(k, n) e^{i\frac{2\pi nk}{N}} \quad (2.17)$$

for *m = 1, 2, …, M* and *n = 0, 1, 2, …N/2* (2.18)

where *M* is the number of data points of the signal *y[m]*, *N* is the width of the window, the signal vector *y[m]* is padded at the beginning or at the end with 0.

3. FOURIER TRANSFORM

The Fourier Transform is one of the oldest and most powerful tools in signal processing. This transform maps the signal in time domain to a frequency domain where certain useful features about the signal can be seen. The Fourier Transform of a signal $y(t)$ is given as:

$$Y(f) = \int_{-\infty}^{\infty} y(t) e^{-i2\pi ft} dt \quad (3.1)$$

and its inverse relationship is:

$$y(t) = \int_{-\infty}^{\infty} Y(f) e^{i2\pi ft} df \quad (3.2)$$A discrete version

of the Fourier Transform called the Discrete Fourier Transform (DFT) of a time series of length *N* is defined as:

$$Y\left[\frac{n}{NT}\right] = \frac{1}{N} \sum_{k=0}^{N-1} y[kT] e^{-\frac{i2\pi nkT}{NT}} \quad (3.3)$$

where *T* is the sampling interval.

The inverse relationship is:

$$y[kT] = \sum_{n=0}^{N-1} Y\left[\frac{n}{NT}\right] e^{\frac{i2\pi nkT}{NT}} \quad (3.4)$$

The time information is obscured in the phase information of the Fourier Transform. The Fourier Transform correctly detects the two frequency components. The Fourier Transform is thus not suitable for analyzing non-stationary signals.

4. SHORT TIME FOURIER TRANSFORM

The Short Time Fourier Transform (STFT) is an attempt to improve on the traditional Fourier Transform. The STFT is the result of repeatedly multiplying the signal with shifted short time windows and performing a Fourier Transform on it. The STFT for a signal $y(t)$ is defined as:

$$STFT(\tau, f) = \int_{-\infty}^{\infty} y(t)w(t-\tau)e^{-i2\pi ft}dt \qquad (4.1)$$

where *w(t)* is the windowing function. The purpose of the windowing function is to dissect the signal to smaller segments where the segments of the signal are assumed to be stationary. The resulting spectrum of this window and the signal is the "local spectrum". This localizing window is then translated along the time axis to produce local spectra for the entire range of time.Due to the constant window used in the STFT, all parts of the signal are analyzed with the same resolution. This constant windowing is a drawback of the STFT because at high frequency, good time resolution is needed while good frequency resolution is needed at low frequency. The frequency resolution is proportional to the bandwidth of the windowing function while time resolution is proportional to the length of the windowing function. Consequently a short window is needed for good time resolution while a long window is needed for good frequency resolution. This limitation is due to the Heisenberg-Gabor inequality that states that

$\Delta t * \Delta f \geq c_w$ (4.2) where Δt is the time resolution, Δf is the frequency resolution and C_w is a constant that is dependent on the type of windowing function used.Consequently in most applications, 2 STFT must be used.

5. WAVELET TRANSFORM AND MULTI-RESOLUTION ANALYSIS

The Continuous Wavelet Transform (CWT) associated with the mother wavelet $\psi(t)$ is defined as:

$$W(a,b) = \int_{-\infty}^{\infty} y(t)\psi_{a,b}^{*}(t)dt \qquad (5.1)$$

where $y(t)$ is any square integrable function, *a* is the scaling parameter, *b* is the translation parameter and $\psi_{a,b}(t)$ is the dilation and translation of the mother wavelet defined as:

$$\psi_{a,b}(t) \equiv \frac{1}{\sqrt{|a|}}\psi\left(\frac{t-b}{a}\right) \qquad (5.2)$$

This CWT [9][10] provides a redundant representation of the signal in the sense that the entire support of $W(a,b)$ need not be used to recover $y(t)$. By only evaluating the CWT at dyadic intervals, the signal can be represented compactly as:

$$y(t) = \sum_{k=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} d_j(k)2^{j/2}\psi\left(2^j t - k\right) \qquad (5.3)$$

where $d_j(k)$ is called the discrete wavelet transform (DWT) of $y(t)$ associated with the wavelet is a scaling function $\varphi(t)$. The scaling function along with the wavelet creates a multiresolution analysis (MRA) of the signal. The scaling function of one level can be represented as a sum of a scaling function of the next finer level.

$$\varphi(t) = \sum_{n=-\infty}^{\infty} h(n)\sqrt{2}\varphi(2t-n) \qquad (5.4)$$

The wavelet is also related to the scaling function by

$$\psi(t) = \sum_{n=-\infty}^{\infty} h_1(n)\sqrt{2}\varphi\,(2t-n) \qquad (5.5)$$

scaling function used to represent the signal as

$$y(t) = \sum_{k=-\infty}^{\infty} c_{jo}(k)2^{jo/2}\varphi(2^{jo}t-k) + \sum_{k=-\infty}^{\infty}\sum_{j=jo}^{\infty} d_j(k)2^{j/2}\psi(2^j t-k) \qquad (5.6)$$

where *jo* represents the coarsest scale spanned by the scaling function.The scaling and wavelet coefficients of the signal $y(t)$ can be evaluated by using a filter bank of quadrature mirror filters (QMF).

$$c_j(k) = \sum_{m=-\infty}^{\infty} c_{j+1}(m)h(m-2k) \qquad (5.7)$$

$$d_j(k) = \sum_{m=-\infty}^{\infty} c_{j+1}(m)h_1(m-2k) \qquad (5.8)$$

Equations (5.7) and (5.8) show that the coefficients at a coarser level can be attained by passing the coefficients at the finer level to their respective filters followed by a decimation of two. The decomposition process is shown in Fig.5.1. For a signal that is sampled at a frequency higher than the Nyquist frequency, the samples are used as $c_{j+1}(m)$. A three level decomposition of time series data obtained from power signal disturbances, sampled at 12.8 kHz is shown below. The approximation coefficients contain the low frequency information while the detail coefficients contain the high frequency information of the oscillatory transient.

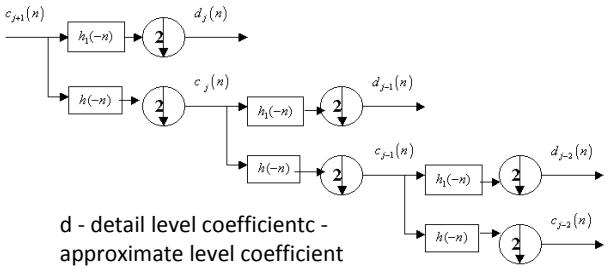

Fig. 5.1. Three level wavelet decomposition

6. A QUALITATIVE COMPARISON BETWEEN DWT AND S-TRANSFORM FOR ANALYSIS OF TIME SERIES EVENTS

The discrete wavelet transform is one of the most popular tools used for power signal time series disturbance classification today due to its multi-resolution capabilities and fast calculation of its coefficients. The feature vectors used for classifying the power signal time series events usually involve performing some kind of transformation on the DWT coefficients, comparing between the DWT coefficients of the disturbance signal with the DWT coefficients of a pure signal, compressed DWT coefficients,and the direct use of the DWT coefficients.The following figures show the 2-D plot of the S-Transform with α=0.2, and the 4-level decomposition of the DWT with db4 wavelet for some of the time series databases like the oscillatory transient, notch, swell, sag, interruption, harmonics, swell with harmonics and sag with harmonics, etc.

6.1 Oscillatory transient

In Fig.6.1.1 an oscillatory transient is clearly detected and localized by the S-Transform. This localization ability of the S-Transform could be used to quantify the duration of the disturbance. By observing the concentric contours, we see that the transient is undergoing damping. While the transient event is detected for all frequency values, it is possible to determine its frequency by noting the location in frequency of the highest contour. Fig. 6.1.2 shows the 4 level decomposition of the same transient with the db4 wavelet. The transient is detected and localized very well by coefficients at all levels. This can determine the duration of the transient disturbances. Unlike the S-Transform, we are not able to pin point the frequency of the transient by looking at the DWT coefficients. This is because each level of the decomposition corresponds to a range of frequencies. For example, in this case with the sampling frequency set at 12.8 kHz, the first level coefficients are obtained by passing the samples to a band pass filter whose frequency band is between 6.4 kHz-12.8 kHz. Level 2 coefficients contain information of frequency band 3.2 kHz-6.4 kHz, level 3 for 1.6 kHz-3.2 kHz and level 4 for 800Hz-1.6 kHz. Observing Fig.6.1.2 we can guess that the transient frequency is in level 3 or level 4 by looking at the magnitude of the detail coefficients at this level. The transient frequency is actually 900Hz.

6.2 Impulsive Transient

The S-Transform correctly detects and localized the impulsive transient as shown in Fig. 6.2.1 The impulsive transient is represented as a long line in the time-frequency plane that last for a wide range of frequencies. This is because a very short impulse in the time domain would result in a Fourier Transform that spreads out for a wide range of frequencies. In Fig.6.3.1, the DWT coefficients at all levels detect and localizes the impulsive transient disturbance.

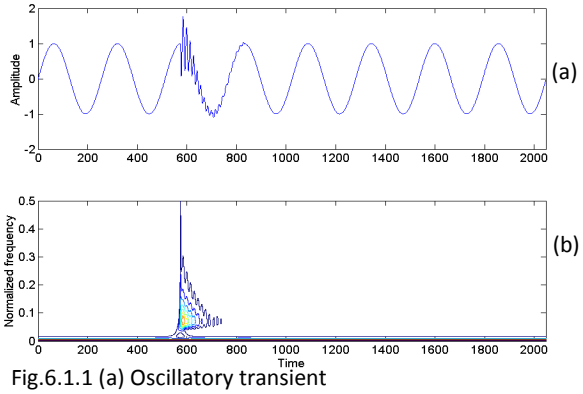

Fig.6.1.1 (a) Oscillatory transient
(b) S-Transform output of oscillatory transient

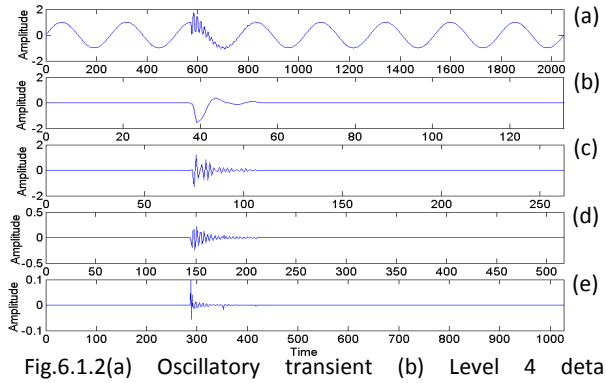

Fig.6.1.2(a) Oscillatory transient (b) Level 4 detail coefficients (c) Level 3 detail coefficients (d) Level 2 detail coefficients (e) Level 1 detail coefficients

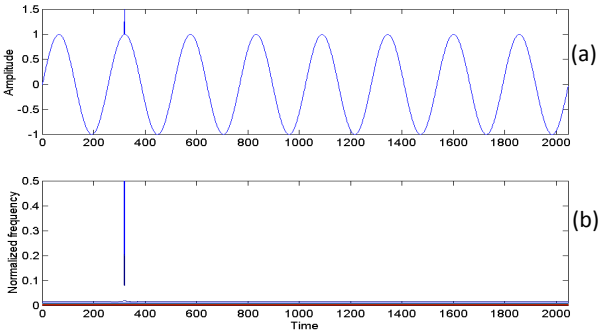

Fig. 6.2.1 (a) Impulsive transient (b) S-Transform output of impulsive transient

6.3 Multiple notches

The S-Transform produces a unique representation of multiple notches. In Fig.6.3.2 the notches are represented as short lines in the time-frequency plane. The length of the lines is much shorter than that for the impulsive transient. Only the deeper notches are detected. This flaw may be remedied by using a smaller value of the factor$\propto$. Fig. 6.3.3 shows that the notches are accurately detected by DWT coefficients at level 1 and 2.

The DWT coefficients look like random noise at higher levels

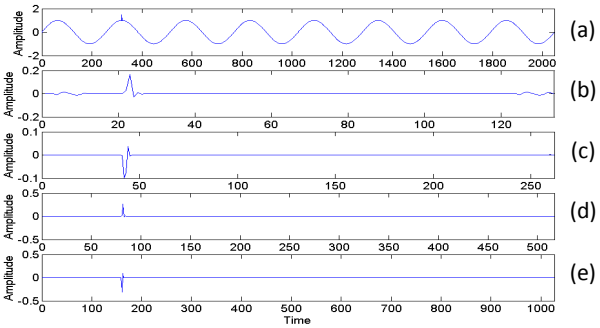

Fig.6.3.1 (a) Impulsive transient (b) Level 4 detail coefficients (c) Level 3 detail coefficients (d) Level 2 detail coefficients (e) Level 1 detail coefficients

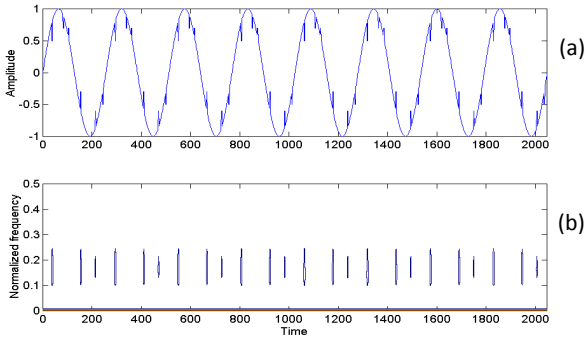

Fig.6.3.2(a) Multiple notches (b) S-Transform output of multiple notches.

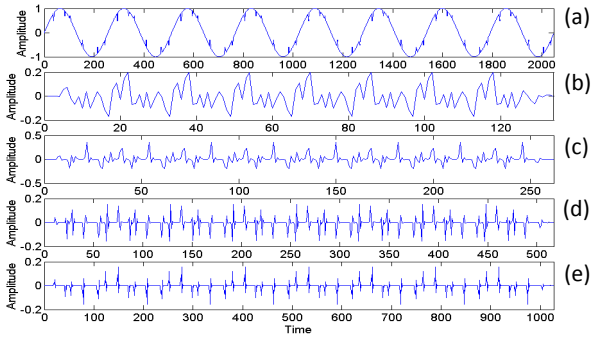

Fig.6.3.3 (a) Multiple notches (b) Level 4 detail coefficients (c) Level 3 detail coefficients (d) Level 2 detail coefficients (e) Level 1 detail coefficients

6.4 Swell

Swell is uniquely identified by the S-Transform in three distinct ways. Looking at the concentric contours in Fig.6.4.1, we see that the highest valued contour is located at the region of the swell. Furthermore if we look at the shape of the lowest valued contour, we see that its frequency values increases at the time of the swell and decreases when the signal is back to its normal amplitude again.. Another easy way to see is to note that the number of contours in the swell region is larger than in the normal region. The DWT coefficients (Fig.6.4.2) detect the time of the rise and fall of the swell disturbance.

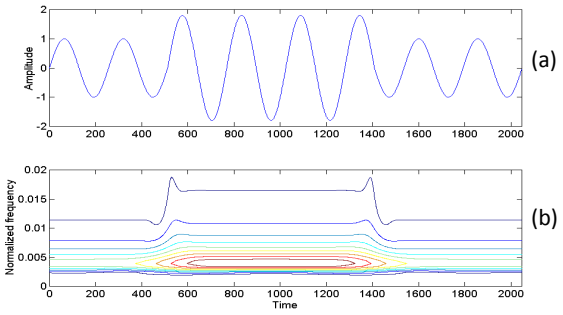

Fig. 6.4.1 (a) Swell (b) S-Transform output of voltage

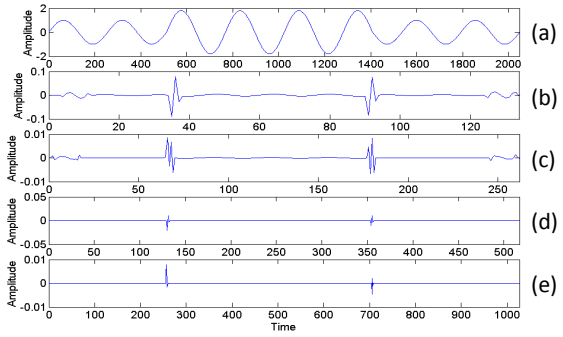

Fig.6.4.2 (a) Swell (b) Level 4 detail coefficients (c) Level 3 detail coefficients (d) Level 2 detail coefficients (e) Level 1 detail coefficients

6.5 Sag

Fig.6.5.1 shows the S-Transform output of sag. The highest contours are located at the beginning and end of the time-frequency plane. The density of the contour decreases in the middle. This shows that there is a decrease in magnitude in the middle of the time axis. The lowest contour maintains a frequency value at roughly 0.017 until the occurrence of the sag in which it rises in frequency for a short time then decreases and settles at0.01 until the end of the sag. At the end of the sag its frequency values rises to a peak and then settles down to its original value of 0.017. The two peaks can be used to quantify the length of the sag. The DWT in Fig.6.5.2 on the other hand does not produce a pattern that can be used to classify the disturbance as sag. Looking at Fig.6.4.2 and Fig.6.5.2, we see that the DWT is unable to distinguish between swell and sag.

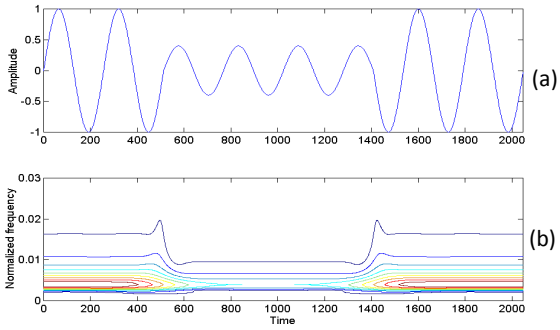

Fig. 6.5.1 (a) Voltage sag (b) S-transform output of voltage sag

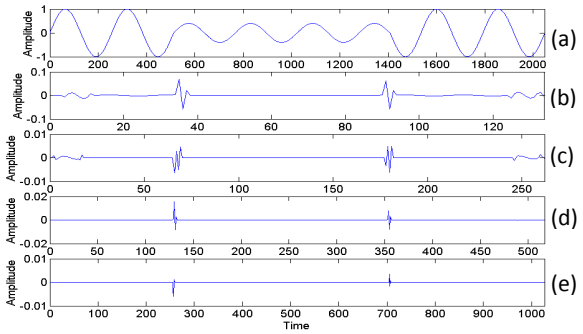

Fig.6.5.2 (a) Voltage sag (b) Level 4 detail coefficients (c) Level 3 detail coefficients (d) Level 2 detail coefficients (e) Level 1 detail coefficients

7. FEATURE SELECTION

7.1 Based on Wavelet Transform

Our approach to time series data mining is to first extract a set of features from each time series and then base further analysis on the feature vectors. There are many different wavelet based methods reported in the literature for feature extraction. They involve performing some kind of transformation on the DWT coefficients, comparing between the DWT coefficients of the disturbance signal with the DWT coefficients of a pure signal, compressed DWT coefficients, direct use of the DWT. In this chapter we will be comparing the performance of and the hybrid FFT-DWT methods for pattern classification of electrical network disturbances along with the new approach based on statistical feature selection by S-transform. These three methods were chosen on the basis that the feature vectors were of low dimensionality and were translation invariant. Each disturbance will give a different signature that can be used for classification purposes.

7.2 Based on S-Transform

The output from the S-Transform is an N by M matrix called the S-matrix whose rows pertain to frequency and whose columns pertain to time. Each element of the S-matrix is complex valued. The S-matrix can be represented in a time-frequency plane similar to that of the wavelet transform. Here α is normally set to 0.4 for best overall performance of the S-Transform, where the contours exhibit the least edge effects and for computing the highest frequency component of an oscillatory transient waveform α is set equal to 1 or 3 as required. The S-Transform performs multi-resolution on a time varying power network signal, because its window width varies inversely with frequency. This gives high time resolution at high frequency and high frequency resolution at low frequency. The detailed computational steps have been outlined in software originally developed by Stockwell [11] and modified by the author.

Feature extraction is done by applying standard statistical techniques to the contours of the S-matrix as well as directly on the S-matrix. These features have been found to be useful for detection, classification or quantification of relevant parameters of the signals. The power network signal is normalized with respect to a base value, which is the normal value without any disturbance.The resolution factor α was set to 0.4 for better time resolution. Ten features were extracted from the S-transform output. They are

1.F_1 = *max (A)+min (A)-max (B).*

where A is the amplitude versus time graph from the S-matrix under disturbance and B is the amplitude versus time graph of the S-matrix without disturbance.

2. F_2= Standard deviation (σ) of contour No.1 having the largest frequency

$$\text{Magnitude}= \sqrt{\frac{\sum_{i=1}^{N}(y_i-\bar{y})}{N}}$$

3. F_3= Energy

4. F_4 = Total harmonic distortion of the signal $= \frac{\sqrt{\sum_{n=2}^{N}|V_n|^2}}{|V_1|}$

where$N$ is the number of points in the FFT, V_n the value of the nth harmonic component of the FFT.

5. F_5 = Maximum Frequency of the signal

6. F_6 = Mean of the lowest contour above twice the normalized fundamental frequency.

7. F_7 = Skewness of Cr. $= \frac{\sum_{i=1}^{N}(y_i-\bar{y})^3}{(N-1)\sigma^3}$

8. $F_8 = \max(Cr)-\min(Cr)$

9. F_9 = Standard deviation of *Cr*.

where*Cr* is the amplitude versus time graph of the S-matrix for frequencies above twice the normalized fundamental frequency magnitude versus time.

10. F_{10}=Average power for frequencies above 2.5 times the fundamental frequency.

These features are found to be well suited to distinguish the twelve (12) classes of power quality events studied here.

8. FEATUREEXTRACTION USING SIMULATED AND RECORDED DATA

The proposed approach is applied to detect, localize and classify signal patterns in electrical power networks. The power network disturbances that produce time varying databases are broadly due to the following events:

(i) Sag
(ii) Swell
(iii) Interruption
(iv) Oscillatory transients.
(v) Spikes
(vi) Notches

To illustrate the relative values of these features, several typical power network time series data from various disturbance events like voltage sag, swell, interruption, harmonics, transients, spikes, and notches, etc. are simulated using the MATLAB code. A random white noise of zero mean and signal to noise ratio (SNR) varying from 50dB to 20dB is added to the signals. A typical SNR value of 30dB is equivalent to a peak noise magnitude of nearly 3.5% of the voltage signal. A sampling frequency of 1.6 kHz is chosen for computation of S-Transforms for the simulated waveforms. As the S-Transform software uses analytic signals, only positive frequency spectrum is evaluated and FFT of the signal samples is automatically computed as a part of the software. The software developed by author was an extension to the original version of the S-transform by Stockwell.Laboratory recorded data is also used for feature extraction and time series classification. The various time series disturbance signals belonging to the described classes namely voltage swell, sag, interruption, transient, spike, and notch, sag with harmonics, swell with harmonics etc. are shown in Figs.6.1.1 to 6.5.2 In these figures also the multi-resolution S-transform output contours along with the instantaneous amplitude of signal versus time graphs are shown. The S-transform contours clearly present the distinguishing characteristics of the non-stationary signal patterns in the data samples that are pertinent features of each individual disturbance class.

Table–1 shows the features of voltage sag, swell and interruption with both simulated and recorded data. The voltage sag and swell magnitude and duration are varied to show the change in feature magnitudes. These disturbances belong to steady state fundamental frequency category and the class C_0 to C_7. Table-II shows the features of the disturbances that include very fast changing signals such as oscillatory and impulsive transients, notches, etc. These disturbances are categorized into classes C_8 to C_{12}.

From Tables-1 and 2 different classes of power network disturbance signals exhibit different feature magnitudes, which can be used for an automatic recognition of their class. It is quite evident that the standard deviation F_2 is generally small and ≤ 0.05 for power frequency disturbances like voltage sag, swell, interruption, etc. ($F_1 = 0.2$ for 20% sag, $F_1 = 1.6$ for 60% swell etc). The feature F_1 is not very much influenced by the noise level up to 20dB (which is quite high). However, when the harmonic content in the signal is within 15%, the exact magnitudes of sag, swell, interruption, etc. are minimally influenced. With more than 20% harmonic content in the signal, an interruption may be misclassified as sag.

From Table-2, it is found that the transient disturbances have a higher value of the feature F_2 ($F_2 > 0.1$) and the feature F_1 lies in the range of .99 – 1.05. The feature F_2 is, however, an important one for distinguishing between oscillatory and impulsive transients and notches since F_1 is less than 0.9 for notches, however when notch is mixed with harmonics the feature value might higher than 0.9. Further the feature F_3 distinguishes between oscillatory transients and impulsive transients as $F_3 <$ .05 for oscillatory transients and >0.05 for spikes or impulsive transients. The factor α influences the size of the window and effects the variation of features F_1 and F_2 during steady and transient disturbances as shown in Table-3. It is also found from the table that as α increase from 0.4 to 1.2, the peak amplitude of the signal (which is 0.6 in this case) becomes inaccurate. The standard deviation also decreases considerably with increasing values of α.

Table 1. Features extracted from S-transform

Disturbances	F_1	F_2	F_3	F_4	F_5
Normal	1.002				
Sag (60%)	0.593	0.053	.031	.0312	50.0
Swell (50%)	1.50	0.0129	.076	.015	50
Momentary Interruption (MI) (5%)	0.0724	0.035	.019	.0350	50
Flicker (5 Hz, 4%)	0.987	.0168	.026	0.0186	55
Notch + harmonics	0.939	0.131	.0529	0.136	56.25
Spike + harmonics	1.065	0.141	.0627	0.1308	56.25
Transient (low frequency)	1.015	0.138	.0163	.01	705
Transient (high frequency)	1.043	0.149	.014	.043	2520

Table 2. Features extracted S-transform with SNR 20 dB

Disturbances	F_1	F_2	F_3	F_4	F_5
Normal	.9963	.001	.052	.028	50
Sag (60%)	.591	.022	.039	.027	50
Swell (50%)	1.503	.012	.076	.029	50
Momentary Interruption (MI)	.070	.0387	.0323	.044	50
Harmonics	1.032	.050	.064	0.25	50
Sag with Harmonic (60%)	0.601	.0228	.0408	0.1139	50
Swell with Harmonic (50%)	1.50	.0219	.079	.1155	50
Flicker (4%, 5 Hz)	.998	.0209	.027	0.1159	55
Notch + harmonics	.940	.1275	.0531	0.198	50
Spike + harmonics	1.072	.141	.066	.204	50
Transient (low frequency)	1.021	0.1473	.0148	.0566	440
Transient (high frequency)	1.0384	.155	.014	.068	3315

Table 3. Effects of Window parameter α

	Steady state variations		Transient variations	
Factor α	**F_1**	**F_2**	**F_1**	**F_2**
0.4	0.602	0.0427	0.9877	0.1654
0.6	0.610	0.0236	0.9898	0.1650
0.8	0.621	0.01735	0.9911	0.1647
1	0.633	0.0145	0.9923	0.1623
1.2	0.645	0.0126	0.9927	0.1600

The tables presented above provide a relative comparison of the relevant features for pattern recognition and classification of the time series events.

9. CONCLUSION

The paper presents several signal processing techniques along with some advanced ones like Wavelet and S-transforms for extracting the relevant features from time series databases that could provide visual and automatic classification of these time series databases. Time series databases occurring in power networks [13] are processed through these time-frequency transforms to provide valuable information of their class and other characteristics for data mining of these events and subsequent knowledge discovery.Overall, wavelet and S-transform based methods are showing promising results in the area of multiple pattern recognition and data mining of time series data from an electricity supply network. Further the approach of utilizing wavelets and soft computing methods [14] for pattern recognition of non-stationary time series data is a general one and can be applied to medical, financial, and other types of temporal data. In a large data set partitioning the data into several sections and applying the above approach the similarity of patterns present in the data set can be determined. The new approach presented in this paper outperforms the other existing techniques as far as accuracy, sensitivity to noise are concerned.

10. REFERENCES

[1] S. Santoso, E.J. Powers, W.M. Grady and A.C. Parsons, "Power quality disturbance waveform recognition using wavelet-based neural classifier. I. Theoretical foundation," IEEE Trans. on Power Delivery, vol.15, pp.222-228, 2000.

[2] D. Borras et al, "Wavelet and neural structure: A new tool for diagnostic of power system disturbances," IEEE Trans. on Industry Applications, vol.37, no.1, pp.184-190, 2001.

[3] C.H. Lee, J.S. Lee, J.O. Kim, and S.W. Nam, "Feature Vector Extraction for the Automatic Classification of Power Quality Disturbances", 1997 IEEE International Symposium on Circuits and Systems, pp.2681-2685, 1997.

[4] B. Perunicic, M. Malini, Z. Wang, Y. Liu, "Power Quality Disturbance Detection and Classification Using Wavelets and Artificial Neural Networks," Proc. International Conference on Harmonics and Quality of Power, pp.77-82, 1998.

[5] A.M. Gaouda, M.M.A. Salama, S.H. Kanoun, and A.Y. Chikhani, "Wavelet-Based Intelligent System for Monitoring Non-Stationary Disturbances," International Conference on Electric Utility Deregulation and Restructuring and Power Technologies, pp.84-89, 2000.

[6] A.M. Gaouda, M.M.A. Salama, M.R. Sultan, and A.Y. Chikhani, "Power quality detection and classification using wavelet multiresolution signal decomposition," IEEE Trans. on Power Delivery, vol.11, no.4, pp.1469-1476, 1996.

[7] A.M. Gaouda, S.H. Kanoun, M.M.A. Salama and A.Y. Chikhani, "Wavelet-based signal processing for disturbance classification and measurement," IEE Proc-Gener. Transm. Distrib, vol.149, no.3, May 2002.

[8] A.M. Gaouda, M.M.A. Salama, S.H. Kanoun, and A.Y. Chikhani, "Pattern Recognition Applications for Power System Disturbance Classification", IEEE Trans. on Power Delivery, vol.17, no.3, pp.677-682, 2002.

[9] S.C. Burrus, R.A. Gopinath, and H.Guo, "Introduction to wavelets and wavelet transform," Prentice Hall, New Jersey, 1997. (Book)

[10] M.Rao. Raghuveer and A.S. Bopardikar, "Wavelet Transforms Introduction to Theory and Applications," Addison-Wesley, Massachusetts, 1998. (Book)

[11] R.G. Stockwell, L. Mansinha, and R.P Lowe, "Localization of the Complex Spectrum: The S-Transform," IEEE Trans. Sig. Proc., vol.44, no.4, pp.998-1001, 1996.

[12] R.G. Stockwell, "S-Transform Analysis of Gravity Wave Activity from a Small Scale Network of Airglow Imagers," Ph.D. Thesis, University of Western Ontario, 1999.

[13] S. Santoso, E.J. Powers, W.M. Grady, and A.C. Parsons, "Power quality disturbance waveform recognition using wavelet-based neural classifier. II.Application," IEEE Trans. Power Delivery, vol.15, pp.229-235, Jan.2000.

[14] Yean Yin, Liang Zhang, Miao Jin and SunyiXie"Wavelet Packet and Generalized Gaussian Density Based Textile Pattern Classification Using BP Neural Network ", Lecture Notes in Computer Science, Volume 6146/2010, 540-547, DOI: 10.1007/978-3-642-13498-2_70, 2010.

[15] Michael Häfner, Roland Kwitt,Andreas Uhl,Friedrich Wrba , Alfred Gangl ,Andreas Vécsei, "Computer-assisted pit-pattern classification in different wavelet domains for supporting dignity assessment of colonic polyps", Journal Pattern Recognition archive Volume 42 Issue 6, Elsevier Science Inc. New York, NY, USA doi10.1016/j.patcog.2008.07.012, June 2009.