Chapter-1

Introduction

Semiconductor devices used in various applications have some harmful effects, which are to be eliminated by power quality (PQ) engineers. Elimination of these unwanted PQ events uses classification methods. Different digital signal processing methods have been used to process the PQ signals and then various features are extracted to fed the classifiers to classify the PQ events.

Various transforms has been used to obtain the harmonic information of the signal. Among these are Fourier transform, short time Fourier transform (STFT) and wavelet transform. Here we used wavelet transform to obtain the analytic signal for instantaneous feature extraction of PQ disturbances and constructing the feature vector which are fed to K-NN classifier [1]. We also compared the classification accuracy by constructing the feature vector using analytic signal by Hilbert transform.

Hilbert transform gives phase shift of 90 degree of given signal which is combined to original signal to give analytic signal. Let s(t) is original signal and g(t) is Hilbert transform of s(t) then analytic signal defined as z(t) = s(t) + jg(t). Using this analytic function we estimate instantaneous amplitude, phase and frequency. They are

$$|z(t)| = \sqrt{|s(t)|^2 + |g(t)|\text{^}2}$$

$$\theta(t) = \tan^{-1}\left(\frac{g(t)}{s(t)}\right)$$

$$f(t) = \frac{1}{2\pi}\frac{d\,\theta(t)}{dt}$$

Using amplitude and phase we have estimated energy, standard deviation of amplitude and standard deviation of phase after passing the classifiers and compared the result which are obtained using wavelet transform.

Here we are generating 135 signals for each disturbance like sag, swell, harmonic and transient. Out of these, 100 cases are taken for testing LDA+K-NN classifier and 35 cases are taken for training.

1.1 Use of Power Quality

Power quality is an important issue for the producers, distributors and consumers. The common problems, like harmonics, short term voltage variations (sags, swells and interruptions), [2] long term voltage variations (under voltages, over voltages and interruptions), transients, unbalance, frequency variations and others, can cause several problems to the consumers which require high levels of power quality for their industrial processes or home use.

Power quality studies are the necessary so that measures can be taken to solve the problems. Several international standards related to power quality and electrical system monitoring, like IEEE 519, IEC 61000 and EN 50160, help to classify and offer possible solutions to the problems described above, but specialized equipment is necessary in order to find which problems affect a given facility [3].

Since many of the commercially available equipment's are either too expensive, or have too many limitations, it was decided to develop a new low-cost power quality monitor that could be an alternative to the equipment in the market.

1.2 Literature Survey

There is a need to analyze power-quality (PQ) signals and to extract their different features to take preventative actions in power systems. [2] Classify PQ signals using Hilbert and Clarke Transforms as new feature extraction techniques. Both techniques accommodate Nearest Neighbor Technique for automatic recognition of PQ events. The Hilbert transform is introduced as single-phase monitoring technique, while with the Clarke Transformation all the three-phases can be monitored simultaneously. [4] The performance of each technique is compared with the most recent techniques (S-Transform and Wavelet Transform) using an extensive number of simulated PQ events that are divided into four classes.

[5] Used generalized S-transform with a variable window as a function of PQ signal frequency to generate contours and feature vectors for pattern classifications. The new classifier provides a complete characterization of both steady state and transient PQ signals

using fuzzy logic-based decision systems the fuzzy PQ classifier uses 14 rules based on trapezoidal membership functions for most of the PQ disturbance events with an accuracy rate close to 98% average.

[6] Presents a brief description of the S-transform and Wavelet transform together with two novel PQ monitoring approaches It was concluded that the K-NN based techniques can offer an effective automatic classification of PQ events. To identify the best number of neighbors in the classification of the events, new feature extraction techniques were examined. The K-NN-DHT method was developed for the single phase analysis of the power systems, and the K-NN-CT was offered for three phase systems. The performance of these two techniques was also compared with the Wavelet and ST techniques, using K-Nearest Neighbor pattern recognition technique. To ensure a higher accuracy in each classifier, the number of neighbours was varied. This provided a pattern based on increasing or decreasing profile of the average error. The results were given between 1 to 50 neighbours. The above 50 neighbors, the profile of the average error followed the same pattern.

[7] Presents S-Transform based probabilistic neural network (PNN) classifier for recognition of power quality (PQ) disturbances. This method requires less number of features as compared to wavelet based approach for the identification of PQ events. The features extracted through the S-Transform are trained by a PNN for automatic classification of the PQ events. This method can reduce the features of the disturbance signal to a large extent without losing its original property, less memory space and learning PNN time are required for classification. Four types of disturbances are considered for the classification problem. The simulation results show that the combination of S-Transform and PNN can effectively detect and classify different PQ events. The classification performance of PNN is compared with a feed forward multilayer (FFML) neural network (NN) and learning vector quantization (LVQ) NN. The classification performance of PNN is better than both FFML and LVQ. These features correctly classify the PQ disturbances, even under noisy conditions, as these features are based on magnitude, frequency, and phase of the disturbance signal. This has advantage over wavelet analysis in which some of the features are noise prone (like the energy at dilation levels 1 and 2) and a de-noising scheme is to be implemented to extract those features accurately. PNN correctly classifies the PQ class with high accuracy. The PNN is compared with both LVQ and FFML and it is found that PNN gives the best result.

1.3 Thesis Organization

Organization of thesis is as follows. Chapter-2 presents brief introduction of wavelet transform and its properties. Chapter-3 presents theory of analytic signal using Hilbert transform and wavelet transform. Chapter-4 presents Empirical mode decomposition and intrinsic mode function in brief. Chapter-5 presents classification, types of classifier, linear discriminant analysis and K-NN classifier.Chapter-6 presents implementation and results. Chapter-7 presents conclusion and future work.

Chapter-2

Wavelet Transforms

2.1 Continuous-Time Wavelets

Consider a real or complex-value continuous time function ψ(t) with the following properties

1. The function integrates to zero:

$$\int_{-\infty}^{\infty} \psi(\mathrm{t})\mathrm{dt} = 0 \tag{2.1.1}$$

2. It is square integrable or, equivalently, has a finite energy

$$\int_{-\infty}^{\infty} |\psi(\mathrm{t})|^2 dt < \infty$$

Where the function ψ(t) is called mother wavelet or it is also called wavelet if it is satisfies above two conditions. The first property states that the function is oscillatory or that has wavy appearance. Thus, it is a "small wave" or a wavelet. The two conditions are easily satisfied and there is an infinitive of functions that qualifies as mother wavelet [8].

2.2 Definition of the Continuous Wavelet Transform

Assume x(t) be any square integrable function. The Continuous wavelet transform or continuous-time wavelet transform of x(t) with respect to a wavelet ψ(t) is defined as

$$\mathrm{W(a,b)} = \int_{-\infty}^{\infty} x(t) \frac{1}{\sqrt{|a|}} \psi^* \left(\frac{t-b}{a}\right) dt \tag{2.2.1}$$

Where a is called scale parameter and b is called translation parameter and both are real, the function $\psi *$ denotes complex conjugation of function ψ**.** Here both function x(t) and ψ^* belongs to set of square integrable functions. It is also called set of energy signals.

We can also write

$$\mathrm{W(a,b)} = \int_{-\infty}^{\infty} x(t) \psi_{a,b}^* (t) dt \tag{2.2.2}$$

Where

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right) \tag{2.2.3}$$

If we take value of a=1 and, b=0 then

$$\psi_{1,0}(t) = \psi(\mathrm{t}) \tag{2.2.4}$$

The normalised factor $\frac{1}{\sqrt{|a|}}$ ensures that energy stays the same for all a and b that is

$$\int_{-\infty}^{\infty} |\psi_{a,b}(t)|^2 dt = \int_{-\infty}^{\infty} |\psi(\mathrm{t})|^2 dt \tag{2.2.5}$$

for all a,b. For any given value of a, the function $\psi_{a,b}(t)$ is shift of $\psi_{a,0}(t)$ by amount of b, along the time axis. Thus the variable b represents time shift or the translation. From

$$\psi_{a,0}(t) = \frac{1}{\sqrt{|a|}}\psi(\frac{t}{a}) \quad (2.2.6)$$

$\psi_{a,0}(t)$ is a time-scaled and amplitude-scale version of $\psi(t)$.

2.3 Properties of Continuous Wavelet Transform

Wavelet transform of a signal contains two variables. They are dilation and translation parameter. On the basis of these parameter continuous wavelet transform has following properties

1 Scaling

2 Shifting

3 Time Reversal

1 Scaling

Since a determines the amount of time scaling or dilation, it is referred to as the scale or dilation variable.

$$\psi_{a,0}(t) = \frac{1}{\sqrt{|a|}}\psi(\frac{t}{a}) \quad (2.3.1)$$

In the above equation if a > 1 then there is starching of function along the time axis, where as if 0 < a < 1, there is contraction of function.

2 Shifting

The wavelet transform is defined as

$$\text{W(a,b)} = \int_{-\infty}^{\infty} x(t)\frac{1}{\sqrt{|a|}}\psi^*(\frac{t-b}{a})dt \quad (2.3.2)$$

For different value of b we can get different wavelet coefficients that is shifting of wavelet

3 Time Reversal

The negative value of dilation parameter gives time reversal in combination with dilation. The get time reversal for negative value of b also.

2.4 The CWT as a Correlation

The CWT has application in area such as pattern detection and classification, the set of square integrable function form a linear vector space under addition and vector multiplication. This vector space comes with the well-defined inner product. If two signals x(t) and y(t) have finite energy , then their inner product are defined as

$$\langle x(t), y(t)\rangle = \int_{-\infty}^{\infty} x(t)y^*(t)dt \quad (2.4.1)$$

The total energy of y(t) is given as $\int_{-\infty}^{\infty} |y(t)|^2 dt$ which is $\langle y(t), y^*(t)\rangle$. The square root of energy is called the norm of y(t) and is denoted by ||y(t)||

CWT can be represented as inner products of signal x(t) and translated and dilated wavelet $\psi_{a,b}(t)$ for all a and b.

$$W(a,b) = \langle x(t), \psi_{a,b}(t)\rangle \tag{2.4.2}$$

The continuous wavelet transform has a cross-correlation representation also. The cross-correlation of two signal x(t) and y(t) are defined as

$$R_{x,y}(\tau) = \int x(t) y^*(t-\tau) = \langle x(t), y(t-\tau)\rangle \tag{2.4.3}$$

Where τ is lag or shift parameter, similarly we can write the wavelet transform as cross-correlation

$$W(a,b) = \langle x(t), \psi_{a,0}(t-b)\rangle = R_{f,\psi_{a,0}}(b) \tag{2.4.4}$$

In other word the continuous wavelet transform is the cross correlation at lag b between x(t) and wavelet dilated to scale factor a.

2.5 Inverse Continuous Wavelet Transforms

The inverse Fourier transform of wavelet W(a,b) defined as

$$X(t) = \frac{1}{C}\int_{a=-\infty}^{\infty}\int_{b=-\infty}^{\infty}\frac{1}{|a|^2} W(a,b)\, \psi_{a,b}(t) da\, db \tag{2.5.1}$$

Where C is the constant from the admissibility condition and it is defined as

$$C = \int_{-\infty}^{\infty}\frac{|\psi(\omega)|^2}{|\omega|} d\omega \tag{2.5.2}$$

And is such that $0 < C < \infty$

In above equation the term $\frac{W(a,b)}{|a|^2}$ provide the weighting function for constructing x(t) $\psi_{a,b}(t)$ the translated and dilated of mother wavelet.

<u>There are 5 steps process to be taken to calculate wavelet coefficient</u>

Step 1

Take a wavelet and compared it to a part at the start of the original signal

Step 2

Calculate a number, C, that represent how closely the signal is correlated with the wavelet. The higher C, is larger the similarity. The result will depend on the shape of the wavelet.

Step 3

Shift the wavelet to the right and repeats step 1-2 until the whole signal is covered.

Step 4

Scale(stretch) the wavele and repeats the steps 1-3

Step 5

Repeat step 1-4 for all scale

There are two types of wavelet transform.

1 Continuous wavelet transform

2 Discrete wavelet transform

Here we have used only continuous wavelet transform.

There are various wavelet families

1. Haar wavelet
2. Daubechies wavelet
3. Symlet wavelet
4. Coeflets wavelet
5. Biorthogonal wavelet
6. Revers biorthogomal wavelet
7. Meyer wavelet
8. Gaussian wavelet
9. Morlet wavelet
10. Complex Morlet wavelet

2.6 Haar Wavelet

Haar wavelet is basis wavelet of wavelet families. Haar wavelet is discontinuous. It is similar to a step function. It is defined as wavelet ψ and scaling function ϕ.

$$\psi(x) = \begin{cases} \mathbf{1} & \text{when } 0 < x < \frac{1}{2} \\ \mathbf{1} & \text{when } 0 < x < \frac{1}{2} \\ 0 & \text{Otherwise} \end{cases} \quad (2.6.1)$$

$$\Phi(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & otherwise \end{cases} \quad (2.6.2)$$

The shifting by b and scaling by a, gives

$$\psi_{a,b}(x) = 2^{a/2}\,\Psi(2^a x - b) \quad b = 0,1,2\ldots.. \quad a = 0,1,\ldots\ldots 2^i - 1. \quad (2.6.3)$$

$$\phi_{a,b}(x) = 2^{a/2}\phi(2^a x - b) \quad b = 0,1,2\ldots.. \quad a = 0,1,\ldots\ldots 2^i - 1. \quad (2.6.4)$$

Coefficient corresponding to $\psi_{a,b}$ are average or sum coefficients

Coefficient corresponding to $\phi_{a,b}$ are difference or detailed coefficients

2.7 Morlet Wavelet

The Morlet wavelet is also called Gabor wavelet is given by

$$\psi(t) = e^{-\beta\frac{t^2}{2}} e^{j\omega_o t} \quad (2.7.1)$$

The value $\beta = \omega_o^2$ and ω_o are defined in such a way so that admissibility condition is satisfied. The function is called Morlet wavelet. This wavelet is used for the time frequency analysis of acoustic signals.

Using wavelet transform we obtain the scalogram defined as

$$|w_\psi^s(a,b)|^2$$

The scalogram is a measure of the energy distribution over time shift b and scaling factor a of the signal. It holds that the energy E_s of a signal s is

$$E_S = \iint_{-\infty}^{\infty} |w_\psi^s(a,b)|^2 \; df db \quad (2.7.2)$$

It is possible to divide this total energy into an energy density over time and over frequency. This is achieved by one integration over frequency or time. The energy density over time is defined by

$$E(b) = \int_{-\infty}^{\infty} \left|w_\psi^s(a,b)\right|^2 df \quad (2.7.3)$$

The energy density over frequency or the energy density spectrum is defined by

$$E(f) = \int_{-\infty}^{\infty} |w_\psi^s(a,b)|^2 \, db \quad (2.7.4)$$

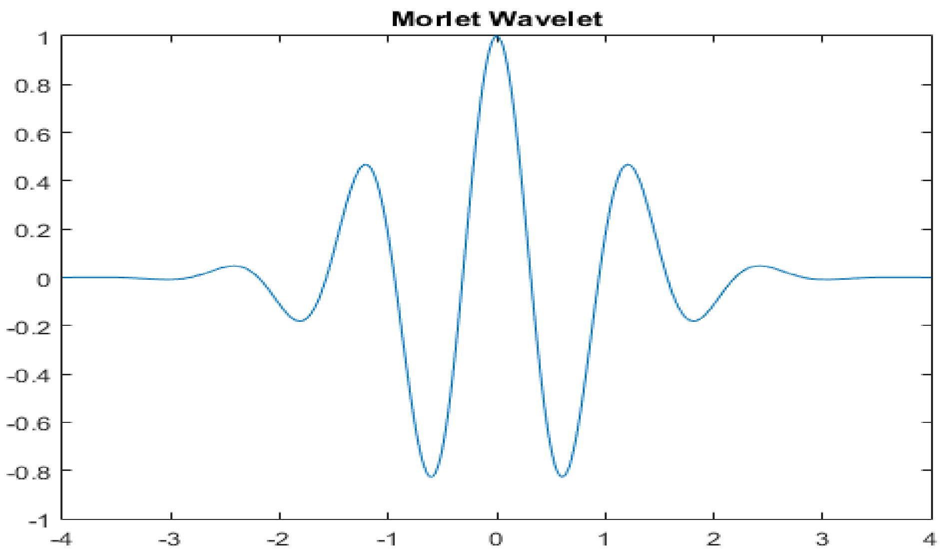

Figure.1.2.2

2.8 Application of Wavelet Transforms

There are various applications of wavelet transforms

• Signal processing

• Data compression

• Smoothing and image denoising

• Fingerprint verification

• Biology for cell membrane recognition, to distinguish the normal from the pathological membranes

• DNA analysis, protein analysis

• Blood-pressure, heart-rate and ECG analyses

• Finance for detecting the properties of quick variation of values

• In Internet traffic description, for designing the services size

• Industrial supervision of gear-wheel

• Speech recognition

• Computer graphics and multifractal analysis

• Many areas of physics have seen this paradigm shift, including molecular

Dynamics, astrophysics, optics, turbulence and quantum mechanics. Wavelets have been used successfully in other areas of geophysical study. Orthonormal wavelets, for instance, have been applied to the study of atmospheric layer turbulence.

Chapter-3

Analytic Signal Using Hilbert Transform And Wavelet Transforms

3.1 Hilbert Transform

Hilbert transform was introduced by Gabor. It is used for time-frequency analysis in different applications of signal processing. It gives analytic signal. Hilbert transform of a real-valued signal x (t) is defined as:

$$\tilde{x}(t) \triangleq H[x](\mathrm{t}) = \frac{1}{\pi}\int \frac{x(\tau)}{t-\tau} d\tau \tag{3.1.1}$$

This $\tilde{x}(t)$ is the convolution of $x(t)$ with $\frac{1}{\pi t}$ over $R \triangleq (-\infty, \infty)$ which is represented in the Fourier domain as

$$\hat{\tilde{x}}(\omega) = -i\; sgn(\omega)\hat{x}(\omega) \tag{3.1.2}$$

Where $\hat{x}(\omega) \triangleq F[x](\omega)$

And

$$sgn(\omega) \triangleq \begin{bmatrix} +1 & \omega > 0 \\ -1 & \omega < 0 \end{bmatrix}$$

Given the Hilbert transform of a real-valued signal $x(t)$, the complex-valued analytic signal is defined as

$$X(t) \triangleq x(t) + i\tilde{x}(t) \tag{3.1.3}$$

3.2 The Discrete-Time Hilbert Transform

The Hilbert transform introduces a 90-degree phase shift to all sinusoidal components. In the discrete-time periodic-frequency domain, the transfer function of Hilbert transform is defined as,

$$\mathrm{H(j}\omega) = \begin{Bmatrix} -\mathrm{j}, & 0 < \omega < \pi \\ \mathrm{j}, & -\pi < \omega < 0 \end{Bmatrix} \tag{3.2.1}$$

The convolution kernel for H(jω) can be calculated through the inverse Fourier transform

$$h[n] = \frac{1}{2\pi}\int_{-\pi}^{\pi} H(j\omega)e^{j\omega n}d\omega. \quad (3.2.2)$$

$$= \begin{cases} \frac{1}{2\pi}\frac{sin^2(\pi n)}{n} & n \neq 0 \\ 0, & n = 0 \end{cases} \quad (3.2.3)$$

h[n] has a infinite support from n= $-\infty$ *to* ∞. In practice, the entire function cannot be stored digitally.

3.3 Application of Hilbert Transforms

1. Propagation time estimation
2. Decay rate estimation
3. To estimate envelope
4. For estimation of instantaneous frequency and phase

3.4 Analytic Signal Using Wavelet Transform

In this section we present the wavelet transform method to obtain an analytic signal It was presented by Gao [9].

Given an analytic wavelet function $g(t)$ and its Fourier transforms $g(\omega)$ satisfying

$$g(t) \in L^1(R,dt) \cap L^2(R,dt) \quad (3.4.1)$$

$$g(\omega) \in L^1(R\setminus\{0\}, d\omega/|\omega|) \cap L^2(R\setminus\{0\}, d\omega/|\omega|) \quad (3.4.2)$$

For a given signal $s(t) \in L^2(R,dt)$ the wavelet transform of $s(t)$ with respect to wavelet $g(t)$ is defined as

$$S(b,a) = \frac{1}{a}\int_{-\infty}^{\infty} s(t)\overline{g}\left(\frac{t-b}{a}\right)dt \quad (3.4.3)$$

Where $t, b \in R, R$ the real number is set and $a > 0, \overline{g}(t)$ is complex conjugate of $g(t)$.

Theorem 1: If $s(t)$ is a signal with finite energy and $g(t)$ is an analytic wavelet function, the $S(b,a)$, the wavelet transform $s(t)$, with respect to $g(t)$ is a complex function with respect to real valued variable b and the scale factor $a(a > 0)$. For a fixed value of a ,the imaginary part of this complex function is the Hilbert transform of the real part i.e., $S(b,a)$ is an analytic function with respect to b .

Theorem 2: If $g(t)$ is an analytic wavelet function with its real part even and

$$C_g = \sum_{0}^{\infty} (gR(\omega) / \omega) d\omega, \tag{3.4.4}$$

With $0 < C_g < \infty$.

For an arbitrary real $s(t) \in L^2(R, dt)$,

$$\frac{1}{C_g} \int_{0}^{\infty} S(t,a) \frac{da}{a} = s(t) + jH[s(t)] \tag{3.4.5}$$

Where $S(t,a)$ is defined as given in equation (3.4.3) and $H[s(t)]$ is the Hilbert transform of $s(t)$. We have

$$\begin{aligned} S(t,a) &= \frac{1}{a} \int_{-\infty}^{\infty} s(b) \overline{g}\left(\frac{b-t}{a}\right) db \\ &= S(t,a) = \frac{1}{a} \int_{-\infty}^{\infty} s(b) \left[g_R\left(\frac{b-t}{a}\right) - jg_i\left(\frac{b-t}{a}\right) \right] db \\ &= \frac{1}{a} \int_{-\infty}^{\infty} s(b) g_R\left(\frac{b-t}{a}\right) db - j\frac{1}{a} \int_{-\infty}^{\infty} s(b) g_i\left(\frac{b-t}{a}\right) db \\ &= S_R(t,a) + jS_i(t,a) \end{aligned} \tag{3.4.6}$$

Where

$$\begin{aligned} g_R(t) &= \operatorname{Re}(g(t)), g_I(t) = \operatorname{Im}(g(t)) \\ S_R(t,a) &= \frac{1}{a} \int_{-\infty}^{\infty} s(b) g_R\left(\frac{b-t}{a}\right) db \\ S_I(t,a) &= \frac{-1}{a} \int_{-\infty}^{\infty} s(b) g_I\left(\frac{b-t}{a}\right) db \end{aligned} \tag{3.4.7}$$

From theorem 1 for an arbitrary $a \in R, a > 0,$ we can get

$$\begin{aligned} S(t,a) &= S_R(t,a) + jS_I(t,a) \\ &= S_R(t,a) + jH[S_I(t,a)] \\ &= \frac{1}{a} \int_{-\infty}^{\infty} s(b) g_R\left(\frac{b-t}{a}\right) db + jH\left[\frac{1}{a} \int_{-\infty}^{\infty} s(b) g_R\left(\frac{b-t}{a}\right) db \right] \end{aligned} \tag{3.4.8}$$

Both side of equation (3.4.8) multiplied by $1/(C_g a)$ and then by integrating it with respect to a, we have

$$\frac{1}{C_g}\int_0^{\infty} S(t,a)\frac{da}{a}$$

$$= \frac{1}{C_g}\int_0^{\infty}\frac{1}{a^2}\int_{-\infty}^{\infty} S(b) g_R\left(\frac{b-t}{a}\right) db da \qquad (3.4.9)$$

$$= s(t)$$

The Morlet wavelet is

$$g(t) = e^{jmt} e^{-t^2/2} \; (m > 6) \qquad (3.4.20)$$

It is easy to verify that $g(t)$ satisfy the theorem 2 in numerically approximate sense [5]

The modified Morlet Wavelet is given as

$$g_r(t) = e^{imt} e^{-(1/2)\left[\sqrt{2}\sigma m/(2\pi\tau)t\right]^2}$$
$$= Cos(mt) e^{-(1/2)\left[\sqrt{2}\sigma m/(2\pi\tau)t\right]^2} + j\sin(mt) e^{-(1/2)\left[\sqrt{2}\sigma m/(2\pi\tau)t\right]^2} \qquad (3.4.21)$$

Let $C = \sqrt{2}\sigma m/(2\pi\tau)$ In equation (3.4.20) and (3.4.21) m is angular frequency, τ is the number of cycles of the carrier wave in an envelope, and σ is a real number. When $g(t)$ is a wavelet, the analytic counter part of a real valued signal can be calculated without cut-off error and it will give more accurate results than the Hilbert Transform. The frequency dependence of the phase is very less as compared to the maximum entropy method. The wavelet given in above equations (3.4.20) and (3.4.21) are not compactly supported, but their amplitude decay at the rate of $e^{-(1/2)t^2}$ so the equation (3.4.15) gives the more accurate result than Hilbert transform.

Using the analytic signal, the instantaneous parameters are defined as

1 Instantaneous amplitude e(t)

$$e(t) = \sqrt{s^2(t) + H^2[s(t)]} \qquad (3.4.22)$$

2 Instantaneous phase

$$\theta(t) = \arctan\left(\frac{H[s(t)]}{s(t)}\right) \tag{3.4.23}$$

3 Instantaneous frequency

$$f(t) = \frac{1}{2\pi}\frac{d}{dt}\left[\arctan\left(\frac{H[s(t)]}{s(t)}\right)\right] \tag{3.4.24}$$

Chapter-4

Empirical Mode Decomposition Algorithm

EMD filters out functions or signals that form a complete and nearly orthogonal basis for the original signal. Completeness depends upon on the method of the EMD [10]-[11]. i.e the way it is decomposed implies completeness. The decomposed functions are known as intrinsic mode function (IMFs). The decomposed functions are sufficient to describe the signal, even though they are not necessarily orthogonal.

The functions in to which a signal is decomposed are all in the time domain and of the same length as the original signal. These functions allow for varying frequency in time to be preserved. Obtaining IMFs from real world signals is important as the natural processes often have multiple causes, and each of these causes may happen at specific time. This type of data is evident in an EMD analysis, but hidden in the Fourier domain or in wavelet coefficients. The EMD method required to reduce any given data in to a collection of intrinsic mode function (IMF) to which further analysis can be applied.

The principle of EMD is to extract oscillatory modes existing in time scales defined by the interval between local extrema. A local maxima and minima point is any point on the signal where its derivative is zero, and its second derivative non-zero. The term local is used to differentiate it from a global maximum and minimum point. There may be several local maxima and minima within an observed window, however only one global maximum and global minimum point may be present. Once the time scales are identified, IMFs with zero mean are sifted out of the signal.

IMF represents a simple oscillatory mode as a counterpart to the simple harmonic function, but it is much more general: instead of constant amplitude and frequency in a simple harmonic component, an IMF can have variable amplitude and frequency along the time axis. The procedure of extracting an IMF is called sifting. The sifting process is as follows:

1. Identify all the local extrema in the test data.
2. Connect all the local maxima by a cubic spline line as the upper envelope.
3. Repeat the procedure for the local minima to produce the lower envelope.

The upper and lower envelopes should cover all the data between them. Their mean is m_1. The difference between the data and m_1 is the first component h_1:

$$X(t) - m_1 = h_{1.} \qquad (4.a)$$

Ideally, h_1 should satisfy the definition of an IMF, since the construction of h_1 should have made it symmetric and having all maxima positive and all minima negative. After the first round of sifting, a crest may become local maxima. New extrema generated in this way shows the proper modes lost in the initial examination. In the subsequent sifting process, h_1 can only be treated as a proto-IMF. In the next step, h_1 is treated as data:

$$h_{1.} - m_{11.} = h_{11.} \qquad (4.b)$$

After repeated sifting up to k times, h_1 becomes an IMF, that is

$$h_{1.(k-1)} - m_{1k.} = h_{1k.} \qquad (4.c)$$

Then, h_{1k} is designated as the first IMF component of the data:

$$c_{1.} = h_{1k.} \qquad (4.d)$$

4.1 Intrinsic Mode Function (IMF)

An intrinsic mode function (IMF) is a function that satisfies two conditions:

1. In the whole data set, the number of extrema and the number of zero crossings must either equal or differ at most by one.

2. At any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero.

The first condition is similar to the traditional narrow band requirements for a stationary Gaussian process. The second condition is necessary so that the instantaneous frequency will not have the unwanted fluctuations induced by asymmetric wave forms. For non-stationary data, the ‘local mean’ involves a ‘local time scale’ to compute the mean, which is impossible to define. So, the local mean of the envelopes defined by the local maxima are used and the local minima is used to force the local symmetry. This is a necessary step to avoid the definition of a local averaging time scale. It introduces an alias in the instantaneous frequency for nonlinearly deformed waves, the effects of nonlinearity are much weaker in comparison with non-stationary.

4.2 Empirical Mode Decomposition (EMD) Steps:-

The steps of EMD method [12] are as follows:

1. Determine local maxima and minima of generated signal, s(t),

2. Implement cubic spline interpolation between the maxima and the minima to Obtain the envelopes e_{max} ande_{min}, respectively,

3. Estimate the mean value of envelope, M(t) $= \frac{e_{max} + e_{min}}{2}$

4. Substrat mean value from original signal $C_1 = s(t) - M(t)$,

5. $C_1(t)$ is an intrinsic mode function(IMF) if the number of local maxima or local minima of $C_1(t)$, is equal to or differs from the number of zero crossings by one, and the average value of $C_1(t)$ as much as equal to zero. If $C_1(t)$ is not an Intrinsic mode function (IMF) , then repeat steps 1-4 on $C_1(t)$ instead of s(t), until the new $C_1(t)$ obtained satisfies the conditions of an IMF,

6. Compute the residue, $R_1(t) = S(t) - C_1(t)$,

7. If the residue, $R_1(t)$, is above a threshold value of error tolerance, then repeat steps 1-6 on $R_1(t)$, to obtain the next intrinsic mode function and a new residue.

An appropriate stopping criterion, in step 5, avoids ‘over-modeling’ in $C_1(t)$ to avoid significant loss of information. The first intrinsic mode function consists of the highest frequency components present in the original signal. The successive intrinsic mode function (IMFs) contains continuously lower frequency components of the signal. [13] To guarantee that the IMF components retain enough amplitude and frequency modulations, we have to determine a criterion for the sifting process to stop. This is performed by limiting the size of the standard deviation, SD, computed from the two consecutive sifting results as

$$SD = \sum_0^{N-1}\left[\frac{|(S_{1(k-1)}(t)-C_{1k}(t)|^2}{S_{1(k-1)}^2(t)}\right]$$

A typical valued of standard deviation is varied between $0.2 < SD < 0.3$, N is number of sample. If n orthogonal intrinsic mode function is obtained in this iterative manner, the original signal may be reconstructed as,

$$S(t) = \sum_n C_i + R(t)$$

The final residue contains general trends of the original signal.

Chapter-5

Classification

Classification is some decision or forecast made on the basis of currently available information, and a classification procedure is then some formal method for repeatedly making such judgments in new situations. The problem concerns the construction of a procedure that will be applied to a continuing sequence of cases, in which each new case must be assigned to one of a set of pre-defined classes on the basis of observed attributes or features. The construction of a classification procedure from a set of data for which the true classes are known has been also termed pattern recognition [14].

5.1 Perspectives on Classification

A wide variety of approaches has been taken towards this task. Three main historical strands of research can be identified: statistical, machine learning and neural network.

5.1.1 Statistical Approaches

Two main phases of work on classification can be identified within the statistical approaches. The first, "classical" phase concentrated on derivatives of Fisher's early work on linear discrimination. The second, "modern" phase exploits more flexible classes of models, many of which attempt to provide an estimate of the joint distribution of the features within each class, which in turn provide a classification rule. Statistical approaches are generally characterised by probability model, which provides a probability of being in each class rather than simply a classification.

5.1.2 Machine Learning

Machine Learning is used to encompass automatic computing procedures based on logical or binary operations that learn a task from a series of examples. Decision-tree based approaches are used in this type of learning, in which classification results from a sequence of logical steps. These are capable of representing the most complex problem given sufficient. Other techniques, such as genetic algorithms and inductive logic procedures (ILP), are also used. This allow us to deal with more general types of data, including cases where the number and type of attributes may vary, and where additional layers of learning are superimposed, with hierarchical structure of attributes and classes and so on. Machine Learning aims to generate classifying expressions simple enough to provide insight into the decision process.

5.1.3 Neural Networks

In neural networks, the excitement of technological progress is supplemented by the challenge of reproducing intelligence itself. Neural networks consist of layers of interconnected nodes, each node producing a non-linear function of its input [15]. The input to a node may come from other nodes or directly from the input data. Also, some nodes are identified with the output of the network. The complete network therefore represents a very complex set of interdependencies which may incorporate any degree of nonlinearity, allowing very general functions to be modelled. In the simplest networks, the output from one node is fed into another node in such a way as to propagate "messages" through layers of interconnecting nodes. More complex behaviour may be modelled by networks in which the final output nodes are connected with earlier nodes, and then the system has the characteristics of a highly nonlinear system with feedback. Neural network approaches combine the complexity of some of the statistical techniques with the machine learning objective of imitating human intelligence.

There are many issues of concern to the classifier. Some of them are:-

a Accuracy: There is the reliability of the rule, usually represented by the proportion of correct classifications

b Speed In some circumstances, the speed of the classifier is a major issue.

c Comprehensibility If it is a human operator that must apply the classification procedure, the

Procedure must be easily understood else mistakes will be made in applying the rule.

Time to Learn Especially in a rapidly changing environment, it may be necessary to learn a classification rule quickly.

5.2 Linear Discriminant Analysis

The objective of LDA is to perform dimensionality reduction while preserving as much of the class discriminatory information as possible.

There are many possible techniques for classification of data. Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) are two commonly used techniques for data classification and dimensionality reduction. Linear Discriminant Analysis easily handles the case where the within-class frequencies are unequal and their performances have been examined on randomly generated test data. This method maximizes the ratio of between-class variance to the within-class variance in any particular data set thereby guaranteeing maximal separability. The use of Linear Discriminant Analysis for data classification is used for classification problem in speech recognition. The main difference between LDA and PCA is that PCA does more of feature classification and LDA does data classification. In PCA, the shape and location of the original data sets changes when transformed to a different space whereas LDA doesn't change the location but only tries to provide more class separability and draw a decision region between the given classes. This method also helps to better understand the distribution of the the feature data.

5.3 Different Approaches to LDA

Data sets can be transformed and test vectors can be classified in the transformed space by two different approaches [16]. Class-dependent transformation: This type of approach involves maximizing the ratio of between class variance to within class variance. The main objective is to maximize this ratio so that adequate class separability is obtained. The class-specific type approach involves using two optimizing criteria for transforming the data sets independently. Class-independent transformation: This approach involves maximizing the ratio of overall variance to within class variance. This approach uses only one optimizing criterion to transform the data sets and hence all data points irrespective of their class identity are transformed using this transform. In this type of LDA, each class is considered as a separate class against all other classes.

5.4 K-Nearest Neighbor Classifier

The K-NN algorithm [17] is a nonparametric method. In this, the proximity of neighbouring Input (x) observations in the training data set and their corresponding output values (y) are used to predict (score) the output values of cases in the validation data set. This output variables can either be interval variables in which case the K-NN algorithm is used for prediction while if the output variables are categorical, either nominal or ordinal, the K- NN algorithm is used for classification purposes. Before getting into details of the K-NN algorithm we need to define the Euclidean distance between two input vectors Xr and Xs. Consider the two input variable case since it is easy to represent in two-dimensional space. The two vectors Xr =(Xr1, Xr2) and Xs = (Xs1, Xs2) represented in Figure below. The distance between these two vectors is computed as the length of the difference vector Xr - Xs, denoted by

$$d(X_r , X_s) = |X_r - X_s| = \sqrt{(X_{r1} - X_{r2})^2 - (X_{s1} - X_{s2})^2}$$

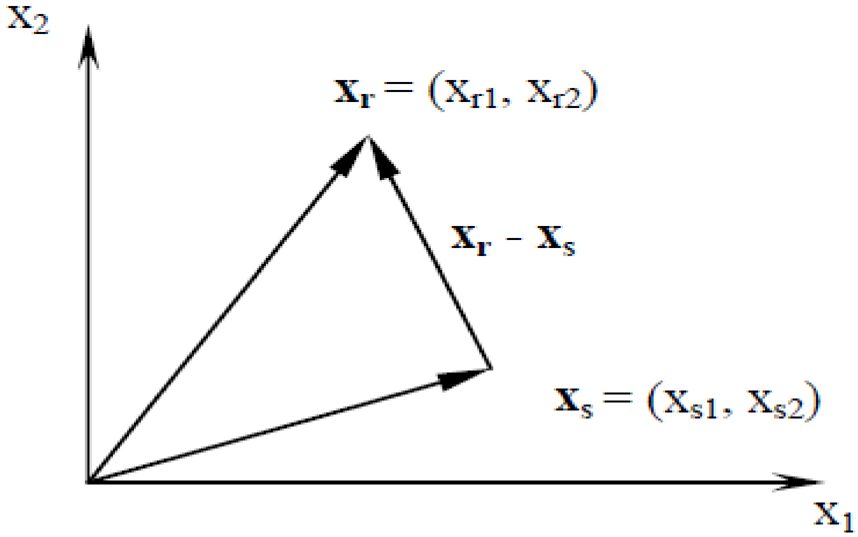

Figure 5.4

More generally the distance between two p-dimensional vectors u = (u1, u2,…,up) and v = (v1, v2,…,vp) is calculated as

$$d(u,v) = |u-v| = \sqrt{(u_1 - v_1)^2 + (u_2 - u_2)^2 + \cdots \dots \dots \dots \dots \dots .. (u_p - v_p)^2}$$

Given that the K-NN method is dependent on distance measures, it is recommended that the input data be standardized first before proceeding to build an optimal K-NN model. The minimum distance between the vectors gives the closest neighbour so it is predicted that it belongs to the same class with the test object.

Chapter-6

Implementation and Results

We generated four signals Sag, Swell, Harmonic and Transient signal in MATLAB, having frequency 3.2 kHz. Then we applied EMD to find IMFs of these signals. We used three IMFs for feature extraction. Then wavelet transform method is implemented to obtain the analytic signal Using these analytic signals, three instantaneous parameters have been calculated. They are energy, standard deviation of the contour, standard deviation of phase contour There are total nine features for an signal because three IMF are taken for one signal and we have extracted three features for one IMF.

Signal which are generated in MAT lab given below

$$C_1 - Sag$$

$$C_2 - Swell$$

$$C_3 - Harmonic$$

$$C_4 - Transient$$

1. Sag signal

```
t1=0.05; t2=0.1;
v =(1-A*(u(t-t1)-u(t-t2))).*sin(2*pi*50*t);
```

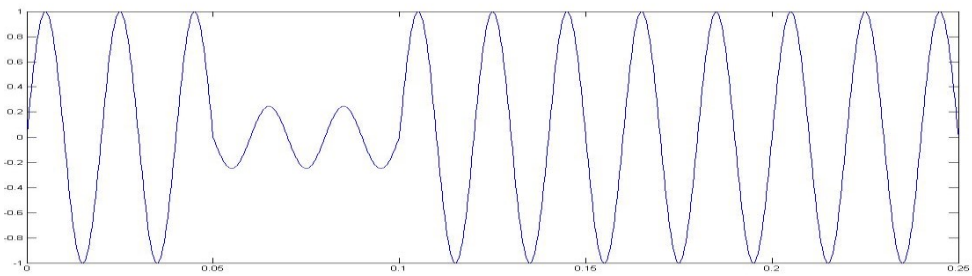

Figure.6.1

2. Swell signal

```
t1=0.05; t2=0.1;
x =(1+A*(u(t-t1) -u(t-t2) )) .*sin(2*pi*50*t) ;
```

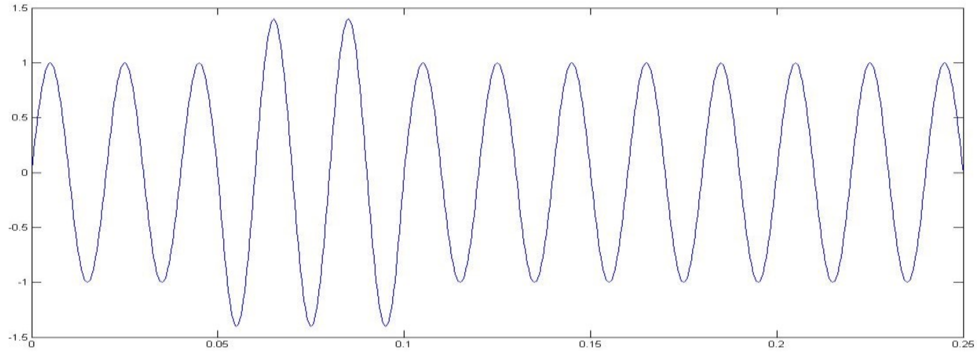

Figure 6.2

2. Harmonic signal

```
fs=3.2*10^3;
t=0:(1/fs):0.25;
a = 0.05;
b = 0.15;
r2 = (b-a).*rand(1000,1) + a;
r3 = (b-a).*rand(1000,1) + a;
r5 = (b-a).*rand(1000,1) + a;
r7 = ones(1000,1)-(r2.^2 + r3.^2 + r5.^2 );
y =
(asin(2*pi*5t)+bsin(3*2*pi*t)+csin(5*2*pi*5t)+dsin(7*2*pi*50*t
));
```

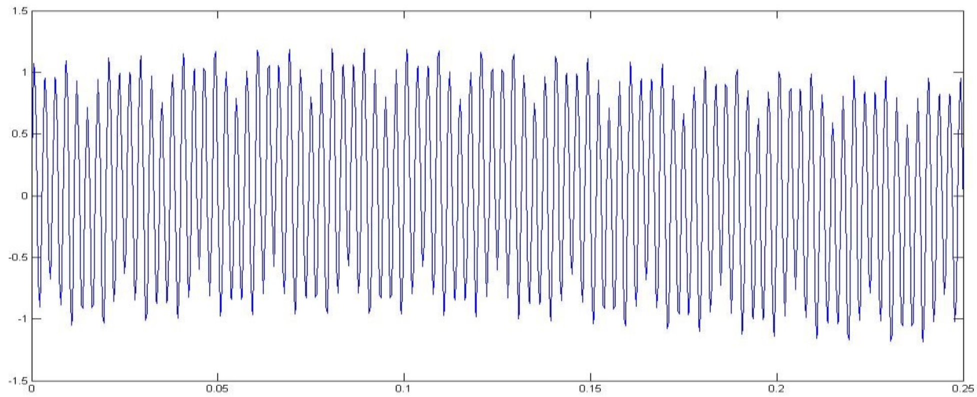

Figure 6.3

4. Transient signal

```
fs=3.2*10^3;
t=0:(1/fs):0.25;
K=0.7;
s=0.0015;
```

```
    a = 900;
    b = 1300;
    r = (b-a).*rand(1000,1) + a;
   t1=0.05; t2=0.1;
 delt=(t-t1);
term3 =(u(t-t2) -u(t-t1) ).*(cos(2*pi*A*delt) );
    z=(cos(2*pi*50*t)+K*exp(-delt/s).*term3 );
```

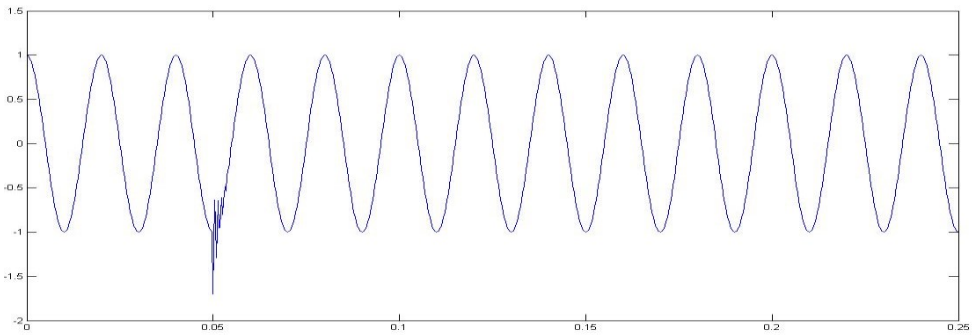

Figure 6.4

The extracted features obtained by taking wavelet transform of IMFs have been used for PQ events classification. We used k-NN classifier in conjunction with LDA algorithm. We calculted the classifier efficiency and compared the results with feature extraction using the Hilbert transform method. Tabel I and II shows the classification using feature extraction by Hilbert transform and Wavelet transform methods respectively.

TABLE –I

6.1 Classification result using EMD and Hilbert Transform

CASE	C_1	C_2	C_3	C_4
C_1	93			2
C_2	3	100		2
C_3			100	
C_4	4			96
Classification Efficiency in %	9	100	100	96
Error Efficiency in %	7	0	0	4
Overall Efficiency in %	97.25%			

$$\text{Overall Efficiency} = \frac{Number\ of\ events\ classified\ correctely}{Total\ number\ of\ event}$$

(Considering all the classes)

TABLE –II

6.2 Classification result using EMD and Wavelet Transform method

CASE	C_1	C_2	C_3	C_4
C_1	93			2
C_2	3	100		2
C_3			100	
C_4	4			96
Classification Efficiency in %	93	100	100	96
Error Efficiency in %	7	0	0	4
Overall Efficiency in %	97.25%			

A maximum classification accuracy of 97.25 % in both the cases i.e. 97.25%.

Chapter-7

Conclusion and Future Work

Feature extraction and instantaneous parameter calculation using wavelet transform method is important as wavelet transform is useful for nonstationary signals. Due to time frequency localization property of wavelet transform it helps in antinoise performance.

The method can be improved using different types of complex wavelets. We can use wavelet transform based compression algorithms to reduce the data size. Also we can use different types of denoising algorithms.

Reference

[1] Stuti Shukla, S. Mishra, Senior Member, IEEE, and Bhim Singh, Senior Member, IEEE "Empirical-Mode Decomposition with Hilbert Transform for Power-Quality Assessment" *IEEE Transaction on power delivery, VOL. 24, NO.4, OCTOBER 2009.*

[2] N. S. Tunaboylu and E. R Collins, "The wavelet transform approach to detect and Quantify voltage sags," Proceedings of the 7th *International Conference on Harmonics and Quality of Power (ICHQP) Las Vegas, Nevada, USA, Octok 16-18.1996, pp. 619- 623.*

[3] José C. C. Costa 1, Ricardo L. Pregitzer 1, Tiago N. Sousa 1, José Batista 2, João L. Afonso 1 "A Case of Power Quality Assessment Using a Development Power Quality Monitor" *CEE'05 – IEEE 1st International Conference on Electrical Engineering Coimbra, Portugal, 10-12 October 2005.*

[4] Direction des Etudes et Recherches Groupe "Qualite de la fourniture Services Entreprises" Clamart, "New Signal Processing Tools Applied to Power Quality Analysis", EDF - *FRANCE IEEE Transactions on Power Delivery, Vol. 11, No. 2, April 1996.*

[5] M. V. Chilukuri, Member, IEEE, and P. K. Dash, Senior Member, IEEE, "Multiresolution S-Transform-Based Fuzzy Recognition System for Power Quality Events" *IEEE TRANSACTIONS ON POWER DELIVERY, VOL. 19, NO. 1, JANUARY 2004.*

[6] Ameen M. Gargoom, Student Member, IEEE, Nesimi Ertugrul, Member,IEEE and Wen L.Soong "Investigation of Effective Automatic Recognition Systems of Power-Quality events" *IEEE TRANSACTIONS ON POWER DELIVERY, VOL. 22. NO. 4, OCTOBER 2007*

[7] S. Mishra, Senior Member, IEEE, C. N. Bhende, and B. K. Panigrahi, Senior Member, IEEE, "Detection and Classification of Power Quality Disturbances Using S-Transform Probabilistic Neural Network", *IEEE TRANSACTIONS ON POWER DELIVERY, VOL. 23, NO. 1, JANUARY 2008.*

[8] Raghuveer M.Rao. Ajit S.Bhopardikar "Wavelet transforms, introduction to theory and applications" Har/Dskt ed. Pearson Education.

[9] Jinghuai Gao, Xiaolong, Dong, Wen-Bing Wang, Youming Li, and Cunhuan Pan "Instantaneous Parameters Extraction via Wavelet Transform" *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, VOL. 37,* NO.2 MARCH 1999.

[10] Patrick Flandrin*, Fellow, IEEE*, Gabriel Rilling, and Paulo Conclaves, "Empirical Mode Decomposition as a Filter Bank" *IEEE SIGNAL PROCESSING LETTERS, VOL. 11, NO. 2, FEBRUARY 2004.*

[11] Nord en E.Huang1, Zheng Shen2, Steven R. Long3,Manli C.W u4, Hsing H. Shih5, Quanan Zheng6, Nai-Chyuan Yen7, Chi Chao Tung8 and Henry H. L iu9 "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non- stationary time series analysis", *Received 3 June 1996; accepted 4 November 1996.*

[12] Stuti Shukla, S. Mishra*, Senior Member, IEEE, and Bhim Singh, Senio Member, IEEE* "Empirical-Mode Decomposition with Hilbert Transform for Power- Quality Assessment" *IEEE Transaction on power delivery, VOL. 21, NO. 4, OCTOBER 2006.*

[13] Md. Khademul Islam Molla, Mostafa Al Masum Shaikh1 and Keikichi "Time-Frequency Representation of Audio Signals Using Hilbert Spectrum with Effective

Frequency Scaling", Hirose *Proceedings of 11th International Conference on Computer and Information Technology (ICCIT 2008) 25-27 December, 2008, Khulna, Bangladesh.*

[14] D. Michie, D.J. Spiegelhalter, C.C. Taylor (eds), "Machine Learning, Neural and Statistical Classification". *ACM SIGAR bulletin VOLUME 7 issue 1, jan.1996*

[15] N.B Karayiannis, A. Venetsanopoulos. "Neural Network Architectures and Learning Schemes Artificial neural networks". *The Springer International Series in Engineering and Computer Science Volume 209, 1993 pp 9-85.*

[16] Normalized LDA for Semi-supervised Learning in Fan Zhen Lei Stan Z. Li Center for Biometrics and Security Research & National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences95 Zhongguancun Donglu, Beijing 100190, China 978-1-4244-2154-1/08/$25.00 _c 2008 IEEE

[17] Classification of Power Quality Disturbances Using Wavelet Transform and K-Nearest Neighbor Classifier Ngo Minh Khoa Department of Technique & Technology Quy Nhon University Quy Nhon, Viet Nam, Dinh Thanh Viet, Nguyen Huu Hieu Department of Electrical Engineering Da Nang University of Science and Technology Da Nang, Vie Nam.